

CAUSALITY IN A MCDOWELLIAN WORLD

Thesis submitted for the degree of Doctor of Philosophy

The Queen's University of Belfast, May 2014

Alan Charles McKay

ABSTRACT

The thesis explores and suggests a solution to a problem that I identify in John McDowell's and Lynne Rudder Baker's approaches to mental and intention-dependent (ID) causation in the physical world.

I begin (chapter 1) with a brief discussion of McDowell's non-reductive and anti-scientistic account of mind and world, which I believe offers, through its vision of the unbounded conceptual and the world as within the space of reasons, to liberate and renew philosophy. However, I find an inconsistency in McDowell's criticism of Davidson's anomalous monism (chapter 2), stemming from a tension between McDowell's naïve common sense view of mental causation and an understanding of ordinary physical causation which I think McDowell ought to embrace, which portrays it as both objective, in the sense of being recognition-independent, and as belonging within the space of reasons. The question of the relation between these two concepts of causation is an aspect of the more general question of the relation between the space of reasons and the realm of law.

In chapter 3 I begin examining the possibility that Baker's account of material and property constitution could form the basis of a bridge between the two spaces, and find it generally promising. However, I find that her defence of her version of non-reductive monism against Kim's causal arguments also runs into problems, which I attribute to the fact that she holds a view of causation as secondary to causal explanation.

In chapter 4 I develop an account of what I call manifest physical causation – of objective causal relations in the world of Sellars’s manifest image. Based upon McDowell’s transcendental empiricism, which takes the existence of the ordinary perceived world as a condition of the possibility of our possession of conceptual capacities, I contend that it is this picture of reality, rather than that of fundamental physics, that should be the starting point of our theorizing. Causation in the manifest image, I argue, covers the behaviour of the familiar physical world as well as that of its well-understood extensions into the special sciences and engineering. Manifest physical causation, on my account, is productive, acts through mechanisms which are almost entirely mechanical, electromagnetic, and/ or chemical, and is causally closed. In my view, normative, semantic, contentful property-instances are not part of the manifest physical causal nexus.

In my final chapter I suggest a modification of Baker’s constitution account, which I call Constituted Causation, whereby higher-level – mental and other ID – causal relations are constituted, in favourable circumstances, by lower-level ones. ID causal relations belong in their own causal nexus but are connected to the manifest physical world through constitution, a relation of unity without identity. Causation and constitution are, respectively, intra- and inter-level relations, and they are non-overlapping. The constituted network of rational and normative relations bears, I believe, striking parallels with McDowell’s view of reality. According to Baker’s view of constitution, the essential properties of constituted entities subsume those of their constituters; extending this to my account enables us to say that the real cause and explanation of someone’s action is that they consciously performed it

rather than that certain manifest causal processes occurred at the lower level.

INDEX

CHAPTER 1. THE MCDOWELLIAN PICTURE.

I. Mind and World.	6
II. The Core Arguments of Mind and World.	13
III. Openness to the World.	34
IV. Conclusion.	52

CHAPTER 2. MCDOWELL, DAVIDSON, AND CAUSALITY.

I. General Causality.	55
II. Reasons and Causes.	62
III. The Space of Reasons and the Physical.	78

CHAPTER 3: BAKER'S CONSTITUTION ACCOUNT.

I. Introduction	107
------------------------	-----

II. The Constitution Account.	111
III. Baker and Higher-Level Causation.	124
IV. Baker's Property-Constitution Account.	140
V. The Distinguishing Features of ID Causation.	159

CHAPTER 4: MANIFEST PHYSICAL AND INTENTION-DEPENDENT CAUSATION.

I. Causation and the Manifest Image.	168
II. Manifest Physical Causation as Production.	198

CHAPTER 5: THE NATURE OF INTENTION-DEPENDENT CAUSATION.

I. ID Causation and Causal Nexuses.	222
II. Constituted Causation.	234
III. Causal Pluralism.	248
IV. ID Causal Relations.	262
V. Constituted Causation and the Space of Reasons.	272
BIBLIOGRAPHY.	285

CHAPTER 1. THE MCDOWELLIAN PICTURE.

I. Mind and World.

(i) Introduction.

Over the last four decades John McDowell has developed and expounded an influential philosophy of mind and world which many find appealing and liberating. Its central component is a construal of human beings as essentially rational and ethical creatures who are nevertheless thoroughly natural. The route to this picture requires that we take as the starting point of inquiry, not the world as described by physical science, but the world as it is revealed to be, according to McDowell, through both the fact that the structure of our thought and language relate to it and the nature of the relation. I begin with *Mind and World* (MW, 1996), which is based on McDowell's 1991 John Locke lectures and is to date his only full-length book. His other publications include four collections of essays and the 1997 Woodbridge Lectures, while several edited collections of essays on his work, with responses by the author, as well as three monographs (de Gaynesford 2004; Thornton 2004; Gaskin 2006) have also been published. His ideas have inspired work in, for example, the philosophy of education (Bakhurst 2011) and natural theology (Wahlberg 2012). McDowell's general methodology is to allay what he calls philosophical anxiety and to dissolve or exorcise apparently intractable problems, not by arguing for one side or the other, but by showing how the assumptions that give rise to them were

misconceived from the outset, and it is this therapeutic, as opposed to constructive, methodology that he applies to the question of our nature. As well as philosophy of mind, the therapeutic project takes him into epistemology, philosophy of language, ethics, and aesthetics, and into deep engagement with several past philosophers, most notably Aristotle, Kant, Frege, Wittgenstein, Sellars and Davidson, while he also acknowledges debts to Hegel, Gadamer, P.F. Strawson, Rorty, and Gareth Evans as well as Putnam and Brandom.

A hallmark of McDowell's philosophy of mind is a radical anti-Cartesian, anti-reductionist, externalist view of mindedness. Many of the arguments of MW have been criticized and challenged (see, for example, Dodd 1995; Bird 1996; M. Williams 1996, 2006; Fodor 1998; Rorty 1998a, 1998b; Wright 1998, 2002; Norris 2000; Bernstein 2002; Brandom 2002; Stroud 2002; O'Shea 2005; Bilgrami 2006; Gaskin 2006; G. Macdonald 2006; Williamson 2007; Haddock 2008; Travis 2008, Margolis 2009). However, in this chapter I concentrate on building a broad picture of his philosophy, keeping critical appraisal to a minimum. My intention is to show how his ideas, most of which were first introduced in earlier work, build gradually into a picture of the relation of mind and world which will support the original form of naturalism that he advocates in the latter part of MW. In subsequent chapters I focus on one aspect of McDowell's thought, his view of causality, and in particular on tensions that I believe arise from the contrast between agent causality on the one hand and causality in the everyday physical world and in the physical special sciences on the other. Accordingly I concentrate on McDowell's philosophies of mind,

perception, and action and of thought, content, and nature, touching only very briefly and occasionally on other aspects of his work.

(ii) *The Myth of the Given and Coherentism.*

Many contemporary philosophers attempt to reconfigure rationality and normativity so as to make them fit into a reductive, scientific account of what it is to be natural. McDowell regards this conception of the natural as too narrow. His approach, by contrast, is to expand the idea of the natural (MW: 109) to include our distinctive features, while maintaining that rationality and normativity are nevertheless radically distinct from the concepts that are appropriate to the narrow conception. This involves a rejection of the dominant materialism which takes the starting point of inquiry to be the world as described by natural science. The approach to the nature of reality begins, rather, with our own presence in the world as rational creatures, and proceeds by examining what follows from that. Reality thus is seen, primarily, not from the scientific perspective but as what, in a non-technical sense, could be called the life-world (Barber 2010). As we shall see, this is far from being a rejection of science. McDowell's view is simply that, despite its enormous power, science delivers only a partial view of reality.

McDowell begins his account of the problems he addresses in MW by introducing a concept and an assumption that he takes from and shares with Wilfrid Sellars. In *Empiricism and the Philosophy of Mind* (EPM: 76) Sellars writes "...in characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or

state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says". McDowell adopts the concept of the logical space of reasons early in MW and enlarges on Sellars' account. He spells out his understanding of the term in 'Knowledge and the Internal' (KI: 408): "The space of reasons is the space within which thought moves, and its topography is that of the rational interconnections between conceptual contents; we might equally speak of the space of concepts". Thus while Sellars is particularly concerned with knowledge, McDowell generalizes the metaphor of the space to include all those contents of our thought and discourse that are characterized by the employment of normative concepts such as meaning, justification, inference, and intention, and that are appropriately expressed using the concepts of the propositional attitudes. These concepts are just those that make up what Davidson (ME: 223) calls the constitutive ideal of rationality (McDowell STE: 44), although as we shall see shortly, McDowell and Davidson differ profoundly on the extent of the logical space that contains those concepts.

The space of reasons is thus the locus of a particular kind of explanation or understanding, which is contrasted with another kind. In a passage that I will refer back to often in this dissertation, McDowell writes,

(T)he concepts of the propositional attitudes have their proper home in explanations of a special sort: explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be. This is to be contrasted with a style of explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen (FAM: 328).

Normative, space of reasons explanations are contrasted, then, with those that apply paradigmatically in physical science but also in the description and explanation of the behaviour of everyday physical objects (in this and the next two chapters I use “physical” loosely for either of these senses and in contrast to “mental”; in chapter 4 I develop a more precise meaning). What is meant by the expression “how things generally tend to happen”, is just that the explanations, belonging to what McDowell calls the realm of laws, do not make appeal to rational or normative factors. It is important to be clear that all *explanations*, including scientific explanations, are, *qua* explanations, rational, and belong in the space of reasons, while it is in their contents that the contrast between the styles of explanation is found. As McDowell writes (MW: 70, n1); “Of course depictions of nature are linked by relations of justification. The point is that there are no such linkages in what is depicted”. I return to this contrast below (II:(iv)).

It is significant that McDowell calls the logical space that contrasts with the space of reasons the realm of *law* rather than of *nature*, reflecting his insistence that both the space of reasons and our ability to respond to its particular form of intelligibility are part of nature. In the course of a normal upbringing and enculturation (“*Bildung*”), McDowell argues (MW: 84ff), we acquire the ability to respond to the intelligibility and demands of the space of reasons, “the kind of intelligibility that is proper to meaning” (MW: 71). McDowell calls our possession of this ability “second nature”, a generalization from Aristotle’s idea that ethical character is acquired when we become aware of the existence of ethical requirements that are already present in the world for those with the ability to see them. He insists that, like the ethical demands in Aristotle’s

account, the conceptual order – the logical space of reasons – is autonomous within its own sphere; the capacities that belong to it therefore require no justification or explanation that involves reference to ways of thinking that do not belong in that space – that involve, in other words, looking at that space from ‘sideways-on’– from a point of view that “places reality outside a boundary enclosing the conceptual” (MW: 82).

Sellars and McDowell share this view of the irreducibility of the normativity of the logical space of reasons. The kinds of concepts and facts appropriate to it, cannot, without loss of essential content, be described in language appropriate to the realm of scientific laws. McDowell coins the term *bald naturalism* for approaches that attempt to re-characterize the space of reasons in terms that do not make use of normative concepts. He does not deny that it is *possible* to give a bald naturalistic – physical or physiological, say – account of the basis of our intentionality. However, providing such an account will not help us with the philosophical problem of how our intentionality *as such* relates to the world. To mark the contrast between bald naturalism and his own approach, McDowell (NPM: 262) introduced the term ‘liberal naturalism’, which has since been gaining currency (see de Caro and Macarthur 2010: 1-19).

McDowell begins by drawing attention to a source of philosophical anxiety - the problem of how empirical content can be possible. We take ourselves, from a pre-philosophical position, to be open to the world, in a sense which paradigmatically involves our thinking, including our empirical beliefs and judgments, being constrained by how things are in

the world, and yet, when we philosophically analyze our relation to the world in a certain way, it appears that it cannot be the case that we are under such constraint. The arguments that lead to this conclusion are closely related to those directed by Sellars and others against epistemological foundationalism. For the classical foundationalist, constraint by the world is unproblematic in that we have basic, non-inferential justified beliefs deriving from sensations, and we inferentially justify all other beliefs on the basis of these and, ultimately, these alone. In EPM, Sellars exposes the weakness of this position, arguing that sensations, being non-epistemic, are not the kind of items that could be held to have a justificatory role. The belief that they can do so is a version of ‘the myth of the given’, a term introduced by Sellars (EPM: 33ff) and taken up with approval by McDowell in MW. ‘The given’ here refers to a bare presence of something to the senses, whether this is thought of as the presence of something indirect like a sense datum or, more directly, of something in the world itself, such as an object. Those who accept Sellars’ conclusions are thereby faced with the problem that, however one construes perceptual experience, our perceptually acquired beliefs cannot be rationally based upon, but are at best *caused* by, extra-perceptual reality. Mere cause, belonging to the non-normative realm of law mode of explanation, cannot thereby play a justificatory role, and to say that our beliefs are merely caused is, in McDowell’s view, to admit that all we have are “exculpations where we wanted justifications” (MW: 8).

This realization, McDowell argues, tempts us towards the other horn of a dilemma - coherentism, the epistemological position that accepts that our perceptual beliefs are caused, not justified, by the world, but claims that

our belief system is nevertheless justified through its own internal coherence. The arguments of Davidson (CTTK), a prominent exponent of this anti-empiricist view, are designed to establish *a priori* that our empirical beliefs must be generally true. McDowell does not attack these arguments themselves, but rather questions Davidson's entitlement to assume from this position that those beliefs, however internally coherent, are empirically contentful - about the world - at all; "when Davidson argues that a body of beliefs is sure to be mostly true, he helps himself to the idea of a body of beliefs" (MW: 68). But, McDowell argues, the absence of any external rational connection means we are not entitled to call this a body of beliefs about a world. Thus McDowell expands Sellars' narrowly epistemological arguments into a claim about intentionality in general, maintaining that the coherentist position threatens not just our entitlement to claim a rational basis for knowledge or empirical justification, but our intentionality and conceptual content themselves. If our beliefs are mediated through blind causal mechanisms, not only can we not know that that which is causing them is the world we take our beliefs to be about, but we forfeit our very entitlement to the idea of their being about anything at all. Further, as William Child points out (1994: 149), if we think of conceptual content as independent of the world in this way we are not even entitled to the concept of causation, or of causal mechanisms, that is in play in the coherentist account. Thus the coherentist model of the interior as isolated from the world "undermines our right to think of it as the realm of thought at all" (KI: 409).

McDowell now aims to show that there is an alternative to these unacceptable positions. Assuming that we *are* open to the world, and

under its rational constraint, as we take ourselves to be, what must be the case for this to be possible?

II. The Core Arguments of Mind and World.

(i) *The Conceptually Structured Nature of Experience.*

One of the core arguments of MW is that the condition that must obtain is that perceptual experience be conceptually structured. Experience, that is, is constituted by concepts, rather than raw sensations. What this means is that in *receptivity* one “takes in, for instance sees, *that things are thus and so*” which is “the sort of thing that one can also, for instance, judge” (MW: 9). We do not *impose* conceptual structure on a non-conceptual intake. We passively take in facts (when we are not deceived; see below) that just *are* the facts that obtain in the world.

As we saw above, McDowell equates ‘the space of concepts’ with that of reasons. The claim that experience is conceptual, then, as the above remark about judging suggests, is the claim that in perceptual experience what we take in becomes assimilated into the network of conceptual knowledge and abilities – the world-view - that, as inhabitants of the space of reasons, we already have in place. More recently (see ENN: 1-14; 260-261) McDowell has modified his view slightly, in that while he still thinks of perceptual experience as actualization of conceptual capacities, he no longer thinks of this conceptual content as propositional – as expressible in ‘that’ clauses. But, he maintains, it becomes propositional when articulated by the subject (*ibid*: 260).

McDowell frames this account of conceptualized perceptual experience in terms of a fusion, or co-operation, of the Kantian notions of receptivity and *spontaneity*. The latter term, in McDowell's usage, refers to "the freedom that empowers us to take charge of our active thinking" (MW: 70) - our capacity to exercise our conceptual capacities freely within their own sphere, or our freedom self-consciously to form judgments and exercise the range of conceptual abilities that constitute our ability to move within the logical space of reasons. As McDowell sees the fusion of receptivity and spontaneity, the former, as the faculty whereby we are impinged on by *intuitions* (perceptual experiences), "does not make even a notionally separable contribution to the co-operation" (MW: 9). Intuitions, that is, already have conceptual content, which thereby becomes integrated into the already existing system of concepts that comprise our spontaneity. The important point is that there is no intermediary either between the world and experience or between experience and judgment. The same seamless pattern of rational interconnections pervades and links both the concepts that constitute our experiential intake and those that are freely exercised in activities like judgment.

As McDowell puts it in his 1998 Woodbridge lectures (WL: 23 & *passim*), the conceptual capacities that are *passively actualized* in perceptual experience are the very same capacities (or, in the light of the above modification, "all but" the same capacities (ENN: 260)) that are *actively exercised* in judgment. Not only this, but, crucially, these two ways in which our conceptual capacities are engaged are interdependent. McDowell writes elsewhere, "(w)e make sense of a consciousness as

having each of those capacities only because we see it as also having the other” (PMW: 427). This two-way dependence is captured in Kant’s claim, “Thoughts without content are empty, intuitions without concepts are blind” (Kant 1929: A51/B75). We could not, *pace* Davidson, have conceptual content if we did not have conceptualized perceptual experience (the first part of Kant’s maxim), nor could we, *pace* the foundationalists, experience an empirical world as a world that exists independently of our experience if we did not already have conceptual content (the second part). De Gaynesford (2004: 21) expresses the central idea thus:

- (i) Concepts are the constituents of experience (so experience requires concepts and conceptual activity).
- (ii) Concepts derive, either directly or indirectly, from experience (so concepts and conceptual activity require experience).

McDowell has called this view “transcendental empiricism” (RN:6; RS:287)), reflecting its derivation from the starting position of the conditions of the possibility of rational constraint. ‘Conceptual activity’ essentially includes self-conscious subjectivity;

the objective world is present only to a self-conscious subject, a subject who can ascribe experiences to herself; it is only in the context of a subject’s ability to ascribe experiences to herself that experiences can constitute awareness of the world (MW:114).

Sellars further emphasizes the interdependence of receptivity and spontaneity, or of the capacity to perceive objects and the possession of a world-view:

I do wish to insist that...if there is a logical dimension in which other empirical propositions rest on observation reports, there is another logical dimension in which the latter rest on the former (EPM: §38).

That is, observational knowledge is possible only for a possessor of a background of other knowledge. McDowell modifies Sellars' account in several ways. First, as noted above, he extends the subject matter from knowledge, narrowly conceived, to conceptual activity, or intentionality, in general. Secondly, in WL I and II, he takes issue with Sellars's insistence that there must be, in, say, a visual perception of a red cube, a sheer manifold of *non-conceptual* receptivity interposed, as it were, between the physical and physiological processes involved in the perception and the subjects having the conceptual content 'there is a red cube in front of me'. As we have seen, McDowell regards the idea that such a non-conceptual item could have a role in the formation of content as an example of the myth of the given, into which Sellars here appears to fall. McDowell traces Sellars' insistence on the non-conceptual element to the latter's view that,

speaking as a philosopher, I am quite prepared to say that the common-sense world of physical objects in Space and Time is unreal – that is, that there are no such things (EPM: §41).

All that exists, on this view, are the fundamental entities postulated in what Sellars, in PSIM, calls the scientific image, equating these entities, according to McDowell (*ibid*: 42n30), with what Kant thought of as 'things in themselves'. Commonsense physical objects like red cubes figure in the contrasting, manifest image. Since, on Sellars's view, they do not exist, these objects of the manifest image cannot be the sources of our sensations of them. Hence the necessity to postulate the non-

conceptual manifest of sheer receptivity, in which the mind – the productive imagination, in Kant’s term – constructs their images.

McDowell believes Sellars is “philosophically misguided”, and misinterprets Kant, when he denies the existence of everyday objects (*ibid.*: 42). On the contrary, on McDowell’s account, it is the presence of worldly objects themselves that directly, in perception, affects our conceptual content, and indeed is a condition of the possibility of our so much as *possessing* conceptual content. The idea of the non-conceptual sensory manifold can be rejected, both because it would represent a return to the myth of the given and because, on McDowell’s account of perceptual experience as conceptually structured, it has no role to play. But how, then, can the presence of a red cube at a particular location in a subject’s visual field be linked to the subject’s acquiring the thought content ‘there is a red cube...’? After all, the visual experience is obviously not a linguistic one, being made up sensations of colour, shape, and location. McDowell’s answer makes use of a further remark of Kant’s:

The same function which gives unity to the various representations *in a judgment* also gives unity to the mere synthesis of various representations *in an intuition* (Kant 1929: A79/B104-5; WL: 30).

While an intuition – “an immediate presence of an object to sense” (WL: 33) is non-linguistic, it is not non-conceptual. Forming the *judgment* that there is a red cube in front of one requires more than that conceptual capacities be actively exercised that enable recognition of redness and cubical shape. As McDowell says (*ibid.*: 30), judging that a red pyramid and a blue cube were in front of one would involve these same capacities. So, in addition, the capacities must be exercised with the right

logical togetherness, such that the object's redness, shape, and location for the subject are linked in a way that is expressible by the sentence 'there is a red cube *there* (in front of me)', that is, a sentence that specifies the object's colour, shape and determinate location. McDowell's claim now is that in a corresponding *intuition*, which may form the basis of a judgment but is not itself a judgment, the very same conceptual capacities, with the same logical togetherness, are *passively* actualized. Thus in an intuition there is a conceptual, but non-linguistic, shaping of sensory consciousness whose logical structure matches that of the corresponding linguistic judgment. As we shall see (III:(iv) below), the conceptual, non-linguistic intuition constitutes the *de re* Fregean *sense* of the object, the red cube. This *de re* sense is a component of the sense of the whole thought that comprises the intuition, 'there is a red cube there', hence enabling the cube itself, in McDowell's term, to *figure* in the thought.

A final way in which McDowell modifies Sellars' idea of the interdependence of observational experience and world-view is by his claim, embodied in transcendental empiricism, that the dependence is transcendental in both directions. For Sellars, like McDowell, the possession of a world-view is a transcendental condition of the possibility of perceptual experience. But while Sellars thinks it is a merely empirical fact that our conceptual capacities are based ultimately on experience, McDowell takes it to be a "transcendental requirement" that

it must be intelligible that conceptual activity has a *subject matter*...the very idea of a conceptual repertoire provides for

conceptual states or episodes in which a subject matter for conceptual activity is sensibly present” (WL: 37).

So far, McDowell’s arguments in this area concern our sensibility - the claimed rational constraint *by* the world on our conceptual contents. But, as he emphasizes, the Kantian aphorism can be framed in a form that applies in the mind-to-world direction as well; “intentions without overt activity are idle, and movements of limbs without concepts are mere happenings, not expressions of agency” (MW: 89). That is, “intentional bodily actions are actualizations of our active nature in which conceptual capacities are inextricably implicated” (MW: 90). Thus *all* our transactions with the world, as persons, or occupants of the space of reasons, are thoroughly conceptualized – “imbued with intentionality” (*ibid*).

In what follows my main concern will be with the mind-to-world direction of interaction – that is, with mental causation and associated issues. The remarks just quoted show that the case that McDowell makes against the myth of the given and for rational constraint upon us by the world applies equally to this direction - to the question of how the world is affected by us. If the myth of the given is the idea that non-intentional items have rational or epistemic effects upon us, its mind-to-world counterpart is surely the idea that contentful items such as John’s believing that..., John’s desiring that..., and so on, affect worldly objects as non-conceptual bare presences. And McDowell’s remarks in MW: 89-90 suggest a counterpart conclusion, that the world that we act upon is *not* non-conceptual.

With the claim of the inextricable involvement of concepts in both perceptual experience and actions, McDowell has shown how we can escape from the dilemma that he identifies between the unacceptable alternatives of the myth of the given on one hand and coherentism on the other. But his radical solution raises further questions. What must the world be like for its impingements on us, or ours on it, to be in conceptual form? How can a picture like this be reconciled with the natural scientific account of the world as thoroughly bound by physical law, the locus of “how things generally tend to happen”? What is the place in this picture of the familiar and surely non-conceptual objects that populate our everyday world? And if the world *itself* presents us with conceptual content, of the form *that things are thus and so*, then how can it be that our perceptions can sometimes deceive us, or that our actions are sometimes not what we think they are? In the remainder of this chapter I will try to explain and elaborate on McDowell’s answers to these questions.

(ii) *The Unboundedness of the Conceptual.*

Profound as McDowell’s claims that the conceptual nature of experience is a necessary condition for thought are, they have consequences that are even more far-reaching. It follows, he argues, from the claims that perceptual experience is already conceptualized and that actions are conceptual ‘all the way out’, that the conceptual is *unbounded* – there is nothing beyond it. This entails that the space of reasons and the space of concepts are co-extensive (MW: 10-14). If the former extended beyond the latter, we would have non-conceptual reasons, in that our judgments about the world would be at least partly influenced by something

impinging on our concepts from beyond their outer boundary – in other words, the given. But on the view McDowell introduces, there is no world that is beyond the reach of conceptualization. It is, on the basis of this argument, transcendentally necessary that the world be *thinkable* – that it be containable within the space of reasons. So the problem of the given, that we seem to be impacted upon from outside the space of reasons, is dissolved, since that space has no outside. In the coherentist picture, on the other hand, the space of reasons and the space of concepts are again co-extensive, but only because both are confined within the mind. On that view, the world beyond our heads, or minds, is devoid of conceptual content, and hence cannot provide us with reasons, but only causal impacts.

The implications of the idea that the conceptual is unbounded are far-reaching. McDowell writes, “The understanding – the very capacity that we bring to bear on texts – must be involved in our taking in of mere meaningless happenings” (MW: 97). Of course, our understanding of each other as rational beings, and of texts and the like, always involves that kind of intelligibility. But McDowell is making the stronger claim that “conceptual capacities, capacities for the kind of understanding whose correlate is the kind of intelligibility that is proper to meaning, are operative also in our perception of the world apart from human beings” (MW: 72). Here we must make some distinctions. The idea is not that we should expect to find ‘meaning’ in inanimate nature in the sense that the latter is actively, purposefully communicating with us. As McDowell points out, the assumption that nature is ‘enchanted’ in that way - full of signs intended for us - was widespread in pre-modern times. Belief in astrology is perhaps the best-known manifestation, but a more vivid one

is the doctrine of signatures, according to which, for example, a visual resemblance between the seeds of a variant of the aconite (Monkshood) plant and the human eye constitutes a sign of the plant's therapeutic efficacy for eye complaints (see Foucault 1966: chapter 2). But this way of thinking was supplanted by the rise of modern science – “a hard-won achievement of human thought” (MW:70). The old view was displaced by the “disenchanted” (*ibid*) image of inanimate nature as meaningless, purposeless, and governed by laws, so that the non-human world is now widely regarded as intelligible only in the realm of law sense. But McDowell's claim is that, even though we cannot return to pre-modern enchantment, nor must we go to the other extreme and think that the only source of meaning in the world is that imposed by ourselves.

The ambiguity in the term ‘meaning’ here is analogous to one that McDowell points out (MW: 28; see *iii*) below) between thought as the act of thinking and thought as content. Only agents, and, derivatively, some of their artifacts like texts or works of art, mean in the purposeful sense of *performing* an act of meaning, and presumably the aconite seed's meaning in this sense was assumed to reside in its being God's artifact. It is this kind of meaning that, at least among most educated people in the developed world, has now irrevocably gone from inanimate nature – few modern theists, for example, think God acts in this way. But McDowell's aim of dissolving the anxiety that springs from the apparent gulf between mind and world – of “reconcil[ing] reason and nature” (MW: 86) - involves the *partial re-enchantment* of non-human nature (e.g., MW: 97). This is achieved by recognizing that nature is the locus of meaning in a second sense, of intelligibility, for rational creatures like ourselves. But here we must be careful to distinguish this view from the

one that, as we have seen, McDowell rejects – the view that it is *we* who project intelligibility on to a world that is outside the conceptual realm. Rather, as McDowell puts it elsewhere, it is *the world* that is “constitutively apt for conceptual representation” (IIW: 310-311). The meaning in question is in nature itself.

Discussing this aspect of McDowell’s thinking, Richard Gaskin (2006: 224-231) introduces “with only minimal metaphorical licence” (*ibid.*: 226) the notion of “the world’s own language”. The world’s having a language, Gaskin argues, is a condition of the possibility of actual and possible empirical languages (*ibid.*). McDowell himself had already rejected this idea; “I do not picture objects as speaking to us in the world’s own language. Objects speak to us...only because we have learned a human language” (WL: 43). But provided we accept that the idea is a metaphor, these positions might not be so very different. That it is a metaphor is clear in that the world can ‘speak’, not, of course, in any literal sense, but in the above sense of being intelligible, of having meaning for us through its conceptual structure, as the latter is made manifest, for example, in the logical togetherness of the colour, shape, and location of the red cube in McDowell’s example. And Gaskin’s suggestion that our having language is conditional on the world’s having its own language can be interpreted as a figurative way of expressing an aspect or component of McDowell’s transcendental argument, above - the claim that our being under rational constraint by the world is a condition of our having conceptual content. McDowell shows the closeness of his view to Gaskin’s when he writes:

A seen object as it were invites one to take it to be as it visibly is. It speaks to one: if it speaks to one’s understanding, that is just what its

speaking to one comes to. “See me as I am”, it (so to speak) says to one; “namely as characterized by *these* properties” – and it displays them (WL: 41).

The fineness of the distinction here is apparent in McDowell’s comment on the passage just quoted, “the point is that objects come into view for us only in actualizations of conceptual capacities that are ours” (*ibid.*: 43). The *capacities* are ours, but they are capacities that, when actualized in perceptual experience, enable us to respond to a conceptual structure that belongs to the world itself, as is captured by the idea of logical togetherness.

(iii) *Thinkables, Facts, and Objectivity.*

The claim that the conceptual is unbounded, then, is not the pre-modern belief that the world is imbued with meaning. Rather, it is the claim that it is conceptually structured and so inherently conceptualizable. One way in which McDowell expresses this is as a Wittgensteinian idea; “When we say, and *mean*, that such-and-such is the case, we – and our meaning – do not stop anywhere short of the fact; but we mean: *this – is – so*” (Wittgenstein 1958: §95). That is, as McDowell puts it (MW: 27-28), “there is no ontological gap between the sort of thing one can mean...and the sort of thing that can be the case. When one thinks truly, what one thinks *is* what is the case...the world is made up of the sort of thing one can think”. We should note here the distinction between the sort of thing one can think and the sort of thing one can think *about* (see Fish and Macdonald 2007: 40). McDowell is not just making the unremarkable claim that we can think about the world, that is, represent it in thought. The claim is the much stronger one that the contents of our

thoughts and what makes up the world are ontologically the same, and that “we conceive the world...as, precisely, everything that can be truly thought or said: not everything we would *think about* if we thought truly, but everything we would *think*” (CCP: 143). But again, it is important to be clear that this is not an idealistic claim that “renounc[es] the independence of reality, as if we were representing the world as a shadow of our thinking, or even as made of some mental stuff” (MW: 28).

Ensuring that the picture is not one of an unacceptable idealism depends upon another distinction, that between thought as an activity, that is, the *act* of thinking, and thought as content – as *what* we are thinking (*ibid*), the first component of the distinction made in the previous paragraph. The world, on McDowell’s account, is made up not of thoughts in the first sense, but of ‘the thinkable’ – that which can be the content of true thoughts, that is, *facts*. What enables the world to be thinkable is that, as for the early Wittgenstein (1922/2001:1.1), the world consists of facts, and for McDowell facts are “Fregean senses, with Fregean senses as their constituents” (RN: 93). I discuss the neo-Fregean aspect of McDowell’s philosophy below (III:(iii), (iv)); for now, we can think of Fregean senses, on McDowell’s account, both as thoughts and as how we think of, or identify, objects. When we think truly about the world, he claims, the worldly fact is *identical* with what we think. So the world is made up of conceptualized items, which are true *thinkables*. This captures the idea that the world is independent and capable of exerting the required rational constraint on us, from “outside *thinking*, but not from outside what is *thinkable*” (MW: 28). *Which* conceptual contents the world presents us with in perceptual experience is not under our

control, but on pain of irrationality or other error we are obliged to take those contents into account in forming our empirical judgments.

We thus have two distinctions within the concepts of thought and thinking in McDowell's account. As we saw, the act of thinking must be distinguished from thought content, which belongs, along with worldly facts, in the realm of Fregean sense, or *Sinn*. Content - what we are thinking - must also be distinguished from what we are thinking about - that is, objects and their properties. I return to how objects fit into McDowell's account in III:(iv) below. The idea of a fact as a true thinkable clarifies McDowell's notion of *objectivity*. In 'Towards Rehabilitating Objectivity' (TRO, 2000), he sets out his objections to Richard Rorty's rejection of any conception of objectivity that goes beyond community consensus. Rorty argues (e.g., 1991:21-34; 1998b) that the view that, both in everyday life and in science, our judgments are answerable to how things are in a reality beyond ourselves, is a last vestige of a religious hankering for a trans-human authority.

As McDowell sees it, Rorty conflates two ideas. The first is the idea that the world constrains our judgments - that there is a way things are in the world that transcends the totality of peoples' opinions. Second, there is the idea that the world is inaccessible to us, in the Cartesian sense that there is a 'veil' between our senses and an unknowable reality. Lynne Rudder Baker, in *Explaining Attitudes* and subsequently (EA: 232-236: AA: 70) usefully puts this contrast as that between recognition-independence and mind-independence, where the latter is to be understood as independence from *thinkability* or recognizability in general, and not just from acts of thought. Both Baker and McDowell

understand objectivity as recognition-independence. To understand it as mind-independence in this sense would, for McDowell, be to accept that there is a way the world is that is independent of, or beyond, possible conceptual experience, which is just the givenist view that he rejects. Rorty, however, rejects both recognition- and mind-independent notions of objectivity, concluding that all that we can mean by the ideas of objectivity, and hence truth, is conformity to current practice – ‘solidarity’.

McDowell (TRO: 212) agrees with Rorty’s acceptance of the claim (attributed to Putnam; Rorty 1991: 6) that “notions like “reference” – semantical notions which relate language to non-language – are internal to our overall view of the world”. But Rorty thinks this means rejecting outright the idea of answerability to the world, and this is just where McDowell disagrees. McDowell grounds objectivity in the idea of truth as a normative concept expressible by Tarskian disquotation. Thus an utterance of ‘Cold nuclear fusion has not yet been achieved in the laboratory’ is true, or warranted, or justified, if and only if cold nuclear fusion has not yet been achieved in the laboratory (TRO: 218; call this sentence or proposition ‘C’). The normativity - warrant or justification - does not consist in the disquoted claim’s acceptance by one’s peers but by whether or not C is true. The point is that this notion of justification does not, as Rorty claims, involve “trying to climb outside one’s own mind” (*ibid*: 218). The truth of the claim ‘C’, then, consists in C’s being a *fact*, a true thinkable, and “whether or not cold fusion has occurred [i.e., whether or not C *is* a true thinkable] is not the same as whether or not saying it has occurred will pass muster in the current practice” (*ibid*: 219). The thought that C is the thought that “things really are a certain

way ... to insist on this distinction is not to try to think from outside our practices; it is simply to take seriously the idea that we can really mean what we think and say from within them” (*ibid*).

To put this slightly differently, truth and objectivity are to be conceived of and viewed from the only perspective available to us, the normative perspective of the place we occupy in the space of reasons. This is all we can mean by the way “things really are”. I think we can interpret McDowell’s view as that the possibility of our being capable of self-conscious empirical thought at all is conditional upon this notion of objectivity’s being available. The same perspectival constraint applies to all the other concepts that he discusses (or denies the need to discuss) in his writings; meaning, intentionality, causality, ethical and aesthetic value, and indeed the idea of the world itself. The common theme is that these, like all concepts, originate from, and presuppose, our practices and our perspective on the world, and cannot be viewed from sideways-on. These points are also brought out in McDowell’s writings on Wittgenstein’s rule-following arguments and their interpretation by Kripke and others (see WFR). McDowell rejects two interpretations; one is that when we follow a rule such as addition in arithmetic the answer is, as it were, *already there* in what Christopher Norris (2006: 185) calls “an ‘ethereal machine’ – a kind of ghostly paradigm – that hovered above our rule-following practices and somehow determined their objective correctness or otherwise”. But, consistently with the above criticism of Rorty, McDowell also rejects the interpretation, associated with Kripke (1982), that that correctness or otherwise derives solely from, and amounts to no more than, conformity with communal practice. McDowell argues that that our warrant for the correctness or otherwise

of the results of previously unattempted additions is that it does indeed derive from “the idea of a custom or practice” (WFR: 242), but, as with the cold nuclear fusion example above, there is more to these notions than simply mutual agreement. Rather, his claim seems to be that the standards governing addition, like the facts in the earlier case, and indeed all the “dictates of reason” are “there anyway” (MW: 91) – not in the sense of Norris’s “ghostly paradigm”, but in a sense that their objectivity (again, as recognition-independence, or “ratification-independence”, as McDowell expresses it in WFR)) is a condition of the possibility of our rationality.

So we have an idea of McDowell’s conception of the structure of the unbounded conceptual world; it is made up of facts, or true thinkables, which are objective in the sense that their obtaining is independent of anyone’s thinking, but whose obtaining is nevertheless internal to our world-view. Our perceptual experience is of facts, whose conceptual structure and objectivity enables them to rationally constrain us.

(iv) The World in the Space of Reasons.

It is a corollary of the unboundedness of the conceptual that the space of reasons includes the world itself. McDowell’s claim that “[t]he understanding – the very capacity that we bring to bear on texts – must be involved in our taking in of mere meaningless happenings” (MW: 97) is made in the context of his criticism of Kant’s insistence on a supersensible reality, a noumenal realm, with which our subjectivity co-operates to create our empirical world. McDowell’s Strawsonian

interpretation of Kant has been challenged (e.g., Bird 1996, Allison 1997, Norris 2000, Friedman 2002) but just how faithfully it reflects Kant's own thought is not the point at issue here. He argues that Kant's mistaken insistence leads him to find in nature "the intelligibility of law, but not the intelligibility of meaning" (MW: 97).

The first part of what McDowell calls "Kant's insight" (TSN: 178) is that the world *must* have an intelligible structure, in contrast to Hume's view of the natural world as "an ineffable lump, devoid of structure or order" (*ibid*). Any empirical world-*picture*, the Kantian argument goes, is clearly conceptually structured. But if the picture is accurate it must then represent the world itself as having a structure that matches "the structure in the space of *logos* [reasons] possessed by accurate representations of it" (*ibid*). It is a condition of the world of nature's being a world at all that it be composed of things that are the case. As McDowell writes;

But mirroring cannot be *both* faithful, so that it adds nothing in the way of intelligible order, *and* such that in moving from what is mirrored to what does the mirroring, one moves from what is brutally alien to the space of *logos* to what is internal to it (*ibid*: 179).

Since "what does the mirroring" is conceptually structured, then, if the representation is accurate, so must be "what is mirrored" – the world. The necessarily intelligibly structured world, then, is *conceptually* structured. It is "not constitutively independent of subjectivity, which has its being in the space of *logos*" (*ibid*: 185).

According to McDowell, Kant fails to reach this conclusion because of his insistence that there must be something beyond subjectivity, or the conceptual, hence his reliance on the “in itself”, the noumenal realm, which, along with *our* subjectivity (on McDowell’s reading), jointly determines the intelligible structure of nature (*ibid*: 180). But then, when the idea of the “in itself” is jettisoned by post-Kantian thinkers in response to the unacceptability of this position, McDowell claims (*ibid*: 179), its role is taken over in the standard, neo-Humean (or bald naturalist) account, by the intelligible, empirical world itself. That world is describable according to the realm of law mode of intelligibility, but “all other intelligible order, all meaning or value...is conceived as partly a reflection of our subjectivity” (*ibid*). But if Kant’s insight that the world is intelligible is retained *minus* the idea of the ‘in itself’, we arrive at McDowell’s picture of the world as exhibiting a meaningful structure that mirrors that of our rationality. McDowell’s argument that this must be the case is essentially the same as his transcendental empiricist argument ((i) above) that our experience must be conceptual. He writes, further, that

the very idea of thought – the exercise of the intellect – presupposes a notion of objectivity that we can gloss in terms of a distinction between being right and seeming right. And the idea of the world, as it figures in [the thesis that the world is not constitutively independent of subjectivity], need not amount to more than an expression of that notion of objectivity (*ibid*: 185).

The notion of objectivity in question here is clearly the recognition-independent one described above, and it is also a normative one. And the world itself is to be understood just in terms of this normative notion of objectivity – as what is actually the case, and as of something whose

existence, and ability to rationally constrain us, is a condition of the possibility of thought. The content of a thought, or utterance, - say expressing a belief - gives the idea of a way that the world is, and thus also the idea of ways that the world is not. The very possibility of rationally entertaining the thought or making the utterance, then, depends on the notion of objectivity that makes possible the distinction between content and how things are in the world. As Michael Luntley (1999: 368n10) puts it, to take this view is “to see the world as that which is required in order for language to be the sort of thing it is: viz. a system with the semantic capacity for being correct/incorrect”.

When we adopt a scientific approach to the world, we abstract away from space of reasons thinking and confine ourselves to the intelligibility of the realm of law. The discovery that the world has this kind of intelligibility – that it answers to this kind of inquiry - and the scientific picture that the exploitation of the discovery has produced are, as McDowell acknowledges, a hard-won achievement of modernity. But its having this kind of intelligibility, he insists, does not point us to the ultimate picture of reality – scientific facts are only some of those that make up the world, conceived as the totality of what is the case. For example, much of the argument of ‘Two Sorts of Naturalism’ is aimed at establishing that *ethical* facts are part of the world. There is a difference, McDowell acknowledges, between how we arrive at scientific and at ethical facts. In both cases, our conclusions, that things are thus and so, are constrained by reason, but in the scientific case, in addition, “investigation has led to that conclusion because of the causal influence of the fact that things are thus and so” (*ibid*: 186-187). But while this gives science a special kind of objectivity, its use of the concepts of the

realm of law mode of intelligibility does not confer on it, as a way of approaching reality, a status that somehow exempts it from the need to justify itself according to the standards of “*logos*” that apply to other approaches. Realm of law intelligibility is still intelligibility, and intelligibility presupposes the normative categories of thought, such as justifiability. The conclusions of science are not about a disenchanted nature. Rather, like those of every other form of thought, including the very notion of causation that McDowell refers to above, they are internal to the space of reasons.

As McDowell writes (*ibid.*: 181) “[T]he investigative stance of science discounts for the effects of features of the investigator, even his humanity...what science aims to discover is the nature of reality in so far as it can be characterized in absolute terms: the content of the view from nowhere”. But the investigative stance of philosophy, on this conception, does not need to, nor should it, aspire to do this. It is not only that the features and point of view of the investigator, and of humanity in general, must be part of the subject under investigation. More fundamentally, the investigator’s standpoint, that of the space of reasons, is presupposed by the possibility of investigation, including scientific investigation, at all.

III. Openness to the World.

(i) Error, Descartes, and the Highest Common Factor View.

I turn now to McDowell’s defence of his position against what he suggests is “the real disease of thought” (STEIS: 246), the neo-Cartesian

idea of the mind as consisting of an autonomous, self-contained system whose operation are essentially isolated from and independent of how things are in the world beyond it. Countering this view involves adopting a disjunctive account of perceptual experience and the development of a neo-Fregean account of how our minds relate to external objects.

McDowell plays down the significance of the fact that our perceptual experiences can sometimes deceive us; “[i]t does not matter much that one can be misled” (MW: 9). But, the question arises, how is the claim that we are directly open to the world in our experiential taking of worldly facts into our conceptual repertoire – that “[e]xperience enables the layout of reality itself to exert a rational influence on what a subject thinks” (MW: 26) - compatible with the possibility of perceptual error? The possibility is generally acknowledged that someone might have perceptual experiences – like, say, Macbeth’s hallucinated dagger, in which it seems to the subject exactly as if the world is a certain way – as if, on the McDowellian picture, certain facts are directly perceived to obtain, but in fact they do not obtain. This would precisely not be a case of perceptual openness to the world. Further, it seems that in two scenarios, in one of which he is confronted with a real dagger and in the other with a hallucination of a dagger, Macbeth has exactly the same visual experiential content ‘there is a dagger before me’. Whether or not the experience is veridical is a matter of how things are in the world, a fact that is cognitively unavailable to the subject, at least at the time of the experience, just because the experiential content is *common* to the veridical and hallucinatory cases. The possibility of error thus seems to ensure that it cannot be a case of there being no ontological gap between the subject’s content, ‘there is a dagger before me’, and a worldly fact. As McDowell puts it himself, “how can one know that what one is

enjoying at any time is a genuine glimpse of the world, rather than something that merely seems to be that?” (MW: 112). It seems “that however good a subject’s cognitive position is, it *cannot* constitute her having a state of affairs directly manifest to her” (MW: 113). Even in the ‘good’ case, in which perceptual experience is veridical, its object is “something that falls short of the fact itself” (CDK: 387).

This argument from error thus seems to lead us back into the problems of the coherentist position discussed in I:(ii), in that, by throwing doubt on the possibility of our perceptual experience’s being directly of the world, it threatens not only McDowell’s account of how the world rationally constrains us but also our entitlement to the claim that our experiences are even about an independent world. So the very survival of McDowell’s picture of mind and world depends on the successful resolution of the problems raised by the possibility of perceptual error. His solution is to reject the account of perception assumed by the error argument in favour of a disjunctive account. In ‘Singular Thought and the Extent of Inner Space’ (STEIS, 1986)¹, he describes the origins and basis of the “fully Cartesian picture” (236ff) that he opposes.

McDowell follows M.F. Burnyeat in tracing this view’s origins to Descartes’ break with the ancient sceptics (239-243). While the ancients recognized that there were sceptical problems associated with our knowledge of the external world, they did not, on this interpretation, have the concept of the inner realm, of how things appear to one, as something about which there could be independent facts or truths in the

¹ For the rest of this subsection, unattributed page references are to STEIS, reprinted in MKR: 228-259.

same way as there are about the external world. It was Descartes who introduced the notion that “how things seem to a subject [is] a case of how things are” (239), in the same sense of ‘how things are’ as applies in the outer world, and formalized the idea that one can be infallible about how things are in one’s inner space. So, while sceptical problems remain with regard to our knowledge of the external world, “we can retreat to the newly recognized inner reality” (*ibid*) and be assured that at least we have knowledge of (paradigmatically perceptual) subjective appearances.

The ideas that there are facts about one’s subjective consciousness, and that one can have knowledge of these facts, do not in themselves threaten “the loss of the world”, McDowell argues (240, 242). That threat, rather, comes from the further Cartesian claim that the inner realm is transparent, “accessible through and through” (240), so that as well as its contents being infallibly knowable by the subject, there are no further facts about the inner realm to which the subject does not have infallible access. Crucially, however, the introduction of this new concept has the effect of driving a conceptual wedge between knowledge of these inner facts and all other knowledge, which, at best, becomes downgraded. In a case of possible perceptual error, for example, it is a consequence of this view that the only fact in play is the one that the subject knows indubitably - that there is an appearance. The answer to whether or not the appearance corresponds to a worldly state of affairs is *not* accessible to the subject in this way, so, *ex hypothesi*, that answer cannot be a fact about the inner realm. McDowell calls this the “highest common factor” view (CDK: 386). Only the situation obtaining in the

world can determine whether or not the appearance is veridical, but this is not the kind of fact to which the subject has direct access.

Thus the claim that the inner realm is transparent leads to the further conclusion that it is autonomous, in that, as McDowell says, “subjectivity is confined to a tract of reality whose layout would be exactly as it is however things stood outside it”. Further, “the common-sense notion of a vantage point on the external world is now problematic” (241). We are confronted by an unbridgeable ontological and conceptual gap between mind and world, with the possibility that for all that the subject can know the world may not exist. This is not only the loss of the world. As we saw above in the discussion of the shortcomings of coherentism (I:(ii)), the absence of the possibility of rational constraint by the world that is entailed by this account makes it the case that we are not even entitled to the notion of intentionality, or conceptual content.

The crucial and illegitimate Cartesian move, as McDowell sees it, is the inference from the undisputed fact that, in the veridical perception versus mere appearance case, everything in inner space seems the same *to the subject*, to the conclusion that (because of the transparency) everything in inner space is the same, *simpliciter*.

McDowell points out that many contemporary models of the mind, while rejecting Descartes’ substance dualism, retain the idea of the isolated inner space, now thought of as identical with, or otherwise a product of, the brain. The kinds of theories he has in mind are those that share “something along the lines of functionalism” (244, 246), and what he objects to in these models is that they insist upon the “self-standingness

of the inner realm” (250), in that they equate the functioning of physical mechanisms in the brain with rationality itself (see NPM: 100-106). They tend to be two-component theories (see McGinn 1999), which combine this view of internal functioning with a causal account of the relation between the self-standing realm and the outer world, and as such, McDowell argues, they suffer from the same insurmountable difficulty as Davidson’s coherentist account of perception (I:(ii) above); since the interior is connected to the world only causally, we are not entitled to assume that its content represents a world (251). Even non-reductionist physicalist views of mind, such as Davidsonian anomalous monism (ME; see chapter 2) suffer from this defect to the extent that they are understood in the context of the kind of coherentist isolation from rational constraint discussed in I:(ii) above. So they are prey to the same crippling problems; “once we picture subjectivity as self-contained, it is hard to see how its states and episodes can be anything but blind” (244); that is, how it can be about a world.

(ii) *The Disjunctivist Solution.*

McDowell’s alternative picture substitutes a *disjunctive* conception, which allows that a subject may not know which of two disjuncts gives a correct description of his perceptual content, together with a picture of openness to the world, in which mind and world, rather than being independent, *interpenetrate*. The essential move is rejection of the conflation of the contents of her inner experience’s *seeming* the same to the subject and its actually *being* the same. Rather, it is not at all unnatural, McDowell suggests, to say that

an appearance that such-and-such is the case can be *either* a mere appearance *or* the fact that such-and-such is the case making itself perceptually manifest to someone (CDK: 386-387).

That is, in the veridical (good) case the experience does *not* fall short of the fact, even though the subject may not be able to distinguish it from the non-veridical, or bad case. The argument depends on the assumption that a subject may be in error about the contents of her own subjective experience. In the bad case, the subject may take herself to be thinking that there is a red cube *there*, where ‘there’ designates a particular position in her visual environment. But she is mistaken; not only is there no red cube there, but according to McDowell, she is not even entertaining that thought. In the absence of the object at which the thought purports to be directed, according to this picture, even though the subject may take herself to be thinking about that object, her thought, however it may be positively characterized, is not in fact about that object. The contents of such thoughts are thus partly constituted by the worldly facts. So the obtaining of a worldly fact is “precisely not blankly external to [a subject’s] subjectivity, as it would be if the truth about that were exhausted by the highest common factor” (CDK: 391).

Thus McDowell disarms the argument from perceptual error. In a case in which Macbeth *correctly* perceives that there is a dagger before him, on the disjunctive account, he is indeed open to the world. The contents of his thought are identical with the worldly facts. We may think of this view as a special case, applying to immediate perception, of McDowell’s broader anti-Cartesian view that singular thoughts – thoughts directed at particular objects, whether they are perceptually based or not – are *object-dependent* (STEIS: 247-249). McDowell, following Evans, whose

posthumous (1982) he edited, develops the account of object-dependent thoughts from a re-interpretation of Frege's notions of sense and reference.

(iii) Sense and Reference.

We saw (II:(iii)) that McDowell regards both thoughts and facts as Fregean senses (*Sinne*) - as he writes, "thought and reality meet in the realm of sense" (MW: 180). In fact, his whole project of replacing the dominant picture of mind and world as occupying separate realms with one in which they interpenetrate is based on a radical interpretation of Frege and neo-Fregeanism.

Frege's main motivation for the introduction of the notions of *Sinn* and *Bedeutung* (reference) is usually taken to have been the need to account for how someone could, without irrationality, take different cognitive attitudes to utterances containing two co-referring terms. Thus (taking the best-known example) we might explain how someone might assent to 'Phosphorus is visible' but dissent from 'Hesperus is visible', even though both names refer to Venus, by saying that the two names, and the sentences containing them, have different senses. As McDowell makes clear (EF: 167-170), this is indeed the most salient application of the distinction. However, it encourages the view that the sense of a name is best understood as the mode of presentation of the object named, and, as Thornton (2004: 243) points out, this in turn tends to encourage a view of sense as something distinct from the object, such that the object lies *beyond* the realm of sense, in the realm of reference. It would be as if the self/reference distinction marked a boundary between the conceptual and

the non-conceptual, a picture that is obviously quite incompatible with McDowell's.

Michael Luntley highlights the inadequacies of this view of sense;

The model appears to be this: We think about objects with names by having before our minds entities called senses. These entities are modes of presentation...But if we think of the way the object is presented to us in thought as an entity associated with the name, then sense turns out to be a blueprint for the object. We never directly think about objects...the realm of thought [is] wholly disengaged from the world (Luntley 1999: 226)².

The picture of senses as standing between us and the world takes us back to something very like that of the self-standing inner realm. This is another example of the “sideways-on” approach (MW: 34ff), the error of trying to understand an aspect of the mind-world relation from a standpoint outside the space of reasons, as if we could say, here the sense, there the object (Thornton *op. cit.*:243). It suggests, also, that to know the sense of a name is to be in possession of some descriptive way of identifying the thing named, so that to know the sense of, say, “Hesperus”, is to know it as the name of the star that appears at a certain point in the evening.

Frege's own view may be vulnerable to this kind of criticism, but a number of later neo-Fregeans, like Evans, McDowell, and Luntley have modified Frege's theory considerably. One of McDowell's stated aims in SRPN is to show how an interpretive truth theory can serve as a theory

² References to Luntley are to his *Contemporary Philosophy of Thought* (1999).

of sense (see Platts 1979: 57-63; Sainsbury 2005: 33-41). A truth theory for a language has axioms of the form

s is true iff p ,

but for a theory to be interpretive, it must also be the case for each axiom that

s means that p .

As I understand it, the point is that a proposition such as

‘Hesperus is rising’ is true iff Phosphorus is rising,

while preserving truth, is not interpretive in this sense since it is possible for a rational speaker to assent to one and dissent from the other of the two clauses. Intuitively, then, we cannot infer from it the further proposition “‘Hesperus is rising’ means that Phosphorus is rising” – the initial proposition, after all, could be about two heavenly bodies that happen always to rise at the same time. So as an axiom in a theory the statement will not contribute towards our ability rationally to understand such a speaker, and so will fail to capture the essential requirement of a theory of sense. For, as McDowell writes;

The point of the notion of sense...is...tied to our interest in understanding behaviour, und ultimately our interest in understanding – fathoming – people (SRPN: 172)

According to McDowell's "austere" theory of sense (*ibid.*: 172), to know the sense of a name, say 'Hesperus', is to know that 'Hesperus' stands for Hesperus, *in such a way* that one is enabled both to use the name oneself and to understand the speech and behaviour of others who use it. Or as Dodd (2008: 54) writes, "[t]o give a name's sense is to specify *what must be known* by someone who understands the name". It might seem that that someone know that, say, 'Hesperus' stands for Hesperus is too insubstantial a condition to be capable of specifying what can serve as a source of rational power. But the point is that for McDowell *Sinn*, like any other concept, can only be understood from within the space of reasons – that is, from the standpoint of a person who is "already competent in the use of the very name in question" (SRPN: 191). Austere though the theory is, it is not trivial, and when viewed in context it is adequate to its task. For example, it can clearly account for differences in rational power due to different attitudes to co-referring terms – someone could know the senses of 'Hesperus' and 'Phosphorus' in this way and not know that they refer to the same object.

Crucially, the theory also picks out the difference between someone who believes that (to modify Luntley's example, 257; see SRPN: 182, 192)

"Vancouver" stands for Vancouver

and someone who merely believes that

"Vancouver" stands for Vancouver' is a truth.

The latter believer might be, say, a non-English speaker who has been told, in her own language, that ' "Vancouver" stands for Vancouver'

states a truth in English. This knowledge alone, however, would not enable her either to use it or understand others' use of it. But someone who knows the *sense* of "Vancouver" is able to place the name in the context of an overall theory, by virtue of being a competent user and understander of the name.

On the austere theory, according to McDowell, "the differences in sense are located no deeper than the failures of substitution" (SRPN: 191). The fact that someone could hold different beliefs about Hesperus and Phosphorus is not *explained* by their having different senses, as if sense reflected the workings of an inner mechanism. Rather, their having different senses is simply a *consequence* of the fact. "As far as names are concerned," McDowell writes, "the ontology of a theory of sense need not exceed the names and their bearers" (SRPN: 175). A name's sense, then, is not an entity. An object's name's sense consists in its contribution to the sense of the sentence in which it occurs, and hence to what Luntley (209ff) calls the "rational power" of the sentence – the ability of sentential content to enable us to make rational sense of speakers' behaviour. This contribution is just that of supplying the reference of the name as the speaker understands it. As Luntley argues (235), it is not possible to think of an object without thinking some whole thoughts about it. The idea that it *is* possible would be a form of the myth of the given. The sense, on this understanding, turns out to be what is *shown* when someone competently manifests her understanding of the name – in the simplest case, by correctly referring to the object named (SRPN: 175). From the point of view of McDowell's project of bridging the perceived gulf between mind and world, a crucial element of this picture is that the sense of a name is not descriptive; "to know the reference of a name.. [is] .. to know that object: acquaintance, perhaps,

but in any case...knowledge of things” (SRPN: 174). The sense of the name is not descriptive but *de re*.

Williamson (2007: 16) claims that McDowell’s claim (II:(iii) above) that there is no ontological gap between thought and world is false because “what one thinks is individuated at the level of sense while what is the case is individuated at the level of reference”. But the austere theory, I think, disarms this kind of objection. Rather than being a “level”, sense, on that theory, is just that *through* which the referent is known.

(iv) *Objects, de re Senses, and Object-dependent Thoughts.*

McDowell says of objects that they *figure* in thoughts as Fregean referents; “...for an object to figure in a thought, a thinkable, is for it to be the Bedeutung associated with a Sinn that is a constituent of the thinkable” (RN: 94-95). “Figuring in” is contrasted with “be[ing] a constituent of” (STEIS: 237). Responding to Travis (ENN: 258-267) McDowell writes that objects

are not conceptual...[b]ut they are not *outside* the conceptual...[w]hen I say reason’s reach extends no further than the conceptual, I am not separating things into two sorts, those that can bear rationally on what a subject should think and those that cannot (*ibid*: 259).

If there is a suspicion of ambiguity in the first two occurrences of “conceptual” in this passage, it may serve to clarify what McDowell means by the claim that the conceptual is unbounded. An ordinary, worldly object is “not conceptual” in the sense that, unlike a thought, it is not ‘of the mind’ – it is extra-mental. But “conceptual” in the second

occurrence seems to mean something more like ‘conceptualizable’ although, as explained above, this must be taken to mean more than just representable in thought. Rather, the sense in which the conceptual, in this sense, is unbounded is that the world, containing objects and their properties, has a structure that *mirrors* the propositional structure of our rationality (TSN: 179). And once we predicate anything of an object it figures in that propositional structure. So although the object is not conceptual, that does not mean it is part of the given. As I understand McDowell, *qua* non-conceptual, the object is something like a ‘bare presence’, which cannot figure in any way in our conceptual economy. So it appears that, for him, all that need, or indeed can, be said about objects is what one says when one says that to be an object is to be a referent associated with a sense, which thereby figures in a thought. Or, returning to the metaphor of the world’s own language (II:(ii) above), we might say that objects figure in McDowell’s picture as that which the world says something about.

The remaining task is to show how objects, thus conceived, *can* be part of our conceptual economy, and can thus play their part in the rational constraint that the world exerts on us. McDowell achieves this by combining the Russellian notion of knowledge by acquaintance with Fregean sense theory, giving a picture in which objects figure in singular thoughts as *de re* senses, in such a way that the thought can only be entertained if the appropriate object exists. Russell (1905) famously restricted the category of logically proper names to demonstratives such as ‘this’ and the first personal pronoun, that is, names of features of sense data – inner items whose existence as such is, according to Russell, not open to doubt. All other names expressed by singular terms –

‘Barack Obama’ or ‘that red cube’ – are, on Russell’s account, disguised definite descriptions. As McDowell explains (STEIS: 228-233)³, firstly, for Russell, knowledge by acquaintance (Russell 1910/11) – knowledge of objects stemming from their immediate presence to one’s mind (231) – is object-dependent in the above sense. This is because such Russellian singular thoughts are partly constituted by the objects, and so could not be thought if the object did not exist. But it is possible to be mistaken in one’s belief that one is thinking about the bearers of names, so if ‘Barack Obama’, say, was a logically proper name, that would amount to its being possible for a subject to be under the illusion that she was having a thought. That possibility is ruled out, for Russell, by his adherence to the Cartesian view that McDowell rejects (III:(i) above) – the view that our thoughts, the contents of our inner space, are completely transparent to us. Treating such names as definite descriptions – that is, in Russell’s (1905) sense, as characterizations purely in terms of quantification and predication – means that in the case where there is no corresponding object, thoughts containing the name are merely false, rather than nonsensical.

If we replace the Cartesian view with McDowell’s, in which inner space is not transparent, we can accept the possibility that a subject can be in error about it. We can accept that she could believe herself to be having a singular thought when in fact she is not having such a thought. This is a generalization of the disjunctive theory of perception described above – the disjuncts are either (a) that the subject has a singular, object-dependent thought about an object, or (b) that in a case in which the object which would have figured in the singular thought does not exist, it

³ For the rest of this subsection, unattributed page references are to STEIS, reprinted in MKR: 228-259.

merely appears to the subject that she is having a thought of the kind she supposes herself to be having. In the latter circumstances, it is not that “h(er) mind is wholly vacant” as “images and words may clearly pass through it”, (Evans 1982: 45-46) which may well be so similar to those which would be present during a genuine case that the subject cannot tell the difference. One might regard such appearances as meta-thoughts, as McDowell does (SRPN: 185-186), or perhaps as really being descriptive thoughts in Russell’s sense. In any case, the subject may thus not be in a position to know which of the two disjuncts applies, and another person may be in a better position than the subject herself to know this – a better position, that is, to know whether or not the subject is having a genuine singular thought. The complete picture of what is happening in the subject’s inner space, on this view, is only visible from a standpoint that includes the world as well as the subject’s thoughts - this is what is meant by the interpenetration of mind and world. As McDowell says, ‘which configurations a mind can get itself into is partly determined by which objects exist in the world’ (230); in other words, the subject’s experiential content is partly *constituted* by how things are in the world. The contrast between this externalist position and the Cartesian picture of the isolated, autonomous inner space now stands out clearly.

The application of the disjunctive theory to singular thoughts, and with it the admissibility of mere illusions of singular thoughts, means there is no longer a need to apply Russell’s restriction of proper names to the very small group. So names of ordinary objects, previously characterized as disguised descriptions, can now become the names of objects that figure in object-dependent singular thoughts. But a different problem now becomes apparent (233) – the problem, discussed above, of co-referring terms. The problem does not arise under Russell’s restriction – it is hard

to imagine a sense datum of, say, redness, to which a subject could rationally adopt opposing propositional attitudes, say belief and disbelief, even on different occasions. But when the restriction is removed, the problem that Frege highlighted emerges for Russellian singular thoughts. McDowell's solution (233-237) is to suggest that,

Russell's insight [that singular thoughts are object-dependent] can perfectly well be formulated...by claiming that there are Fregean thought-constituents (singular senses) that are object-dependent, generating an object-dependence in the thoughts in which they figure. Two or more singular senses can present the same object; so Fregean singular thoughts can be both object-dependent and just as finely individuated as perspicuous psychological description requires (233).

That is, replacing the Russellian idea that objects are constituents of object-dependent thoughts with the idea that Fregean singular senses figure in them allows us to account for how a rational subject can take opposing attitudes to, say, 'Hesperus is visible' and 'Phosphorus is visible', and, more generally, to understand the rational import of subjects' utterances and behaviour through an understanding of their thought contents.

On Russell's restrictive account, in which the only objects referred to directly are the bearers of logically proper names, there is no difficulty in accounting for *how* the objects become part of our thought contents – on his Cartesian view, sense data, the contents of immediate memory, and one's self-awareness are immediately present to the mind. If we follow McDowell's account, however, we need to understand how an object,

the bearer of an ordinary proper name, through its sense, can become part of the conceptual order – we need, in other words, some filling out of the notion of ‘figuring’. With such an account in place, we would have what McDowell needs, a direct, conceptual link between mind and world.

McDowell achieves this through a modification, or perhaps an interpretation, of Fregean theory which allows thought to refer directly to objects. Rather than, with most interpreters, taking the senses of singular terms to be descriptive, McDowell (DRS: 214), adopts Evans’s view of object-dependent senses of singular terms as *de re*. That is, in a successful thought about an object, the object *figures* in the thought by virtue of the *de re* sense of the object’s being a *constituent* of the thought, as in the example of the red cube in II:(i) above. Because the sense is *de re* – essentially involving the object itself, rather than a specification of it (Sainsbury 2005: 41) – the thought is object-dependent.

This view still needs defence, however, against purely descriptivist interpretations of Fregean sense. McDowell argues that the key to seeing the superiority of the singular thought approach is to seek for the conditions of possibility of our being under rational constraint by the world. What is important is that object-dependence – “the contextual presence of an object itself” (255) - be thought of as “a feature of a thought’s intentional nature” (257) – that the object itself be integral to the thought. A descriptive specification could not do this – it could only add to the contents of an inner realm.

The main source of resistance to his view, McDowell argues, stems from a mistaken idea of what it is to be conceptual. Tyler Burge (2007: 44-65 (1977)) argues that to attribute a *de re* belief, say, to a subject is to relate her to a conceptual, predicative sub-propositional component and a non-conceptual *res*, or object. On this account, the object is extrinsic to the fully specified belief. But Fregean senses are fully conceptual, so on this view they cannot be *de re*. McDowell argues against this position in DRS, tracing Burge's conclusion to a conflation of two meanings of 'concept', between, firstly, concepts as the contents of thoughts or utterances and secondly, concepts as the vehicles of those contents – that is, as words or symbols. If concepts consisted only of vehicles, this would mean that conceptual content must be fully linguistically codifiable, and that would exclude thoughts having *de re* senses as constituents from being conceptual. But if concepts are to be thought of as contents, or what is expressed, such *de re* thoughts can be part of that content alongside linguistically codifiable items. McDowell writes (MW: 107), “the right gloss on ‘conceptual’ is not ‘predicative’ but ‘belonging to the realm of Fregean sense’”. So the *de re* senses of objects, as they feature in thoughts, are conceptual but non-linguistic elements of content, as exemplified in the “logical togetherness” of shape, colour, and location of the red cube which matches the propositional structure of the judgment “there is a red cube in front of me” (WL: 30; II:(i) above).

With the argument that non-linguistic items – Fregean *de re* senses – are part of the space of reasons, McDowell has in place the final component that underpins his account of rational constraint, and the interpenetration of mind and world. Since objects are able to figure in thoughts in this way, we can say that our perceptions of them and their properties are in conceptual form. The Cartesian and neo-Cartesian model of perceptual

experience as mediating between thought and world, with all the problems that, according to McDowell, are associated with that view, is superseded. In its place we have a picture in which perceptual experience can be seen as a *conduit* connecting mind and world, rather as our visual apparatus normally functions not as a barrier but as that through which our visual sense and the world are directly connected. The elements of the conduit are the concepts – Fregean *de re* senses and others through which objects and the world are presented to us. With this argument McDowell shows how the account of our direct contact with the world, through perception of objects, is incorporated the austere picture of sense outlined above.

IV. Conclusion.

In this opening chapter I have tried to provide an overview of the essentials of those aspects of McDowell's philosophy that are most relevant to my task in subsequent chapters. My account has necessarily been almost completely uncritical, even though almost every aspect of McDowell's work has been subject to intense criticism, especially, perhaps, his denial that there is non-conceptual content, his radically anti-Cartesian content externalism, and his anti-reductionist claims on behalf of the space of reasons. Despite his professed aversion to "constructive philosophy" (MW: *xxiii*), the label "constructive" arguably applies to several of the ideas he introduces – the dichotomy of logical spaces, the unboundedness of the conceptual, object-dependent *de re* senses, and the idea of the interpenetration of inner and outer space, for example.

However, when one looks at these aspects of his account in the light of his overall aim, they are seen to be necessary. The removal of the

Cartesian (and neo-Cartesian) barrier between mind and world has, potentially at least, the effect of restoring the idea of the world as primarily *our* world, rather than as, say, primarily the world of fundamental physics. If, for example, some content is non-conceptual, such that we are caused to have it in the realm of law way that McDowell calls *merely* causal, we could not be open to the world in the required way, and the same goes for the disjunctive theory, object-dependency and the other components of the account.

The world of McDowell's picture is the world seen from within the space of reasons – a world that not only contains ordinary objects, colours, sounds, and smells, but also meaning and aesthetic and ethical content. This world has, of course, been here all along, but the dominance both of the scientific picture and of the philosophical notion of the isolated inner realm have tended to make us lose sight of the fact that it is a necessary condition of the existence of both (see Barber 2010: 447-454). Restoring the open relation between our minds and this lived world is therefore, I believe, a supremely worthwhile enterprise.

There are, however, some caveats of which I think we need to take notice. We may accept, with McDowell, that science has nothing to say about the space of reasons. But if philosophical claims made within that space seem to be out of line with intuitions about the physical world which, at least to some extent, owe their prevalence to the credibility of the scientific world-view, then the credibility of those philosophical claims will accordingly be thrown into question. As I explain in the next chapter, I think a problem of this kind arises with McDowell's view of causality, and the rest of this dissertation is devoted to finding the way to resolve this problem on his behalf.

CHAPTER 2. MCDOWELL, DAVIDSON, AND CAUSALITY.

I. General Causality.*(i) McDowell's Naïve View of Causality.*

We have seen that McDowell rejects Davidson's coherentist view according to which experience is conceived as "an alien force, the causal impact of the world, operating outside the control of our spontaneity" (MW: 8). The non-epistemic notion of causality in play here is one that is in sharp contrast to rationality. In McDowell's alternative picture, our experience is not *merely* or *brutely* causal, but is already conceptually structured in a way that makes it capable of contributing rationally and normatively to our conceptual content and behaviour – paradigmatically, in a form that enables it to be the basis of a judgment about the world.

Yet, according to McDowell, when our conceptually-laden experience provides us with the basis for factual judgments about how the world is, or when such judgments furnish us with reasons to act in the world, these processes are also causal. He calls the logical space that contrasts with that of reasons the realm of law rather than of cause because, as he says, "reasons might *be* causes" (MW:71n2). *Prima facie*, then, it looks as if McDowell is acknowledging two different kinds, or at least two different *concepts*, of causality, rational and non-rational, say, or one appropriate to the space of reasons, the other to the realm of law.

In this section I try to clarify McDowell's views on causality in general, mainly by contrasting them with Davidson's. In section **II** I turn to mental or intentional causality, and show how, despite his claim (RN:

92) to “follow Davidson” in holding that “explanations in terms of someone’s reasons are a species of causal explanations”, McDowell in fact interprets that claim very differently. These differences will be seen to contribute to McDowell’s rejection of Davidson’s anomalous monism, which rejection in turn is a reflection of their contrasting views on the place of intentionality in the world. In section **III** I consider the nature of the picture of the relation between intentional and non-intentional causation that might be entailed by McDowell’s views of mind and world.

McDowell persistently refuses to align himself with any particular theory of causality. In his response to Richard Bernstein (RS: 270), he takes issue with Bernstein’s suggestion that ‘we need a deep analysis of causality’, in order to understand how it fits in McDowell’s naturalism of second nature. And when Hessbruggen-Walter (1999) challenges him to say what theory of causality he has in mind in his claim (WL:16) that perceptual conceptual capacities are triggered – i.e., caused – by environmental impacts, McDowell replies:

To this I respond, no doubt frustratingly, that I see no need to embrace any particular theory of causality. The concept of something’s being caused to happen is perfectly intuitive. We acquire it at our mothers’ knees, when we acquire concepts such as those of dropping, breaking, denting, wetting...what is supposed to be conceptually problematic about the idea, in particular, of actualizations of conceptual capacities that occur as effects of environmental circumstances? (RN: 92).

His position, then, is that the notion of cause is a basic, commonsense concept which has no need of further analysis. This is consistent with his similar views on the irreducibility of concepts such as meaning, truth, and normativity (chapter 1). McDowell's thought here becomes clear when one sees that for him all these concepts, including that of "environmental circumstances", belong within the space of reasons. They can only be grasped from within that space, therefore only by an occupant of that space. It would be futile – an example of the sideways-on approach he decries – to try to reach something like an absolute, mind-independent analysis of a concept like cause.

This picture of the origins of our concept of cause in everyday agency echoes the views of Anscombe (1993) and Strawson (1985), who emphasize that our understanding of causality in non-human nature is based on extrapolation from our own everyday experiences of causing things. Our fixation on the term 'cause' in theoretical discussion is, Strawson (*ibid*:120) writes, "a case of premature generalization". What we actually have is a host of special causal concepts which we acquire as we learn to speak. As Fair (1979: 221) writes, 'cause' stands in for a large class of transitive verbs that we can call 'causatives'. In our most basic experience of ourselves as causal agents, in our pushings, pullings, and liftings, causal power is literally felt as physical pressure, or resistance to pressure – thus, on this view, causality is part of our direct experience, and for McDowell all our experience is bound up with our second nature, our status as occupants of the space of reasons. Causality is thus, for him, a *sui generis* and irreducible category, whether it is causation by intentional agents or causation in the inanimate world. In section II I identify what I believe are inconsistencies inherent in this view.

(ii) *Davidson's Nomological View of Causality.*

Strawson is among the many who claim that, despite the origins of our causal concepts in simple agency, we are also inclined to think causality is “a natural relation which holds in the natural world between particular events or circumstances” (Strawson 1985: 115). Of observable examples of mechanical causation, he adds (*ibid.*: 131), “We do indeed suppose there to be exceptionless and truly general mathematical laws which bear on the case”. Davidson’s formulation of the Principle of the Nomological Character of Causality (PNCC); “where there is causality, there must be a law: events related as cause and effect fall under strict, deterministic laws” (ME: 208) is in line with this view. Many regularities are observed in nature, but we can distinguish true causal regularities, such as, say, the relation of a fall in ambient temperature to below zero degrees Centigrade to the freezing of water, from non-causal regularities, such as the succession of day and night (Davidson LC: 205), or the progression of symptoms in a disease. The most common way of marking the distinction is with the claim that causal regularities, unlike accidental regularities, instantiate laws.

Controversially, Davidson claims that, because causality is extensional (see below), the strict law requirement applies to all mental or psychological causation, even though mental causal explanations or descriptions do not allude to such laws. Jennifer Hornsby (SMH: 10) objects that the PNCC claim “is introduced from outside of psychological understanding...and it gives us the sense of ‘physical’ by reference to which we are to understand Davidson’s physicalist thesis”.

As we shall see, this is also at the core of McDowell's disagreements with Davidson. McDowell would join Hornsby in

resisting the idea that the events we recognize in taking a view of minded beings are available to a conception of how things are in nature independent of the minded beings there – of how things are, not naively, but according to, so to speak, objective metaphysics (Hornsby: *ibid*).

Hornsby regards the 'strict law' condition on causality as an example of a McDowellian sideways-on approach, of illegitimately importing a condition on causality from outside, as it were, our standpoint on it.

(iii) Causation, Nomologicality, and Explanation.

I return to McDowell's view of the PNCC below, but first I want to emphasize its role, for Davidson, in underpinning the extensionality of causality, and hence the difference between causation and explanation. For Davidson, causation itself is extensional - a causal event is causal no matter how described. However, very few causal descriptions contain expressions of the strict laws that, according to the PNCC, govern all causal relations. Some singular causal statements (statements using singular terms), such as 'the hurricane was caused by the formation of an area of extreme low pressure' (see ME: 214) imply the operation of a generalization, but even a general statement like 'hurricanes are caused by the formation of low pressure areas' is not strictly lawlike. To find the strict laws that, on Davidson's account, must underpin the truth of both the singular causal statement and the causal generalization, we need to look to the physical laws governing the micro-components of the

systems involved and their aggregate behaviour. We can call an explanation in terms of these micro-components a nomological causal explanation.

For Davidson, then, there is a complex relation between cause and causal explanation. The relata of causal descriptions are, on Davidson's (as on many other) accounts, individual, token *events*. These descriptions, however, as in the case of the hurricane, need not make any mention of strict laws. Furthermore, although one might be tempted to suppose that those descriptions which mention the strict laws – that is, generally, microphysical descriptions – are in some way basic, or privileged over those that do not, Davidson insists that the extensionality of causality ensures that this is not the case. This claim reflects his view of events as unstructured particulars, according to which an event's being the cause of another event is a basic and unanalysable feature. In particular, events are not causes in virtue of *properties*. As we shall see (III:(i)), this is essential to Davidson's account of *mental* causation.

However, explanation is an epistemological notion, relating facts or propositions in ways to which the notion of strict laws need not apply (Davidson CR: 161). Only some of the possible descriptions of a causal event or process are explanatory. In Davidson's own well-known example (ARC: 17), 'the event reported in Tuesday's *Times*' (a hurricane) is given as the cause of 'the event reported in Wednesday's *Tribune*' (a catastrophe). This is a true causal statement, or description, which is not an explanation. Moreover, we saw above that true causal descriptions may or may not be given in terms of strict laws, and explanatory descriptions can fall under either of these categories. One type of causal explanation is formulated in terms of the operation of the

strict laws in the particular case, and is thus deductive-nomological (D-N), in that the occurrence of the particular event is deductively explained by stating the initial conditions and subsuming the event under the covering laws. Other kinds of causal explanation, in contrast, need make no reference to laws, and such explanations are typically singular, as when we say that the hurricane was caused by an area of low pressure. But for both D-N and singular descriptions to be explanatory, the descriptions of the causally related events must be capable of conforming to a pattern of causal relations involving events of the appropriate type. This is straightforward in D-N explanations, in which the pattern is provided by the covering law. But the causal description of the occurrence of the hurricane must also, to be explanatory in the way we want, relate the events described to a causal *pattern* involving hurricanes and low pressure areas as types. The statement that low pressure caused the hurricane is only explanatory against a background in which the possibility of such a causal relation makes sense. To be explanatory, that is, descriptions must relate events as types.

Causal explanation thus has much richer resources than would be the case if it were confined to D-N explanations based on the PNCC. I return to the relations among different kinds of causal claim in chapter 4. I now turn to causality in the philosophy of mind, arguing that McDowell's professed agreement with Davidson on the question of reasons and causes masks deeper differences between their views. These differences, however, need to be understood against the background of Davidson's solution to the problem of mental causation.

II. Reasons and Causes.

(i) The Problem of Mental Causation.

Before discussing Davidson's solution to the problem of mental causation, I will briefly set out the problem's general form, since it will be prominent in later discussion. It has become standard to present the problem in the form of a number of plausible assumptions that turn out, as a group, to be mutually inconsistent (see e.g., Crane 1995; Kim PM, MPW; Baker 2007 (MEL); Bennett 2007; Vision 2011). The following formulation is due to Jens Harbecke (2008: 16-31). On Harbecke's account, four premises form an inconsistent tetrad:

(MC) Mental events cause physical events.

(CP) The realm of the physical is causally complete.

(NI) Mental events are not identical with physical events.

(NO) Physical events are not pervasively, or systematically, causally overdetermined.

(MC) is taken by many to be absolutely central to our concept of ourselves and our place in the world. Burge (1993: 97) argues that our reasons for believing that what we do makes a causal difference in the world are much stronger than any possible metaphysical argument to the contrary, while Fodor (1989: 77) famously comments that if this is false "practically everything I believe about anything is false and it's the end of the world". (MC) is arguably the non-negotiable premise for which the combination of the other three creates a problem. Although as stated (MC) is about mental to physical causation, it may be taken to make the same claim about physical to mental.

(CP) is often expressed more explicitly as ‘any physical event that has a cause at t has a complete physical cause at t ’. The principle expresses another of our deeply held intuitions, based both upon science and, as I shall argue in chapter 4, upon our understanding of how the everyday objects around us behave. (CP) is couched in terms of completeness rather than closure so as not to rule out the possibility of systematic overdetermination without considering it separately (Harbecke *ibid*: 19). Harbecke (*ibid*: 24) points out that a consequence of (CP) is that “if the mental is causally efficacious with respect to the physical at all, it cannot be so by exerting additional fundamental powers onto the physical”. This hints at the possibility, which I exploit in chapters 4 and 5, that mental causes may be of a distinct kind.

(NI) is rejected by type physicalists, who believe that mental *properties* have causal effects through being identical with physical properties, and by eliminativists, who believe there are no mental entities. The commonest objections to type physicalism are the possibility of multiple realization of mental properties by divergent types of physical properties, and the claim that mental content is externally individuated. I will not discuss type physicalism or eliminativism further. Many philosophers of mind are convinced by our intuition that mental properties, events, or both are neither identical with nor reducible to physical properties or events.

(NO). Systematic overdetermination, such that every time an agent affects the physical world there are two causes, a mental and a physical, is regarded by many as ontologically unacceptably extravagant. Perhaps the strongest argument against it is that it implies that, assuming the two causes are both sufficient, that if one of them were absent the effect would still occur. This is implausible, especially if we imagine the

physical cause being absent, since it would seem to imply some kind of telekinetic causation (Kim MPEE: 247).

It is a straightforward matter to show that the four premises cannot all be true. For example, if (MC), (CP), and (NI) are true – that is if mental events have distinct effects, yet the physical is causally complete - then the effect must be doubly caused, and (NO) must be false. If (NO) is accepted, then given (CP) and (NI), mental events, being distinct, must be causally inert – (MC) is false. Acceptance of (MC), (NI), and (NO), on the other hand, means that mental events must break into the supposed completeness of the physical realm. Finally, if the physical *is* complete, we can only have non-overdetermined mental causation if (NI) is false, which amounts to an acceptance of type identity.

I discuss the problem of mental causation below and in later chapters using the different terms in which it has been formulated by, for example, Davidson and Kim, but in each case it can be reformulated in the terms of Harbecke's four premises. For example, if McDowell's views on agent causation ((iii) below) are expressed in terms of the four options that each follow from rejection of one of the premises, it would seem that he rejects (NO); he agrees that the physical is causally complete (personal communication, March 2012) and, as we shall see ((iv) below), that "that intentional items are causally efficacious in their own right" (RAB: 69), which asserts (MC) and implies (NI). But I do not think McDowell himself sees his view as involving overdetermination. It is more likely that he agrees with Hornsby in denying that there *is* a problem of mental causation.

(ii) *Davidson's Causal Arguments.*

Davidson's account of mental causation begins with his arguments for the causal theory of agency in 'Actions, Reasons, and Causes' (ARC, 1963). We can think of these as aimed at establishing that reasons (or, more accurately, the events that are the onsets of reasons, or of subjects' coming to have reasons) *must* be causes, while the argument for anomalous monism in 'Mental Events' (ME, 1970) then purports to show *how* they can be causes. Two arguments for the causal theory can be distinguished in Davidson's account in ARC; one argues from the need to distinguish between competing reasons, the other from the physical character of action in the world and the causal closure of the physical. The first is well known: it can be tempting to think that when someone acts for a reason the pattern of rational justification that explains the action is sufficient to explain why the action occurred. But it is not, "for a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it. Central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the reason." (*ibid*: 9). There is a second argument, expressed by Child thus: "(i) an action explanation is an explanation of *why* something happened: but (ii) no non-causal explanation can explain why something happened; so (iii) action explanation must be causal" (Child 1994: 92, italics added). So even if there was only one possible reason for an action, we still need to invoke cause, if we accept Child's claim (ii). In that claim "something happened" refers to an event, considered as a non-intentional item occurring in the physical world (see I above), and implicit in the claim is the assumption that physical events have complete and only physical causes (see Kim and the Physical Causal Closure Principle, chapters 3 and 4). So the problem boils down to that of how a reason, an intentional item, gets transformed into a physical change in the world, a non-

intentional item. Appealing to further intentional items, such as justifications, cannot help here, the argument goes. The potential action remains trapped, so to speak, at the level of intention, until we invoke cause.

So we have an argument that reasons *must* be causes. If this can be backed up by an account of *how* reasons *can* be causes, that is, a theory of mental causation, we can explain why, and how, certain mental events cause physical actions to be performed in the non-intentional world. Davidson provides an account of mental causation in the form of the well-known token identity theory (Davidson ME), and the conclusion of the argument just discussed, that reasons are causes, more broadly expressed as the Principle of Causal Interaction (PCI) (*ibid*: 208), forms one of the three principles from whose apparent inconsistency he argues for that theory.

For McDowell, the PCI “seems unquestionable” (FAM: 339). In a footnote, he adds that “it would be very strange if the causally interconnected psychological systems I envisage were causally disconnected from the “physical” world” (*ibid*: n16). As the discussion above suggests, his naïve view of intentional causation allows him not to acknowledge the problem that Davidson’s anomalous monism is an attempt to solve. It is just at this point, I think, that a gap opens between what Davidson understands, and what I believe McDowell, given his general account, *should* understand, by the PCI.

(iii) *Anomalous Monism.*

Davidson's account of mental causation (ME) is based on token identity of mental and physical events. It is presented as a solution to the apparent mutual inconsistency of three principles that he takes to be established – PCI, PNCC, and a third, the Principle of the Anomalism of the Mental (PAM), which asserts that “there are no strict deterministic laws on the basis of which mental events can be predicted and explained” (*ibid*: 208). In support of PAM and the radical disparity between physical and mental descriptions, Davidson argues (*ibid*: 213-223) that the two kinds of descriptions are heteronomic; that “no purely physical predicate, no matter how complex, has, as a matter of law, the same extension as a mental predicate” (*ibid*: 215). Given the PCI, that the mental events cause physical events (and *vice versa*), and given the PNCC, that “events related as cause and effect fall under strict deterministic laws” (*ibid*), it appears that PAM must be rejected. Davidson's solution, accommodating all three principles, depends on the extensionality of causality. Each mental event, he claims, has a physical description, since it is also (*i.e.*, is identical with) a physical event. All physical events, including those that are also mental, are bound by the PNCC, and of course, being physical, they interact with other physical events (PCI). However, if each token of a particular type of mental event was always identical with a token of the same type of physical event, relations among mental predicates (Davidson avoids ontological commitment to mental *properties*) would share the patterns of causal interaction of those physical events, thus violating PAM. The suggestion, then, is that each token mental event is identical with *some* token physical event, and that this identity is unsystematic. In Harbecke's terms, Davidson solves the quadrilemma by denying (NI), non-identity, for mental event *tokens*.

For Davidson, while causality is extensional, nomologicality is intensional. So, while each physical/ mental event has a physical description according to which it is an instantiation of a strict causal law, it also has a mental description. This separation of cause from explanation or description allows the claim that descriptions of events as mental which constitute intentional explanations do not allude to strict laws, but instead make use of quite different patterns of parameters – those of what Davidson calls the constitutive ideal of rationality (*ibid.*: 223). These patterns exhibit the same attributes that characterize what Sellars and McDowell call the space of reasons – normativity, rationality, and holism.

(iv) McDowell's View of Reasons and Causes.

A crucial point at which Davidson's and McDowell's views diverge is that McDowell does not accept the universal nomologicality of causality, regarding the PNCC as a scientific prejudice. As we saw, he claims to take Davidson's line: "I follow Davidson's 'Actions, Reasons, and Causes'...in taking it that explanations in terms of someone's reasons are a species of causal explanations." (RN: 92), but for McDowell, unlike Davidson, "we need not see the idea of causal linkages as the exclusive property of natural scientific thinking" (NPM: 92) – that is, as exclusively nomological.

The most complete expression of McDowell's position on reasons and causes is contained in a passage in his response to comments by Bilgrami (RAB: 69):

It is true that I think Davidson's monism is a mistake...the way to eradicate [Davidson's residual naturalism] is not...to drop the very idea that intentional items belong to a causal nexus...The right move is to drop the idea that for intentional items to belong to any causal nexus at all is for them to belong to "the causal nexus that natural science investigates" in a way that would need to be spelled out by redescribing them in non-intentional terms. That is the idea that drives Davidson's argument for his non-reductive monism. When Davidson argues that the understanding provided by ordinary explanations of action is causal understanding, he says enough to display intentional items as belonging to a causal nexus in their own right...The naturalistic picture of *the* causal nexus that underlies Davidson's monism stands in tension with the idea that intentional items are causally efficacious in their own right...we should extract the idea that intentional items are causally efficacious from the setting Davidson puts it in, the monistic picture of *the* causal nexus (*ibid*).

Clearly, then, McDowell takes reasons (as "intentional items", whose exact nature is left unclear; we can safely assume that McDowell does not mean to reify notions such as desire and intention) to be causes, but his claim that they are so "in their own right" distances his view quite profoundly from Davidson's tying of causality to physical nomologicality. McDowell's rejection of Davidson's "residual naturalism" constitutes, amongst other things, an outright rejection of any role for the PNCC in intentional causation.

McDowell's position seems to be that, while he "follows Davidson" in the claim that 'reasons are causes', he rejects the theory – anomalous

monism - that Davidson advances to account for *how* they can be causes. For McDowell himself, the need to provide a metaphysical basis for the interaction of mental and physical events, which motivates Davidson, does not arise. Recall (1:II:(iii)) that McDowell rejects the idea of an interface between mind and world – for example, there is for him no ontological gap between true thoughts and worldly facts.

McDowell's references to the "causal nexus" are interpretable as implying that the causal nexus that natural science investigates and the causal nexus to which intentional items belong are distinct. If so, this would entail that there is no causal interaction between the two. So events belonging to the intentional causal nexus would not causally affect events belonging to the other – call it the nomological - causal nexus, and *vice versa*. McDowell himself is clear, however (email correspondence January 2012) that there is only a single causal nexus.

In his response to Graham Macdonald McDowell writes:

I do not imply that there is *no* room in nature...for conformity to law... [but]... I am quite happy to suppose there are two kinds of happenings in nature; those that are subsumable under natural law, and those that are not subsumable under natural law, because freedom is operative in them (RGM: 238).

He goes on to claim that this is a distinction, not a dualism, because it does not imply a dualism of substances. As it stands, however, the claim seems to imply that, say, a free human action that results in a change in the spatial location of an everyday object, which McDowell certainly accepts is *causal*, is nevertheless not *nomologically* causal, while the

very same object, at the same time, is also subject to other causal influences such as, say, gravity, which *are* nomologically causal.

Bilgrami (2006a: 244ff) points out the inconsistencies of this position. It implies rejecting the claim that the PNCC applies to intentional causation while acknowledging that nomologicality is a feature of our causal understanding of the subject matter of physical science. Yet McDowell's position is that the same everyday notion of cause is operative in both situations. As Bilgrami writes,

[I]t is odd to think of [the notion of cause that is operative in agency] as the same as the one in the Humean picture, only minus the Humean element of implied generality or nomologicality...[or] to think that when dispositions happen to be mental or intentional dispositions, they just simply do not integrate with any other dispositions of any other science (*ibid*: 261).

We should bear in mind, however, that for McDowell the concept of cause that matters is the everyday one of causation as occurring within the space of reasons (I:(iv) above). Under that concept, that we cause happenings in the world through our conscious agency is an irreducible and *sui generis* fact. The fact that some causal happenings are describable in nomological terms is of secondary importance and does not threaten the core meaning of 'cause', which is primarily epistemic and inseparable from that of causal *explanation*. So when McDowell writes of the causal nexus and "happenings that are not subsumable under natural law", he means also the causal-*explanatory* nexus. This would allow the idea of a causal nexus to be understood in a relaxed

way, so that it can include both nomological causation and causation by intentional items. The idea would be that once the notion of cause that comes from ‘objective metaphysics’, as Hornsby puts it (I:(iii) above) – that is, the PNCC conception - is dropped, there is no need to mark a significant distinction between cause and causal explanation.

For McDowell, the world we concept-users inhabit is a world of intentional relations – of true thinkables. Our interaction with the world is just as fully conceptualized in agency as it is in perceptual experience (MW: 89 – 91; see 1.II:(i) above)); “intentions without overt activity are idle, and movements of limbs without concepts are mere happenings, not expressions of agency” (*ibid*: 89). What this implies for mental causation is that, on McDowell’s account, as opposed to Davidson’s and indeed any physicalist account, an intention does not have to be identical with a physical item, that is, one that is conceived as occupying a different ontological realm, in order to make a difference in the world beyond the mind, since that world is not beyond the thinkable. Intentions and actions are both within the space of reasons, both “intentional items”. So an intention to act does not *need* to be thought of as sharing its identity with a physical event in order to, so to speak, cross a gap to interact with another ontological realm. The realm of the conceptual is unbounded (chapter 1): – there *is no* such other ontological realm.

Thus there is no need, on this picture, for an intentional item to have a physical description. On the contrary, as McDowell sees it, intentional items, being native to the space of reasons, are *sui generis*; they are not visible at all from the physical standpoint, and have no physical characterization. Intentional items, perhaps, are the patterns that only those who resonate, as McDowell would say, to the space of reasons – to

meaning -, are able to discern in the world. This is compatible with the common sense observation that an enabling physical basis – a person with a brain – is a prerequisite of intentionality, as McDowell concedes. But, firstly, as Davidson would agree, there is no question of *intentionality's* being redescribable in physical terms, if by intentionality we mean the patterns of relations that constitute intentional *content*, and, secondly, and in contrast to Davidson's account, the term 'physical' here is to be understood as capturable within the scope of the space of reasons.

(v) The Problem of Mental Causation Within the Space of Reasons.

But, I will insist, despite the strategy of enclosing causality within the space of reasons, Hornsby's and McDowell's shared position on intentional causation remains inconsistent. The position Hornsby adopts in her essay "Causation in Intuitive Physics and Commonsense Psychology" (CPCP, 1992) is, I believe, largely shared by McDowell. Hornsby compares the neuroscientific approach to commonsense psychology with eliminativist accounts of the existence of everyday objects. The "atomic physicist" who argues that tables, as such, do not exist because only microparticles and empty space really exist, and the eliminative materialist who argues against the existence of propositional attitudes are both, according to Hornsby, mistakenly looking at their subject-matter from an inappropriate standpoint – one from which neither tables nor persons come into view. Rather, commonsense psychology is the appropriate standpoint from which to view persons and their attitudes, while objects like tables are best viewed in the framework of "intuitive" as opposed to scientific physics. Furthermore, Hornsby claims, the frameworks of commonsense psychology and intuitive

physics share a common notion of causality; “the everyday physical and the everyday mental are equally parts of a single causal world view” (*ibid.*: 186). This seems to describe essentially the same naïve notion of causality that McDowell espouses.

Intuitive physics, as Hornsby presents it, describes the naïve conception of a physical world that children acquire in normal development, concomitantly with, and inseparably from, their acquiring an understanding of commonsense psychology. In McDowell’s terms, it is part of the pre-philosophical world picture of concept-using human beings who have acquired second nature; in Sellars’s (PSIM) term, it is part of the original image of man in the world, which preceded the manifest image. Thus intuitive physics is a concept of the physical world as viewed *from within*, in the sense that it belongs to a way of experiencing and interacting with that world as something that is not separable from our mental lives, but rather as, perhaps one could say, an *extension* of our mental lives or of ourselves, hence the common conception of causality that Hornsby identifies. It is above all a concept of the world from a *personal* point of view. Hornsby contends that it is only when we adopt the distinct standpoint of a science-influenced metaphysics, from which personal agency does not come into view, that the problem of mental causation arises. As Kim (MPW: 59), opposing this view, puts it, the idea is that “worries about mental causation arise out of our misplaced philosophical priorities”. Hornsby writes elsewhere (SMH: 12), “The world in which mind is accommodated by the naïve naturalist is naïvely natural...no peculiarly scientific method is required to have knowledge of it”.

However, one can argue, also without recourse to any peculiarly scientific method, that the naïve view contains tensions that are not easily resolved. We only need to move a small distance away from the prelapsarian state of immanent involvement with the physical world that Hornsby describes in order to reach a more sophisticated standpoint, one from which we can distinguish ourselves, seen as persons among others, from objects like tables. This new view is that of Sellars's manifest image, which "emerges as a result of the gradual depersonification of the original image" (Tuomela 1985: 10). It is important to emphasize that to say that objects are distinct from us is not to imply that they are 'things in themselves', outside the conceptual sphere, or the space of reasons. It is simply to say that they are objective in McDowell's sense of being outside thinking, but not outside the thinkable (MW: 28). Beliefs and judgments about them are true or false independently of what anyone believes or judges.

It appears, then, that Hornsby's depiction of intuitive physics combined with commonsense psychology does not prevent us from seeing that there is also a sense in which "the everyday physical and the everyday mental" are *not* "parts of a single causal world view". So it is not "scientific physics" alone that intuitive physics should be contrasted with. There is also a contrast with a view of the inanimate world that is only slightly more sophisticated than that of intuitive physics, one that almost everyone, and not just the scientifically sophisticated, can readily adopt. Here it is important to distinguish between the space of reasons / realm of law contrast and our understanding of the contrast (1:I:(ii)). The behaviour of the inanimate world when left to its own devices belongs, of course, to the realm of law. Our ordinary, pre-scientific understanding of the inanimate world, however, as a "depiction of nature" (MW: 70n1)

belongs in the space of reasons. What I am arguing is simply that this is a depiction that is part of the manifest, rather than the scientific, image (see chapter 4).

This loss of innocence is, I think, all that is required to make us susceptible to the force of the arguments (see *(ii)* below) that lead to worries about mental causation and, more generally, the place of our subjectivity in the world. We realize, for example, that the only kind of causal interaction possible between human beings and the everyday physical world is ultimately that which occurs through physical force exerted by our muscles and the action of various forces upon our bodies.

Causality itself must, I argue, share in the objectivity of the world with which we interact. To reject the PNCC, as Hornsby and McDowell do, is to reject, for causality, the kind of objectivity that, impossibly on their view, encroaches from beyond the thinkable. But this does not mean there are not objective facts about causality in the acceptable sense of ‘objective’. Consider the facetious suggestion (due to the comic poet Ogden Nash (1983: 133)) that the wind is caused by trees waving their branches. This is a causal explanation, couched entirely in the realm of law mode, which is plainly false. On the question of what makes it false, it is not open to McDowell to say, for example, that the distinction between true and false explanations of what causes the wind is based upon that which will command community consensus. As we saw (1:II:*(iii)*), in TRO he argues against Rorty’s espousal of this very view and in favour of there being a real difference between what one’s peers assent to and what is the case. I think McDowell’s position must entail, then, that behind the correctness or incorrectness of a causal explanation, there is something else, an actual causal fact, for instance the fact that it

is atmospheric pressure gradients, rather than trees' moving, that cause the wind. It is in virtue of this fact that the causal explanation is correct. But McDowell's claim that events that are and events that are not subsumable under natural law are both part of the causal nexus cannot then be interpreted as just a claim about how we deploy a *concept* of cause deriving from our practice of causal explanation. It has to entail that both nomological and non-nomological causes act together in the same world. Thus I think we are forced to the conclusion – my central claim in this thesis - that McDowell's position on this point is inconsistent. I have suggested that it is unlikely that he would think that ordinary causal relations consist of the successive actions of a mixture of kinds of causes upon the same objects, now nomological, now not – as Ansgar Beckermann (2001: 64) puts it, this would make physics “into a science with some general laws and a whole bunch of exceptions...a *weird* science”. As we know, McDowell is a leading critic of the myth of the given (see 1:II:(i)), yet in defending the naïve view of intentional items as causes in their own right, he appears to support an analogous myth in the philosophy of action.

McDowell writes of the action of raising a glass:

I don't think it's particularly problematic that some happenings (e.g. the glass getting to be in a certain position) figure both in the scientific image and in the image that traffics in concepts of agents and their doings. There doesn't seem to be a particularly pressing need to suppose that the microphysical explanation of the glass's getting to be where it gets to be would exclude any space for an explanation in terms of the fact that I raised it (email 4.5.12).

But again, if I am right about the objectivity of causality, and of the PNCC, *within* the space of reasons (I:(iv)), then the change in the position of the glass not only has two explanations, but, apparently, has two causes. As we shall see, McDowell's view of the world as within the space of reasons allows him to give a coherent account of intentional causation *as such*, but the very coherence of the account raises other issues. In particular, he says nothing about what he calls the physical (causal, realm of law) explanation of the glass's movement, other than simply that there is one. But if causality is objective in the sense I have suggested, then on the account McDowell gives it is true both that the glass changes position because he raises it and that it moved because of the operation of physical forces. Both common sense and science give us every reason to accept the second of these, but this leads to puzzlement. If we accept the first causal account, what is the relation between the two? I return to this issue in Chapter 4:I:(iv).

III. The Space of Reasons and the Physical.

(i) Problems with Anomalous Monism.

I have suggested that McDowell's attempted replacement of Davidson's token identity theory of mental causation with one based on the idea that "intentional items are causally efficacious in their own right" does not succeed because his naïve concept of causation runs up against a commonsense, rather than an inappropriately scientific, notion of physical causality. However, Davidson's theory has other difficulties that are decisive for many who are more sympathetic than McDowell is to Davidson's token physicalism. The most widely perceived defect of anomalous monism, pointed out by McDowell himself (MW: 75-76) and

many others (see, for example, Honderich 1982, Follesdal 1985, Stoutland 1985, McLaughlin 1985, 1993, Kim SMK, Child 1994), is that while it may secure the causal efficacy of the mental, it does not thereby secure its causal or explanatory relevance. Suppose, say, John goes to the fridge in order to get a beer. This single action (event) has (1) potentially, a physical description in which complex physical events instantiate strict causal laws, and (2) an intentional or psychological description involving familiar generalizations relating beliefs, desires, reasons, and actions. Both of these causal descriptions can function in causal explanations; in (1) it will be a D-N type of explanation which, if we knew them, would cite the strict laws governing the particular physical events, while in (2) it will be in terms of the belief/desire generalizations. There is a natural tendency here (II:(iii) above) to take the description that figures in the type of explanation in (1) to be the basic description, which captures and explains the real causal efficacy of the event. Stoutland, for example (1985: 53), writes, “But on Davidson’s view an agent’s reason *cannot* cause his behavior in the right way because it cannot cause it *qua* reason; it can cause it only *qua* physical event since it is only in virtue of physical properties that events are causes”. However, this distinction between *qua*, or in virtue of, reason and *qua* physical event is just what Davidson (e.g., TC: 13) rejects. For him, because of the extensionality of causation and the nature of events as unstructured particulars, no description, even one that actually cites the strict laws that are assumed to be present, enjoys a privileged status.

McLaughlin (1993) argues, to the contrary, that Davidson’s first claim, that causality is extensional, is quite compatible with an event’s causing another in virtue of some, but not other, of the event’s properties (or perhaps descriptions), and hence with the rejection of Davidson’s claim

that no description is privileged. Davidson's error, according to McLaughlin (*ibid*: 33), is to conflate the claim

Event *c* causes event *e* in virtue of *c*'s having *F*

with the claim

Event *c*'s having *F* causes event *e*.

That is, an event's causing something *through* its having a particular property is confused with the event's *possession* of the property's being the cause. But '*c*'s having *F*' is a state of affairs, not an event, and not the sort of thing that can be a cause. That *a* weighs less than *b* in virtue of weighing 10 pounds, McLaughlin points out (*ibid*: 34) does not imply that *a*'s weighing less than 10 pounds weighs less than *b*. "But if *b* weighs 11 pounds and *a* weighs 10 pounds, then *a* weighs less than *b* in virtue of weighing 10 pounds" (*ibid*), and this is so under any description of *a* and *b*. The only way of escaping this conclusion, McLaughlin argues (*ibid*: 32), would be to claim that that *c* and *e* are causally related is a brute fact, which seems "quite implausible". So, by analogy, 'John's desiring a beer caused John's going to the fridge in virtue of its (the desiring's) being identical with a physical event' is a valid claim, while 'John's desiring a beer's being a physical event caused John's going to the fridge' is not. This criticism seems in line with our causal intuitions, and if it is right the objection that the token identity account fails to preserve the causal efficacy of content *qua* content stands.

(ii) *The Biofunctional Analogy.*

In an attempt to resolve the question of causal/explanatory relevance, Macdonald and Macdonald (1986; 1995b; 1995c; 2006; 2010:149 -154: 2011; G.Macdonald 2007),⁴ develop a metaphysics of properties and events which shows how different properties – mental and physical – can be *co-instanced* in one and the same event. This involves adopting a different ontology of events – as property-instantiations, or instantings, in objects at times, instead of Davidson’s construal of events as unstructured particulars. This account of events is broadly Kim’s (EPE), as refined by Lombard (1986, 1998) and C. Macdonald (2005), according to which an event is an exemplifying of a property in an object at a time, which, for property P, object x, and time t, can be designated {Pxt}. We can distinguish between two ways in which events are associated with properties. An event’s *constitutive* property is the property P in the above formulation, and as such is a property, not of the event, but of object x. The event itself, on the other hand, has the *characterizing* property of being the kind of event it is, and the key to the Macdonalds’ account is the claim that an event may instantiate more than one such property. And while mental and physical *properties* are distinct, so that something could be true in virtue of one property type rather than the other, this distinction disappears at the level of *property-instantings*. Since, when properties are co-instanced, the relevant event then *is* an instancing of a mental property, the causal efficacy of the mental is ensured.

This account, however, still will not suffice to account for mentalistic *explanation*, because explanation essentially reaches beyond individual instances of causation. The causal and explanatory relevance of the

⁴ For the remainder of this section ‘1995b’ and 1995c’ will refer to the essays by Macdonald and Macdonald in their 1995a, eds.

mental is a separate issue. As the Macdonalds point out (1995b: 61), the relata of mentalistic or intentional explanations are events under those descriptions, that is, *qua* having mental properties. Causal efficacy is a feature of property-instances; the latter are, as they put it (Macdonalds 2011: 15), the worldly items required for actual causation to occur. Causal and explanatory relevance, on the other hand, are in virtue of properties, and the causal efficacy of property instances is insufficient for the causal relevance of those properties (1995b: 67). When two properties are co-instanced on an occasion, there is no implication that they are co-instanced in general (1995b: 68 –70), that is, that they conform to a *pattern*, which is what would be required to make co-instancing explanatory (see I:(iv) above).

Some patterns, such as those that would figure in a *physical* explanation of, say, John's deciding and going to the fridge to get a beer, are strictly nomological. The intentional explanation of this event, however, requires appeal to a quite different pattern, that constituted by a distinct, rational "network of relations between properties" (1995b: 70). The Macdonalds use the example of functional biology (1995b: 71; 1995c: 103 - 104) to illustrate how natural selection gives rise to the emergence of new *functional* patterns in nature. Over time, the regular recurrence of certain physical causal patterns brings about their own perpetuation through reproductive advantage, resulting in the development and establishment of new, stable, but not strictly nomological, functional causal connections. Thus novel causal patterns become apparent, superimposed upon but distinct from the underlying physical-nomological patterns, which make available different explanations.

As an illustration of the distinctness of biofunctional patterns, Fred Dretske (2004) gives a partly imagined example from biology. A plant in the phlox family, Scarlet Gilia, has petals which change colour from red to white each year in mid-June. This change is explicable by evolutionary theory, in that the plant has a different pollinator, attracted to white rather than red, after midsummer, so that in the past plants which by random mutation made the colour change acquired a selective advantage. Dretske asks us to imagine a molecular twin of Gilia which undergoes exactly the same colour change in midsummer, but this time the evolutionary explanation is the regular arrival at the crucial time of rapacious beetles which are only attracted to red petals. We assume that the physicochemical causal processes, bound by physical causal laws, are identical in the two cases, which will thus have the same physical, or causal-nomological, explanation. But the evolutionary explanation is clearly quite different, essentially involving a difference in the *history* of Gilia's eco-system, despite the fact that the difference leaves no trace in the structure or present behaviour of the two physically identical plants. So here we have a hypothetical, but surely possible, example of a situation in which two different histories have given rise to different functions, one of maximizing pollination, the other of discouraging predators, which, however, are instantiated in identical types of physical structures. As the Macdonalds write, "we have a case of different causes, from the biological perspective, of [processes] which would be typed as identical from the physical perspective" (1995b: 72). Functional and physical properties, on this analogy, are co-instanced in *events* in Gilia, thus ensuring the causal efficacy of both, while the functional properties' conforming to the functional patterns in nature ensures their causal relevance in functional explanation.

The historical explanation – involving the change of pollinator, in the case of the real Gilia - has several important features. Firstly, we are inclined to say that the explanation is clearly causal in just as robust a sense as is the physical explanation. But again, if we accept Davidson's stipulation that "events related as cause and effect fall under strict, deterministic laws", it appears that we will not find those laws in the historical causal explanation. If this *does* involve laws – the laws of natural selection, say – they are not strict in the required sense. Each historical explanation is of a single evolutionary event, rather than a repeatable regularity, and even if we subsume these events under types we at best have rough generalizations, which must be supplemented by *ceteris paribus* clauses. So if, as Davidson would insist, there are strict laws in the vicinity, they must be those that are invoked in the physical explanation, which makes no mention of evolutionary history.

Secondly, the historical explanation brings in extrinsic, or relational, features of the events under discussion. The historical facts do not figure in a description of the events that uses only non-relational predicates, of the kind that would be relevant to a physical explanation. Yet the historical sequence of events gives every sign of being real, so that the causal explanation that we give it seems correct, again in just as robust a sense as in the physical cause and explanation.

The price of the Macdonalds' co-instantiating account, if it is one, is that in the intentional case it may seem to deny causal efficacy to functional or semantic *content*, which is individuated extrinsically or relationally with respect to the subject's internal states (1995c: 87), so as regards causal and explanatory relevance, it seems doubtful that the account has any advantage over Davidson's. The separation of cause and explanation

assigns to content an explanatory, but not a direct causal role, conflicting with our intuitions that, for example, when John goes to the fridge for a beer it is his wanting a beer, *qua* wanting a beer, that causes his going to the fridge. But this account may require that our intuitions be overruled. The Macdonalds' answer to the objections of Stoutland and McLaughlin is that causal *efficacy* is not *qua*, or in virtue of, anything, being simply a relation between instances. 'In virtue of' enters the picture only in explanation, and here the existence of 'higher level' patterns ensures causal and explanatory relevance. Where causation as such is concerned, there is no higher level, and "at the level of instances, the world is flat" (Macdonalds 2010: 156).

As the Macdonalds see things, then, functional and intentional explanation share many features. The causal relevance of intentional properties would clearly be achieved through their conforming to a *rational* pattern or patterns in nature (1995b: 7 –73; 1995c: 103-105). These patterns would be thought of as autonomous with respect to the causal-nomological patterns exhibited by physical properties, and as irreducible to them; "intentional properties have a contentful nature which is not exhausted by their causal powers...in virtue of [which] such properties exhibit a rationalistic pattern, or network of relations among themselves" (1995c: 103-104). The defining attributes of the pattern will be those of the constitutive ideal of rationality, or of the space of reasons – normativity, holism, consistency and coherence. And the *patterns* are irreducible despite the fact that their causal powers are exhausted by those of the physical properties with which they are co-instanced (*ibid*).

McDowell's non-reductionist picture of the world is actually well suited to accommodate the rational patterns that the Macdonalds describe. Like

values and secondary qualities (McDowell, MVR: essays 3–10), rational patterns, we could say, are *there anyway* in the world (recognition-independently, we might say), which is within the space of reasons; but as the space of reasons is essentially such that it only comes into view for concept-users, it is only for concept-users that the patterns are visible. The same is true, of course, for causal-nomological patterns – they also come into view when we adopt the scientific stance, in which nature appears as disenchanted. Biofunctional patterns seem to occupy an intermediate position. When we adopt the biofunctional, rather than the causal-nomological stance, those thought contents and facts include *normative* concepts, reflecting the normative nature of the phenomena – not in the full ethical sense of ‘normative’, but in a sense which at least allows us to recognize purpose and to distinguish successful from unsuccessful functioning. We might say that on the biofunctional view the disenchantment of nature is less complete than on the causal-nomological.

But as we have just seen, the Macdonalds think of higher patterns as non-causal, so from McDowell’s point of view, it looks as though the Macdonald account will, in the end, not do, because, even with co-instantiating, it retains the divide between rational patterns as explanatory and physical events as causally efficacious. It does not provide what McDowell wants, an account of “intentional items” as causally efficacious in their own right.

(iii) Biofunctional and Mental Events.

The biofunctional analogy can shed light on questions raised by a further objection to anomalous monism, independently presented by William

Child (1994: 80-89) and Tim Thornton (1998: 197-204). This objection is based on the claim that commitment to token identity implies commitment to an isomorphism between causal-nomological and rational processes, in a way that is incompatible with psychophysical anomalism (principle PAM). The conclusion is that we either must reject PAM, in which case we are led towards a reductive type-type physicalism, or, if we accept PAM, we must accept that we are left with no explanation of how the causal role of a mental/physical event, described physically, matches its normative role when described intentionally (Thornton *ibid.*: 199). Both authors conclude that this means that token identity theory must be rejected.

According to anomalous monism, the argument goes, each mental event is identical to a physical event, and the relation of each such event to preceding and subsequent events is describable as falling under a strict law or laws. So mental events, under their physical descriptions, are interconnected causally-nomologically, and chains or sequences of mental events, under this description, are explainable deductively, on the D-N model. But under their intentional descriptions, each mental/physical event is connected to the previous and the next by *rational* links in ways that, given mental anomalism or the irreducibility of intentional patterns, cannot be captured by a D-N account. As Child argues (*ibid.*: 81–82), “[if] a single mental state is to be a physical state, even on an occasion, then something stronger must be true – that there is a general isomorphism between a person’s mental states and her physical states”.

If each mental event is identical with a physical event, then according to the objection, we are forced to admit that the purportedly irreducible

semantic and normative patterns instantiated by the sequence of mental events in a chain of reasoning *mirror* the causal-nomological pattern or structure instantiated by the same events under their physical descriptions. Thornton writes:

The problem is this. If the dispositions of physical states described in causal terms do not *define* or *set* the normative standard [as would be the case if reductionist type physicalism was true], then they must instead *conform* to that standard. Thus, if a token mental state is identified with a token physical state, the causal role of that mental state when described physically must match its normative role when described in intentional terms... [Yet] in the absence of psychophysical laws, there is no explanation of this harmony (*ibid*: 199-200).

I think the claim of necessary isomorphism is refutable, given Davidson's metaphysical assumptions. Davidson's argument was that events are causal under any description, while those physical events that are also mental have a further, distinct description. Relations between events of this kind are describable according to two distinct patterns, one causal-nomological, the other normative and rational. But on the Davidsonian account these are intensional *descriptions* of patterns of relations, neither of which, as descriptions, need exclude or be in competition with the other (see Heil 2013: 23-30). Davidson himself writes (ME: 215), "Mental events are mental only as described", reflecting his view of events as unstructured particulars rather than property-instances, so that characterization of events as mental is a matter of predication rather than property-attribution. So while the

causal-nomological descriptions, on Davidson's account, reflect real causal-nomological events, mental descriptions are no more than descriptions, hence the latter need not mirror or otherwise answer to the structure of the former.

Nevertheless, Child and Thornton do draw attention to an important point. Our two causal explanations of John's going to the fridge, one D-N, the other in terms of a belief/desire pair, make use, according to anomalous monism, of two descriptions of what is in fact the same extensional causal relation, yet since there is no isomorphism the two explanations, as explanations, share no features in common. The theory asks us to accept as a brute fact that on every occasion on which someone either acts rationally, or forms a rational perceptually-based belief, a causal-nomological chain of events occurs which also instantiates a pattern that is recognizable by us as conforming to the constitutive ideal of rationality. Even if there is no matching of causal and normative roles, an unexplained harmony of a sort remains. The mystery can be at least partially resolved, I argue in later chapters, with the aid of constitution theory.

However, to return to the isomorphism argument, consideration of the reason why (if I am right) mental patterns need not mirror causal-nomological patterns leads to conclusions that might render the whole concept of a mental event, as it is depicted in many versions of non-reductive physicalism (although not, perhaps, in Davidson's own), problematic. To illustrate this, I want to develop the concept of a *biofunctional event*. On the Macdonalds' account, biofunctional properties are "plausibly viewed as co-instanced with physico-chemical properties" (1995c: 104), analogously to the co-instancing of mental and

physical properties, in what we can, exploiting the analogy, call a biofunctional event, which is also a physical event. The Child/Thornton claim is that token identity implies an isomorphism of intentional with physical processes, in a way that is incompatible with psychophysical anomalism or the irreducibility of intentional patterns. Biofunctional patterns are also “generally acknowledged *not* to be reducible to their physico-chemical bases”, in that, for instance, they make use of normative notions (1995c: 104). So they too should be susceptible to the objection that, analogously with Child’s picture of reasoning, each step in the playing out of a functional process, proceeding, *ex hypothesi*, in accordance with functional patterns, is isomorphic with a causal-nomological process. To put it another way, the claim would be that token identity on the basis of co-instanting of biofunctional and physical properties must be rejected because it would require that each co-instanting be linked with the previous and the next in accordance with *both* causal-nomological *and* with biofunctional patterns, and that this possibility is ruled out by the irreducibility of biofunctional properties and their anomalism from the point of view of causal nomologicality.

Here an obvious disanalogy between the biofunctional and intentional cases appears. We can form a concept of a mental event, as it were, from the inside. We have a perfectly intuitive idea of what we mean by a mental event both from our own experience and our interpretations of others when we and they make inferences, form intentions, and so on. We have a point of view on the world; indeed McDowell, Hornsby, Lynne Rudder Baker (see chapter 3), and many others would agree that without this we would have neither the idea of intentional patterns or of mental events, nor, for that matter, of any other kinds of patterns or events. We thus have a way of individuating mental events, and a reason

for thinking of them as particulars, independently of any theory that might *postulate* their existence in order to account for, say, mental causation. But nothing corresponds to this in the case of biofunctional patterns, as we observe them in nature. We can say (probably greatly oversimplifying) that the colour change of some component of a *Gilia* petal from red to white is both a physical and a functional event, a co-instanting of functional and causal-nomological properties. But, in contrast to the intentional case, the subjects of biofunctional events generally do not *experience* themselves ‘from the inside’ as subjects of those events. An exception might be, perhaps, when *we* consciously experience ourselves as, say, breathing, but the vast majority of the world’s biofunctional events either occur in life-forms that lack a point of view or else are events in our own bodies, like the filtering activity of the kidneys, which we do not directly experience. In those cases we have no way of characterizing a biofunctional event except as a physical event that we recognize is also part of a biofunctional process. In other words, what we *see* are biofunctional patterns, and we label a biofunctional event as such only when we already see the pattern. Looking at a beating heart, for example, we recognize a physical event, say an individual contraction, as a biofunctional event only against a background of knowing what the heart’s function is. The biofunctional property that is co-instanted or in some other way correlated with a physical property comes into view only when the system is viewed as a whole, so that the patterns become apparent, and the biofunctional event is describable as such *only as* part of the pattern. Failing recognition of the pattern, there are just causal-nomological events. The distinction between the two kinds of event or process is thus an intensional distinction between descriptions or explanations.

An advantage of using the biofunctional analogy is that, compared with our understanding of what the physical/mental events postulated by Davidsonian token identity theory might be like, we have a reasonable understanding of the relations that actually obtain between the physical and the biofunctional in at least some cases. Thus if we think of the causal-nomological processes at work during, for example, a series of cardiac ventricular contractions – severing of actin-myosin linkages, opening and closing of ion channels, and so on - it is evident that, while we can think of each of these causal-nomological events as correlated with a biofunctional event, the description of the biofunctional *pattern* is *not* isomorphic with that of the causal-nomological pattern. The biofunctional pattern does indeed have a causal-nomological description, but that description, or explanation, will have to refer to a much wider group of causal-nomological events than the local and immediate ones that comprise, say, Gilia's colour change. This wider group will include all the historical and environmental physical events that were and are essential to the colour change's having the biofunctional significance that it does – those upon which the biofunctional description *supervenes* (see 3:III:(ii)). The description would include, for example, an explanation of how the system comes to be “hooked up” in the way that it is. If one wished, one could say that there is an isomorphism between *these* two descriptions, since there will be a causal-nomological description corresponding to each element of the biofunctional account.

I think, however, that what is more significant is that the full causal-nomological account of the relations that we call biofunctional that is available, at least in principle, need not refer to biofunctional patterns or events at all. *We* detect what appear to us to be causal relations among the biofunctional properties - we say, for example, that random colour

changes in earlier generations of *Gilia* that happened to coincide with the change of pollinator caused increased reproductive fitness in the individuals that made the change, or that the contraction of a myocardial cell causally contributes to the delivery of oxygen to the tissues. That is, we discern the patterns of biofunctional relations, but for the causal-nomological processes of the world to unfold, it does not seem necessary that those patterns exist. Biofunctional events seem to be just causal-nomological events viewed in a particular way. A biofunctional event or property-instantiation now appears as an abstract concept, which would be nothing in the absence of the concept of a biofunctional pattern.

Suppose, then, that we can treat mental events analogously – suppose, that is, that mental events such as the onsets and offsets of beliefs and desires are, as Hornsby (ACE: 150) puts it, *aspects*, rather than ontological components, of reality. In fact, this appears to be close to what Davidson meant by his remark, quoted above, that token physical/mental events “are mental *only as described*” (ME: 215, emphasis added), and also:

...[i]n my view the mental is not an ontological but a conceptual capacity...To say of an event, for example an intentional action, that it is mental is simply to say that we can describe it in a certain vocabulary – and the mark of that vocabulary is semantic intentionality (PEA: 114).

On the evidence of this passage, my argument from the biofunctional event analogy would be misdirected if aimed at Davidson. The passage also shows that, as I argued above, Davidson’s account need not entail an isomorphism between mental and physical patterns. His suggestion

seems to be that that mental events are only mental for *us*, or creatures like us, who have the capacity to recognize and describe them as such – that is, for rational beings who are themselves possessors of conceptual capacities, who are capable of resonating, as McDowell would put it, to the rational structure of the world, and hence who can form a *concept* of a mental event.

My argument might rather be directed against the view of those, such as the Macdonalds, who think of mental events as actual instantiations of properties. As I have been arguing, accounts of that kind can be seen as conflating two notions of mental events: first, an intuitive concept derived from our subjective experience, and second, a contrived, derivative one originating in our view of mental patterns from outside – one could say from ‘sideways-on’. The resulting uncertainty about the precise identity conditions of mental events is explored by Hornsby (WPEME), who points out that even if we had access to detailed neurophysiological data accompanying a mental event such as an action we still could not specify which neurophysiological events mark the beginning or end of the action, and hence which events can be said to be identical with it (*ibid*: 68-70).

(iv) Rational Patterns and Minimal Physicalism.

In ME (213–214), Davidson identifies four possible positions one could occupy on the mind-body relation according to one’s view of, firstly, the possibility of psychophysical laws and secondly, the identity of mental and physical events: they are nomological monism, nomological dualism, anomalous dualism, and Davidson’s own position, anomalous monism. If McDowell were to be assigned to one of these positions, on his own account it would have to be anomalous dualism, since he

accepts psychological and psychophysical anomalism but rejects monism (FAM: 339–340; RAB; RGM). But he does not, of course, reject the role of the physical altogether. He can reject *Davidsonian* monism, to the extent that it presupposes the universal applicability of the PNCC, while still eschewing substance dualism.

McDowell (PMM: 281) distinguishes acceptance of the obvious fact that a properly functioning brain is necessary to mental life from what he claims does not follow – “that the proper functioning of that organ is what mental life, in itself, is”. The mind is not an organ, either identical to the brain or “immaterial”; rather, “talk of minds is talk of subjects of mental life” (*ibid*: 280-281). Minds are essentially related to their environment; they are not “in the head” (*ibid*: 276). These remarks suggest that McDowell’s view is close to what Child (1994) calls the “most relaxed”, or minimally physicalist, possible view of psychophysical correlations, according to which “we should not, for example, expect the physical story to be isomorphic with the corresponding mental story, with a one-one correlation... we should not even expect to find identities between token mental and physical events or states” (*ibid*: 113). A similar view is expressed by Hornsby (ACE), with her argument that actions are inaccessible from the impersonal standpoint. Of the impersonal, or, in McDowellian terms, the realm of law point of view, Hornsby writes (*ibid*: 150), “...if it is a claim to survey the whole of space and time and deal with every portion of what it surveys, then nothing stands in its way”; but actions (and, I think we could add, intentionality in general) are, for her, an aspect of reality, not a portion of it (by ‘reality’ here she evidently means what is co-extensive with the ‘whole of space and time,’ while an ‘aspect’ of reality would seem to be how reality appears from a particular perspective – here, the

perspective of a rational being). We should not assume that “to everything we speak of from the personal point of view, there attaches a piece of vocabulary apt for describing things impersonally” (*ibid*). McDowell himself criticizes “both the idea that sub-personal cognitive psychology might supersede ‘folk psychology’ and the idea that it reveals the hidden depths of something whose surface ‘folk psychology’ describes in a rough and ready way” (FAM: 339 n13). For him and Hornsby, as Jose Bermudez puts it (2005: 52), “the *explananda* of commonsense psychology do not feature in any way at all at the subpersonal level”. In PPD (1978) McDowell claims (acknowledging Davidson’s similar argument in ME) that the incommensurability between intentional and physical explanations is such that “there is no threat to the completeness of physics ...if...semantic laws and their special conceptual content cannot be physically mirrored” (*ibid*: 150-151). He allows that this is compatible with “the events that comprise linguistic behaviour hav[ing] physical descriptions...under [which] they are explicable as instances of the way the world works” (*ibid*: 153). What he denies is that a physical characterization can be given of the “point of the grouping”, that is, of semantic content (*ibid*: 152).

A clear statement of this view that we cannot expect mental entities to be physically “mirrored” is given by Stoutland:

Intentional behavior, of course, belongs not only to the space of reasons but also to the realm of physical law. To describe action and reasons for action is also to describe (no doubt very obliquely) events, entities, and properties in the realm of physics (including neuro-physics)...The realm of physical law is not underneath the space of reasons in the sense of being more basic to our grasp of

human action. There is no identity between the events, entities or properties in the realm of physics and those in the space of reasons...token identity is either innocuous – a fancy way of saying that we *can* give a physical description of anything – or false, because we individuate events and entities so differently in physics than we do in the space of reasons (Stoutland 1998: 53).

What is being presented is a form of physicalism according to which not only are the patterns of rationality, even in principle, indiscernible from the physical standpoint (this much is shared with the Davidsonian view), but also according to which mental and physical properties are not co-instantiated or otherwise token identical. As regards propositional attitudes, the position resembles the eliminativism proposed by Churchland (1984) and others, in that beliefs, desires, etc, are envisaged as having no physical correlates. But while the eliminative materialists take this as a reason for thinking, even hoping, that ordinary talk of propositional attitudes might in time be rendered obsolete by advances in neuroscience, the thinkers I am interested in here take an opposing view. The centrality of the subject and her viewpoint on the world, where ‘world’ does not mean merely a physical environment but a socio-culturally characterized milieu, is irreducible, and the personal standpoint *consists* largely in the holding of propositional attitudes. The world of Sellars’ manifest image, or of McDowell’s second nature, is a cultural entity which has emerged from the co-evolution of persons and cultural communities over many millennia, out of which the idea of the scientific image, or the realm of laws, has grown only in the last four hundred years. From this point of view, then, propositional attitudes are real, although not in the sense of being identifiable with physical states.

This picture of the attitudes fits well with McDowell's image of the space of reasons, and second nature, as *sui generis*.

Stoutland's claim, above, that token identity is either innocuous or false is best understood in the context in which it is made, which is that of an account of action (Stoutland 1998). His arguments in that essay, generalized to intentionality as a whole, contain ideas remarkably similar to McDowell's. On Stoutland's account, actions are generally responses to something. Most reasons for action are external situations – that is, situations in the world which (to put it in the terms McDowell uses in MW) exert a rational constraint on us. This is in contrast to the Davidsonian view of reasons as internal states – belief/desire pairs – that are causally produced by the world. So for Stoutland, as for McDowell, reasons – contentful items – are in the world, and when we apprehend them they become part of our content. There is, to paraphrase McDowell (MW: 27), no ontological gap between the reason in the world and our reason. Reasons for action need not be determinate, causally interacting inner states. They are not, typically, propositional attitudes like belief-desire pairs. Roadside stop signs, to use Stoutland's example, have content when we see them *as* stop signs (*ibid*: 46), and this is why we can respond *directly* to them, as reasons. Stoutland also adopts the Sellarsian image of the space of reasons; for him, agent, action, world and reason are all normative, space of reasons concepts, the first three linked by the notion of reason (*ibid*). The ability to respond to reasons is culturally acquired, much as on McDowell's account we acquire second nature. Like McDowell, Stoutland rejects the picture of our relation to the world on which the received view (the representational version of the causal view that he rejects, while still maintaining that reasons are causes) depends. This is the picture, discussed above, in which there is a

sharp distinction between a contentless, non-propositional, causal world and our minds, whose relation to that world is only through inner mental representations. Its claim, rejected by Stoutland, is that “(e)ternal situations can serve as reasons only in virtue of [propositional] attitudes which represent them” *ibid*: 64), and that they can serve as real reasons for actions only if they cause the actions.

In Stoutland’s contrasting picture, obviously closely similar to McDowell’s, the space of reasons extends outwards to *include* the world, which, when one thinks truly, is thus part of content directly, and not as an internal representation. Actions are thus rational, judgment-involving direct responses to reality. Things like roadside stop signs are not merely contentless physical objects; they form part of the space of reasons, and “carry purposes and beliefs independent of the current attitudes of individual agents” (Stoutland, *ibid*: 64–65). However, Stoutland, (*ibid*: 61) does acknowledge that, for example, in cases where someone acts on a false belief, such that there is no external situation to serve as a reason, *their believing* is their reason for acting. He does not deny, then, that propositional attitudes *can* be reasons. What he does deny is the inference from the fact that they can be reasons in circumstances like false belief to the claim that they are also always reasons in cases of veridical belief. Again, the parallels between this account of reasons for action and McDowell’s disjunctive theory of perception are clear.

McDowell’s view, shared by Stoutland, Hornsby, and others, that the propositional attitudes, and commonsense psychology as a whole, is visible only from within the space of reasons, for those who have acquired second nature, seems to be supported by what the biofunctional analogy tells us. On this view, as Stoutland *et al* claim, it would be a mistake to try to find *physical* events or states - that is, items that are

identifiable independently of the space of reasons standpoint - that are identical to or otherwise individually correspond to the attitudes. In addition, we have seen that Stoutland (1998) and McDowell are content externalists. McDowell expresses the idea variously as the unboundedness of the conceptual, our openness to the world, and the claims that experience is already in conceptual form, that we are under rational constraint by the world, that there is no ontological gap between thought and the world, that the mind is not in the head, and that the world is composed of true thinkables. On this view the patterns and relations of rationality and normativity are not confined to our minds but fill the whole of our reality – the world of thinkables. The *attitudes*, of course, *are* confined to our minds, being features of acts of thinking rather than of thought, or content. It is content – that which appears as the contents of the propositional attitudes – which is shared by both mind and world. We thus have a convergence with the conclusion of the biofunctional analogy; if there is an isomorphism involving rational/normative patterns it will be with the causal-nomological structure, not of an individual or her brain, but of the much wider reality upon which her mental content may be said to supervene (chapter 3).

All the authors I have mentioned who support views of this kind are in one sense physicalists, even if it is only the sense that they are not substance dualists. None deny, as Stoutland puts it above, that “we *can* give a physical description of anything”. What is being advocated is a variety of non-reductive physicalism that denies even token identity. But what exactly does this denial amount to? McDowell, Stoutland, Hornsby, and Child would agree that, say, John’s deciding to go to the fridge for a beer is associated with his being in *some* physical state, and probably even that there may be a reliably repeatable correlation. So what are they

denying – what does the difference between their position and token identity amount to?

Part of the answer must lie in the externality of content. If content is not confined within our heads but pervades our whole world, then obviously it will not be possible to identify the token occurrence of a propositional attitude, with its content, with the token occurrence of a physical state of the subject or her brain. Recall McDowell's object-dependent view of thought (chapter 1), according to which singular (rather than descriptive) thoughts and their contents are *individuated* not merely by the contents of the subject's mind but also by the world, consisting of true thinkables, beyond her mind. Suppose Jane mistakes a dark object under a tree for her neighbour's cat, and so takes herself to entertain the thought "that cat belongs to my neighbour". On this view, she is mistaken about the contents of her mind, since in the absence of the appropriate object, the thought is not available to be had (chapter 1; see McDowell, STEIS). Jane is thus not the ultimate authority on the contents of her thoughts. If that is right, the possibility that that the thought is identical with a physical state of Jane's brain is ruled out.

Perhaps we could make a distinction between acts of thinking – the acquiring or holding of propositional attitudes – and thought content. Then we might claim that the first, but not the second, are identical to physical states. To put this in terms of the Macdonalds' interpretation, the idea would be that, if Jane decides to pick up and return the cat, the property of being Jane's *deciding* is co-instanced with a causal-nomological property, but the content of the decision is not, since the latter depends on the identity of the object (the argument is unaffected if the relation is taken to be something other than co-instancing, for example constitution). One might, however, question whether it makes

sense to separate the deciding from its content in this way – how do we characterize a deciding in abstraction from the content of the decision? This is McDowell's view; he argues, against Putnam, that our thoughts and imagery are not characterizable apart from their content (PMM: 286-287), but are essentially direct representings of the world (see also Burge 1993: 108-110). Once we agree that a deciding is associated with an event or state with some causal-nomological description, in fact, it seems reasonable to assume that the causal-nomological event that accompanies a deciding to, say, get a beer from the fridge is distinct from that accompanying a deciding to get a coffee from the machine, and that these two events are probably more alike, in some sense, than either is to a deciding to attend a philosophy seminar. Thus it seems inescapable that what appear to us as differences of content are reflected in differences in physical characterization. So, again on the Macdonalds' account, it looks as if in the act of deciding there is a co-instancing, with a physical property, of the property of being a deciding with a particular content. But in the case of Jane, this content is either that of intending to pick up the cat or it is not, and which of these it is depends on the situation in the world beyond Jane's mind. So the nature of the content is not fully individuated by the physical realization or constitution of the mental act or event. If it is correct, then, that we cannot characterize the mental act independently of its content, we have no choice but to conclude that the relation between the mental act and the simultaneously occurring physical event is not one of identity. So a further difference between McDowell *et al* and token identity theorists is that the former reject token identity.

But if the above argument is on the right lines, this loose construal of physicalism, which denies even token identity, must, nevertheless,

concede that the causal-nomological events that occur when Jane decides, on an occasion, to pick up a cat must differ from those that occur when, on an occasion, she decides to stroke a dog. So on the two occasions we have two distinct physical patterns and two distinct rational patterns. According to McDowell's kind of content externalism, the content that accompanies the instantiations of the rational patterns is partly externally determined, and according to the Macdonalds' account the relations that constitute the rational patterns are not reducible to – not characterizable in terms of – the physical patterns.

It seems clear that content is a feature of the rational patterns. Again, as such it is only discernible by beings who are capable of recognizing it, that is, beings who themselves are bearers of content - in McDowell's terms, occupants of the space of reasons. The Macdonalds' way of ensuring that content has causal efficacy is to claim that each instancing of a particular, special kind of physical event is *also* an instancing of a mental event, in that the instancing shares physical and mental properties. Then the occurrence of a chain of such physical events, each following the other in accordance with physical laws, *is also* the occurrence of a chain of mental events, which together form a pattern of a quite different kind – one that we, somehow, recognize as constituting content. I have been suggesting, however, that the notion of an individual mental event, considered “from the outside”, in isolation from the rational pattern of such events of which it is a member, may be rather unhelpful – its identity conditions are entirely derivative from, first, recognition of the rational pattern, and second, the assumption that physical and mental properties are co-instanced in individual events.

On the looser construal, the claim need only be that when the appropriate causal-nomological processes occur, a different pattern of relations

somehow becomes discernible, to beings like ourselves, as constituting content. A picture something like this, I suggest, should capture the views of McDowell, Stoutland, Hornsby, and others, including, with important modifications, Lynne Rudder Baker, whose work will be examined in the next chapter. The most pressing difficulty with this minimal physicalist view is that it raises again the specter of epiphenomenalism; how could content, as part of a pattern which is discernible only against the background of causal-nomological processes that exemplify a different pattern, be claimed to have any causal efficacy in its own right? A reductive physicalist like Kim might criticize, say, Stoutland's account, above, by claiming that the idea that a stop sign has a rational significance is merely conceptual, reflecting our interests, and that the real cause and explanation are located at the physical level. Again, what is needed is a discussion of the ontological status of these higher-level causal relations, and it is here that Baker's constitutional view might be expected to help.

(v) *Conclusion.*

We now have further clarification of what a fully worked out McDowellian view of the relation of the mental to the physical might look like, and an idea of how it differs from the Davidsonian account. In one sense we can call McDowell a physicalist, since he allows that physical structures and processes necessarily accompany exercises of intentionality. If so, however, he is a highly idiosyncratic, minimalist, and radically non-reductive kind of physicalist, since for him, firstly, *content* is not physically characterizable, and secondly the physical is *within* the space of reasons. As he puts it in Sellarsian terms, the "real order" is "embrace(d)...within the conceptual order" (WL: 63). Content, for possessors of second nature, forms the fabric of the world, but when

we adopt the narrow standpoint of the realm of law, as we do in both cognitive psychology and neuroscience, content, an essentially space of reasons concept, does not come into view at all. In another sense, then, McDowell is not a physicalist in any standard way.

Nevertheless, given that McDowell does acknowledge an intimate and necessary connection between our participation in the *sui generis* space of reasons, or our possession of second nature, and the physical, or first nature, we are justified in asking what the nature of this connection is. In II:(v) above I mentioned that if, as I claim, we must acknowledge that the truth or falsity of causal claims is objective in the recognition-independent sense, then we must accept that when someone performs an action like raising a glass, two distinct kinds of causality – intentional and nomological – seem to be simultaneously in play, and it is hard to see how this could be so. If, as McDowell acknowledges, the intentional is anomalous, or *sui generis*, then clearly a scientific account, no matter how complete, can never explain rationality, normativity and the other hallmarks of intentionality. In that case, science is permanently debarred from *ever* giving the answer to the question ‘if the relation between the physical and the intentional is not one of even token identity, then what is it?’ One way of putting it is that while science might conceivably provide a full account of how our intentionality is possible, or is physically *enabled*, *ex hypothesi* it can have nothing to say about what our intentionality *is*. McDowell is quite prepared to accept this and leave the matter there, but just because he insists that the scientific account of nature is not a full account, this does not *ipso facto* exonerate him from showing how the account it *does* undoubtedly provide can sit comfortably within his expanded naturalism. And, as I have argued in this chapter, where it appears to sit most *uncomfortably* is in the area of

causality, in the tension between the naïve space of reasons view of causality and a rival one which, although it is backed by science, arises not from science alone but, I have argued, from the view of the physical world from within the space of reasons itself. In the next chapter I begin examining the role a constitution account may have in resolving this difficulty.

CHAPTER 3: BAKER'S CONSTITUTION ACCOUNT.

I. Introduction

In the previous chapter I made some suggestions as to the shape of an account of mind-body relations that would underpin a McDowellian account of causality within the space of reasons – one that allowed intentional items to be causes in their own right, in his words (RAB: 69). The picture that emerged was one of what Child had called minimal physicalism – a rejection of token identity and of Davidson's imposition of the PNCC -, but also one in which world, as the totality of facts, and mind form a seamless conceptual space in which content is shared. Causality in this picture would be constrained by the rational and normative patterns characteristic of the space of reasons, which are discernible only to occupants of that space. However, as we saw (2:II:(i); III:(v)), we face the problem that this account seems to conflict with our intuitions about ordinary nomological causation – nomological in a sense that belongs within the space of reasons, and hence cannot be dismissed (as both Hornsby and McDowell are inclined to do) as irrelevant to that space. In this chapter I begin the task of developing an account that aims at resolving this difficulty by expanding the constitution account expounded by Lynne Rudder Baker. I believe this account can provide the basis for an acceptable conceptual link between the realm of law and the space of reasons – the link that McDowell denies is needed.

In five books (1987, 1995, 2000, 2007, 2013) and many articles and other contributions, Baker has developed a mature non-reductionist metaphysics over the last two decades or so. I will mainly be concerned

with Baker's account of the *constitution* of entities at higher ontological levels by those at lower levels, which she develops in *Persons and Bodies* (PB; 2000) in relation to persons and expands in *The Metaphysics of Everyday Life* (MEL; 2007) to a general account applicable to all macroscopic objects, properties, and property-instances, including mental properties and property-instances. Her most recent monograph, *Naturalism and the First-Person Perspective* (NFPP: 2013) concentrates on the irreducibility of the first-person perspective from our ontology but also develops earlier themes, such as higher-level causation.

In her earlier books, *Saving Belief* (SB; 1987) and *Explaining Attitudes* (EA; 1995), as well as in essays such as 'Are Beliefs Brain States?' (ABBS, 2001) and 'Against Reductive Physicalism' (ARP, 2006), Baker sets out her case against what she calls the "Standard View" on intentional content – the view that propositional attitudes are physically realized in the brain. She advocates instead a position she calls "practical realism", which, when applied to beliefs, (taking beliefs as representative of propositional attitudes generally) claims that they are real (in the sense that persons really believe, or have the property of believing) and causally explain behaviour, but according to which "there is no metaphysical requirement that attitudes be constituted by particular brain states" (EA: 6–7). The views Baker rejects include not only type identity reductive physicalism and the token identity forms of non-reductive physicalism, but also eliminativism – the last because the eliminativists assume that the entities, beliefs, whose existence they deny would, if they *did* exist, have to be physical entities. But she also denies that propositional attitudes are "in any sense immaterial" (*ibid*). Her position here seems related to the kind of minimally physicalist view I discussed at the conclusion of the last chapter – a view that, while rejecting

immaterial entities, nevertheless denies that there is a relation of identity between beliefs and type or token physical or neural states. For Baker, for a person to entertain a belief, in the sense of, say, coming to believe that p , there must occur certain activity in that person's brain, but *beliefs* are "not spatiotemporal entities or internal states at all...the term 'belief' is just a nominalization of 'believes that'" (EA: 21). Having a belief is an aspect or property of a *person* rather than of, say, a brain. The basis on which a belief is attributed to a person, furthermore, is solely in terms of its possible behavioural or intentional consequences, quite independently of the state of her brain; "S believes that p if and only if there are certain counterfactuals true of S, where the content of the counterfactuals may be intentionally characterized" (*ibid*). The latter clause ensures that this is not an attempt to give a non-intentional, for example purely functional, account of belief. Nevertheless, it is worth noting that Baker avoids any deeper discussion of, for example, what grounds the counterfactuals, or makes it the case that they apply. We will see below (section IV) that she adopts a similar attitude to causality.

In PB (2000) and MEL (2007) Baker develops her general ontology and her account of persons, the two being closely connected *via* the concept of constitution, which I discuss in detail in the following sections. While Baker uses the idea of constitution at least partly in order to develop her account of persons, she emphasizes that the material constitution relation is quite general, applying to all objects other than those at the most basic level, whatever that may turn out to mean. Her overarching theme is that the world is ontologically rich and that macroscopic objects, including those whose existence depends on the existence of intentional beings like ourselves, are irreducibly real. Ordinary objects, including non-human animals and persons, together with their properties, are part of the

ontological structure of the world. This is made possible by the relations of material constitution, by means of which one thing, in favourable circumstances, comes to constitute another thing at a higher ontological level, with new, irreducible causal powers, and property-constitution, whereby the *properties* of an entity, say an object, come in favourable circumstances to constitute higher properties, again with new causal powers. Many constituted objects are individuated by their *relational* properties (see **II** below), and many constituted *property-instances* are not only relational but are such that their possession by an entity presupposes the existence of creatures such as ourselves. That these properties and their instances are real, and are at least as fundamental to the structure of our ontology as are non-relational properties, is a central claim of the constitution account. As I will emphasize below, the constitution relation is not a supervenience relation. The new properties acquired by constituted entities do not supervene on properties of their constituting entities alone. However, as we shall see, supervenience has a role within the constitution account.

Among the irreducibly real constituted objects that the world contains are persons, who are constituted by, but not identical to, human organisms, and who are distinguished by their adopt what Baker calls a strong first-person perspective – to refer reflexly, and, provided one is competent, infallibly to oneself as the subject occupying the perspective in question. The ability to adopt this perspective, Baker argues (PB:163), is the essential and defining characteristic of persons, in that a person could not permanently lose the property of being able to adopt the first person perspective without ceasing to be a person. She writes (MSVK:166); “A first person perspective is the ability to conceive of oneself, from ‘within’ so to speak, without any name, description or

demonstrative”. For my purposes, the most important part of Baker’s account of persons will be her claim that, in common with other constituted objects (see below), “a person has causal powers that a body would not have if it did not constitute a person” (PB: 109).

In section **II** of this chapter I describe Baker’s constitution account, beginning, as she does, with material constitution. Then in sections **III** and **IV** I critically analyze her use of the account, and particularly of the notion of property constitution, to underpin a theory of mental and other higher-level causation. It will turn out that Baker’s account of causation suffers from much the same difficulties I have identified with McDowell’s; however, her own constitution account supplies a way of resolving the difficulties.

II. The Constitution Account.

(i) Material and Property Constitution.

Constitution, on Baker’s account, is a relation of *unity without identity*, a category that lies between identity and separate existence without being either. I discuss Baker’s defence of this notion in *(iv)* below. To look first at material constitution, when one object constitutes another, as for example a when piece of marble (called ‘Piece’; PB: 29ff) constitutes Michelangelo’s *David*, there are not two objects but one, a unified “statue-constituted-by-a-piece-of-marble” (MEL: 166). There is not separate existence; “[a]s long as x constitutes y , x has no independent existence” (PB: 46), but neither are Piece and *David* identical. Certainly, they share (albeit, crucially, in different ways – see *(iii)* below) many properties –being worth n million euros, serving as an inspiration to

would-be sculptors, and having the same height, weight, and spatial location. But Piece, unlike *David*, could exist in a world without art, indeed without intentionality at all. While constitution is a contingent relation, the identity relation, Baker insists, is such that if $x = y$, then *necessarily* $x = y$ (PB: 31). Therefore identical things cannot differ even in how they *might* be – they cannot differ in their modal properties, such as that of possibly existing in an artless world. Therefore Piece and *David* cannot be identical. Nor is *David* identical to Piece plus some other thing (*ibid*) - Piece is not a *part* of *David*. Rather, Piece *constitutes* *David*. The constitution relation enables us to say, in fact, that Piece *is* a statue – *David* – in a distinct sense of ‘is’ that is neither the ‘is’ of identity (‘Piece is a piece of marble’) nor of predication (‘Piece is in Florence’) but of constitution (PB: 54). For Baker, this use of ‘is’ does more than assign a predicate. To say that Piece *is* a statue in this constitutional sense is, as we shall see below, to say not only that Piece and *David* share a spatiotemporal location but that, in the particular way I describe in the next subsection, they share all their properties.

Integral to the notion of constitution is the idea that when certain things with certain properties are in certain *favourable circumstances*, new things, with new properties, come into existence (PB: 32). This is how, ultimately, everything that exists (except whatever entities might exist at a hypothetical base level) is constituted. Favourable circumstances, in Baker’s technical sense, consist of extrinsic or relational properties that must be instantiated if the constituting object if it is to constitute the higher object in question. Thus Piece constitutes *David* due to its instantiating properties such as having been deliberately created in a world in which art is appreciated. Baker (EA: 63-65) defines a *relational* property thus,

R is a relational property if and only if: x 's having R entails that there is some y distinct from x .

Being a planet is a relational property, while being a lump of rock, even one molecularly identical to a planet, need not be, since the lump, but not the planet, could be the only thing in existence. For Baker, who, like McDowell, is a content externalist, this constraint applies also to propositional attitudes such as having a belief that water is wet, since if subject S was the sole existent, there being no water, she could not have that belief.

An essential feature of Baker's account is that many properties – both properties as types instantiated by objects and property-instances such as events – are not only relational but also *intention-dependent* (ID). These will feature prominently in the discussion to follow. ID *properties* are either propositional attitude properties – believing, etc – or properties whose instances presuppose that there are entities that are bearers of propositional attitudes (Baker ARP: 6), while ID objects are either such entities or objects whose existence presupposes their existence. If something microstructurally identical to a car spontaneously assembled in outer space it would not be a car, because it is of the essence of a car that it has been designed by intentional agents for transportation (MEL: 12). Instances of ID and other relational properties are bearers of new and irreducible causal powers.

Favourable circumstance, Baker concedes, are too varied for there to be a general account of them (MEL: 160), but she suggests that in a particular case, for example a piece of cloth's constituting a national

flag, the favourable circumstances may be specified by a list of open sentences true of something when it constitutes a flag: ‘*x* is in a context in which there are conventions of national symbols’, ‘*x* is flat, rectangular, and nonrigid’, ‘*x* is deliberately constructed of coloured cloth with a specific pattern’, and so on. An *x* that satisfies all these constitutes a national flag.

A continuous series of constitution relations leads all the way up from the most fundamental level to complex ID objects and to persons. Baker adds some clarificatory points: (i) the constitution relation is asymmetric in that Piece constitutes *David* but not *vice versa*. This is what gives rise to the hierarchy of ontological levels mentioned above. Piece’s identity is *subsumed* in that of *David* (PB: 33) for as long as it constitutes it, giving *David* the greater ontological significance, or “pre-eminence” (MEL: 36). (ii) The relata of constitution are individual things – objects in the case of material constitution and property-instances in the case of property constitution (see below). This brings out the contrast between constitution and (mere) composition. While *David* is composed of marble, it is constituted, not by marble as ‘stuff’ but by Piece – an individual piece of marble. The relation is close to identity, and an identity theorist would not say that *David* was identical to ‘marble’, since *David* is a particular and ‘marble’ is a universal. (iii) Very many, but by no means all, constituted objects are ID ones. Genes are constituted by DNA molecules, and a certain combination of chemicals constitutes an organism (MEL: 36), but a gene or an organism is not an ID object.

According to Baker, each concrete individual is of exactly one *primary kind* – the kind that it could not cease to be of while continuing to exist

as that thing. It is in virtue of their primary kind that things are constitutionally related (*ibid*). Primary kinds are the determinants of the persistence conditions of constituted objects. *David's* primary kind is *statue*, and being a statue is its primary kind *property*; *David* could not cease to be a statue without going out of existence, while *Piece* could (in the artless world, for example). *Piece's* primary kind property is that of being a piece of marble. Although it *is* a statue, this is not its primary kind. Even though being a statue is a primary kind property that *Piece* has, it is not *Piece's* primary kind property. Rather, *Piece* has the property of being a statue only *derivatively* (see below).

The following is a slightly modified version of the thesis, derived from a combination of Baker's formal schema for material constitution in PB: 43 and the updated version in OMTU (2002) and MEL:161:

Let *being an F* be *x's* primary kind property, and let *being a G* be *y's* primary kind property, where these are distinct properties. Let *F** be the property of *having the property of being an F as one's primary kind property* and let *G** be the property of *having the property of being a G as one's primary kind property*. Then:

- (C) *x* constitutes *y* at *t* = (df). There are distinct primary kind properties *F* and *G* and *G*-favourable circumstances such that:
- (a) *F*x* and *G*y*; and
 - (b) *x* and *y* are spatially coincident at *t*, and $\forall z(z \text{ is spatially coincident with } x \text{ at } t \text{ and } G^*z \rightarrow z=y)$, and
 - (c) *x* is in *G*-favourable circumstances at *t*; and

- (d) It is necessary that: $\forall z[(F^*zt \ \& \ z \text{ is in G-favourable circumstances at } t) \rightarrow \exists u(G^*ut \ \& \ u \text{ is spatially coincident with } z \text{ at } t)]$; and
- (e) It is possible that: $\exists t\{(x \text{ exists at } t \ \& \ \sim \exists w[G^*wt \ \& \ w \text{ is spatially coincident with } x \text{ at } t])\}$; and
- (f) If x is of one basic kind of stuff, then y is of the same basic kind of stuff.

In the Piece and *David* example, primary kind property F is that of being a piece of marble and primary kind property G is that of being a statue, and Piece's G-favourable circumstances are as described above. Clauses (d) and (e) then read:

- (d) It is necessary that: if anything that has being a piece of marble as its primary kind property is presented as a three dimensional figure in an artworld, given a title, and put on display at t , then there is something that has being a statue as its primary kind property that is spatially coincident with the piece of marble at t .
- (e) It is possible that: a piece of marble exists at t and that no spatially coincident thing that has being a statue as its primary kind property exists at t (PB: 44).

Definition (C) is concerned with constitution as it applies to *objects*, that is, material constitution, a relation that defines the conditions under which one object can come to constitute another, with new properties. In more recent work (MEL: chapter 5; Baker ARP, NRM) Baker introduces a similar account of *property-instances*, the property-constitution (PC) view, to explicate and defend her views on higher-level causation and

nonreductive materialism. (PC) defines the conditions under which instantiation of one property (not necessarily a primary kind property), in a single entity, can come to constitute the instantiation of another property in that entity. (PC) can be thought of as an extension of (C), in that while the (PC) formulation applies to the constitution of objects (we can say an object's instancing the property of being a piece of marble, for instance, constitutes in favourable circumstances an instancing of the property of being a statue), it can also be made to apply to other kinds of particulars, such as states and events. As we shall see in the next section, Baker advocates a form of nonreductive materialism (NRM), based on the (PC) view, which she claims avoids Kim's criticisms of standard versions. Baker's schema for property constitution is (NRM:122):

- (PC) x 's having F at t constitutes x 's having G at $t = \text{df}$
- (a) x has F at t and x has G at t ; &
 - (b) x is in G -favourable circumstances at t ; &
 - (c) It is necessary that: $\forall z [(z \text{ has } F \text{ at } t \ \& \ z \text{ is in } G\text{-favourable circumstances at } t) \rightarrow z \text{ has } G \text{ at } t]$; &
 - (d) It is possible that: x has F at t & x lacks G at t .

As already alluded to, constituted objects have some of their properties *derivatively* and others *nonderivatively* (PB: 46-58; MEL: 37-39). A preliminary statement of the idea of having a property derivatively is: " x has H at t derivatively if and only if x 's having H at t depends wholly on x 's being constitutionally related to something that has H at t independently of its being constitutionally related to x " (PB: 47). Baker adapts the concept of having properties derivatively from Chisholm's notion of 'borrowing' properties (MEL: 37). She does not subject the idea to further defence or analysis, seemingly taking it as simple and

basic. As Baker's use of 'constitutionally *related*' in her definition suggests, on her account possession of properties derivatively, in contrast to the constitution relation itself, is symmetrical; "if x constitutes y , then *both* x and y have some of their properties derivatively" (*ibid*).

It will be relevant in later discussion (see chapter 5:II:(*ii*)) that some categories of property cannot be had derivatively, on Baker's account (PB: 48 – 49). These excluded properties are (i) alethic properties; those expressed by 'essentially', 'primary kind' or a modal expression such as 'possibly', (ii) properties expressed by variants of 'constitutes', 'is identical to' or 'exists', (iii) properties F such that necessarily x has F at t only if x exists at some time other than t , such as the property of *having been* F at t , and (iv) hybrid properties, property conjunctions that either entail or are entailed by two or more primary kind properties, such as being a cloth flag.

(*ii*) *Ontology and Ontological Levels.*

Baker's formulation of the constitution account is largely motivated by her ontological concerns. The account both presupposes and supports an ontology that insists upon the irreducible reality of everyday, macroscopic objects and properties and their instances – the commonsense conception. Baker claims that reality contains a hierarchy of ontological levels, both of objects and properties, linked by the relation of constitution. As opponents of this view, she identifies influential accounts (e.g., Merricks (2001); Sider (2002)) which claim that there is only one ontological level, and that only the fundamental entities of physics really exist. For philosophers who hold this kind of

view, a predicate such as ‘...is a mountain’ can only be a predicate, not an attribution of a property. It is convenient and probably unavoidable for us, living as we do in the macroscopic world, to speak of such objects as if they were real, but, on this view, metaphysical analysis shows that ‘mountain’ and ‘carburetor’ are merely concepts that reflect our interests.

Baker’s ontological claims are part of her practical realism, according to which “metaphysics should not swing free of the rest of human enquiry ..[it].. should be responsive to reflection on successful cognitive practices, scientific and nonscientific” (MEL: 15). “It is not that science tells us *what* exists; science tells us what *else* exists” (*ibid*: 18). For Baker, the indisputable existence of ID phenomena is a reason for rejecting the usefulness of the distinction between mind-dependence and mind-independence, and with it the idea that only what is mind-independent (or non-ID) in the sense of being part of “an in-itself reality independent of our minds and even of our existence” (*ibid*: 18, quoting from Sosa 1993) is to be admitted to our ontology. Ontology, therefore, “need not be wholly independent of our language, our activities, our conventions and practices” (*ibid*: 20). McDowell, as we have seen, takes the rather stronger view that it would be incoherent to suggest that ontology *could* be wholly independent of these; for him it is a condition of the possibility of our being bearers of content that our perceptually-based content should have objective purport – that the world that we perceive should exist. In any case, to repeat, a key element of Baker’s thought is that ID objects, properties, events, and so on with their essential natures, are as fully a part of the world’s ontology as anything else, including fundamental particles. Baker’s reasons for insisting on their reality are mostly self-avowedly practical and pragmatic, rather

than being based upon *a priori* analysis. It is the fact that ID phenomena figure indispensably in our everyday lives and discourse that drives her ontological commitment to them.

Probably the most significant and powerful ontological claim in Baker's account is her assertion that "The identity of the constituting thing is submerged in the identity of what it constitutes. As long as x constitutes y , y encompasses or subsumes x " (PB: 33), so that " x has no independent existence" (*ibid*: 46). The 'is' of constitution is, on this understanding, like the constitution relation itself, asymmetric. One might say that while to say that *David* 'is' *Piece* is simply to apply a predicate, the 'is' in '*Piece* is *David*' is the 'is' of constitution. The upshot is that constituted entities enjoy a higher, more real, or more significant ontological status than their constituters.

Baker's notion of constitution, then, presupposes that reality contains multiple hierarchical ontological levels or layers, and the notions of higher- and lower-level, as applied in particular to properties and property-instances, will feature heavily in the discussion to follow. The expository use of *descriptive*, or conceptual, levels, from fundamental physics up to, say, economics, is of course commonplace (see Kim LWMC for a discussion of 'The Layered World'). Baker, of course, views them as forming an *ontological* hierarchy (MEL: 112). The reason Kim and others hold the opposing view, according to Baker (MEL: 234-237), is that their conception of levels is mereological; "The objects at level $L+1$ are mereological sums of the objects at level L ...cells are sums or fusions of molecules, which are sums of atoms, and so on" (*ibid*: 234-235). Since sums, or aggregates, and their parts are on the same ontological level, on this view there are no different ontological levels of

reality, but only of description. This contrasts sharply with the constitution view, according to which sums or aggregates, under favourable circumstances, can come to constitute things of fundamentally different, ontologically higher, primary kinds. Baker defines higher-level primary kind properties thus (*ibid*: 236):

G is a higher-level primary kind property than F if: there are some x, y, t such that:

- (i) x's primary-kind property is F and y's primary kind property is G, and
- (ii) x constitutes y at t.

Thus the lower and higher levels are differentiated in terms of constitution, suggesting that it is the constitution relation that gives rise to or underpins the levels, rather than *vice versa*. That is, it is not that there is a pre-existing, ordered hierarchy of ontological levels waiting, as it were in a platonic realm, to be occupied by constituted entities. Also, different ontological hierarchies may not map on to each other. As Baker says, there is no answer to whether robots are on a higher ontological level than sea slugs (*ibid*). On the other hand, the idea of a level is not merely a metaphor to capture what a group of entities constituted in similar ways have in common, or how they relate to their constituters. Constituted entities (persons, say) interact, causally for example, with other similarly constituted entities, and it is in this 'horizontal' interaction that the idea of a level comes into its own.

As I said, Baker's own justification for the claims that reality is ontologically rich, diverse, and hierarchically structured is mainly practical or pragmatic, appealing to our need to make sense of the

everyday world. There are, however, other considerations that may be invoked in her support. According to the radical naturalists Ladyman and Ross (2010, see 4:I), the fundamental particles that are taken to be the sole existents by Baker's austere opponents do not exist at all. On the other hand, Ladyman and Ross advocate "rainforest realism" (see also Ross 2000) – a rich ontological pluralism of "real patterns"; "for a pattern to be real...it must be such that a community of inquirers who wished to maximize their stock of true beliefs would continue to be motivated to track the pattern" (*ibid.*: 36). While it seems unlikely that these self-avowedly scientific authors would accept the constitution account, it is interesting that Baker's ontology receives some support from this direction.

However, although in what follows I help myself to the structural framework of Baker's constitution account, with its notions of unity without identity, derivative and nonderivative properties, ontological levels and so on, I do not believe the account I will present depends crucially on acceptance of Baker's ontological claims. This is primarily because I take seriously McDowell's neo-Kantian transcendental empiricist arguments (1:III:(i)), according to which a condition of there being conceptual activity – of there being the space of reasons that we occupy – is that there be perceptual experience, and that experience is of precisely the ontologically rich world of Baker's philosophy. When we are not deceived, we are, on McDowell's account, *directly* in touch with the *facts* that make up the world. On this view, if any ontological level or category can be described as basic, it is that occupied by persons, that is, the space of reasons, which is the necessary origin of our encounters with everything else, and all other levels - the level without which we could have no concept of fundamental particles or anything else.

Science, viewed from the standpoint we must occupy, is itself encompassed by the space of reasons, together with the normative constraints that govern its practice while coming from outside its subject matter. My focus will be on constitution, and especially causality, *within* the McDowellian world – the world of the manifest image, co-extensive with the space of reasons. There is no question ordinary objects’ not existing in that sense, while, as I discuss further in chapter 4:I:(iii) & (iv), the fundamental particles that Baker’s opponents argue are the sole true existents do not figure as such in the world thus construed. The familiar ontological levels – physical, chemical, biological, and so on – are, I suggest, also real in the framework of a McDowellian ontology. In chapter 4 I argue that the entities and activities which together make up the mechanisms that are the bases of the special sciences are part of the world of the manifest image, and as such have the same ontological status as do ordinary perceivable objects. My central claim in this thesis (chapter 5:II:(iii)) depends on the idea that the *intentional* comprises a distinct level –that, indeed, the distinction between the intentional and the physical is the most important one.

Furthermore, although I take both Hornsby and McDowell to agree with what has been called Davidson’s “post-ontological philosophy of mind” (Ramberg 2000) – that the intentional is an *aspect* (Hornsby ACE: 150) rather than a component of spatiotemporal reality, I do not think this precludes it from being treated as a *level* of reality. As clause (f) of definition (C) of material constitution shows, Baker lays great stress on the claim that everything that is materially constituted is “of the same basic kind of stuff” ((i) above) – that is, everything is material, or physical (see IV:(ii) below). The account I will present makes no

particular ontological commitments, although I do claim that mental and ID properties do not *causally* interact with physical properties.

III. Baker and Higher-Level Causation.

(i) *Baker's Analysis of Kim's Causal Arguments.*

I now turn to a critical discussion of Baker's views on causality which will develop, after modification, into what I believe is a promising constitution-based account of mental and other higher-level (ID) causation. It will turn out that the account provides a picture of reality that bears striking resemblances to McDowell's.

Baker sets out a number of 'Theses About Human Persons' (PB: 105 – 110) that follow from her claims. Most of these follow fairly straightforwardly from her account, but one is, I think, particularly important:

For any objects x and y , and time t , if x is (nonderivatively) a human person at t and y constitutes x at t , then x has causal properties at t that y would not have had if y had not constituted anything at t (PB: 110).

Examples of these causal properties or powers are using a passport, enjoying a close friendship, voting in an election (*ibid*) - that is, they are ID properties. Bodies, which constitute persons in favourable circumstances, do not have these causal powers except derivatively. This is the aspect of the constitution account that Jaegwon Kim objects to; "implausibly... spatially coincident objects can, and do, have different

causal powers” (PSNE: 89n). What is usually seen as the problem of mental causation is rather, for Baker, the problem of causation by persons (see PB: 12), and the claim that persons have independent causal powers, if it were sustainable, would dissolve that problem at a stroke, while at the same time allowing Baker to maintain her claim that persons are entirely material entities (PB: 22). In this section and the next I discuss Baker’s exposition and criticism of the arguments against the possibility of independent mental causation of which Kim has been a prominent proponent (see Kim: SMK, PM, MFW, PSNE, EMM).

Kim holds that nonreductive materialists should accept Alexander’s dictum, from the British emergentist Samuel Alexander; ‘To be real is to have causal powers’ (ESC: 202ff; Alexander 1920). Baker maintains, of course, that constituted entities are real; further, as we shall see, she claims that they have causal powers that are independent of those of their constituting entities. If Alexander’s dictum is accepted, then clearly Baker’s ontological claims, forming the core of her constitution account, stand or fall on whether or not her causal claims can be substantiated. It is in the context of Kim’s view of causality as it is deployed in his arguments against nonreductive materialism (NRM) that Baker sets out her own view (ARP; MEL: 97 - 120).

In MPW and PSNE Kim argues, building upon earlier work, that NRM cannot provide an account of mental causation. His conclusion, in brief, is that mental properties must be either reducible to physical properties or be causally impotent, or epiphenomenal. In the terms of Harbecke’s inconsistent tetrad (2:II:(ii)), the choice is between rejecting either premise (MC), that mental events (or properties) cause physical effects, or (NI), that mental and physical events (or properties) are not identical.

His own solution, we shall see, is to claim that mental properties are indeed reducible, although in a functional, species-specific (or perhaps even individual-specific) way. Baker's reconstruction of Kim's anti-NRM argument involves her identifying from his writings six metaphysical assumptions that underpin six principles or theses, as follows (MEL: 100; ARP: 2-3; NRM: 113-114):

- (1). The Physical Realization Thesis: a mental property is instantiated only if it is realized by a physical property. If P realizes M, then P is nomologically sufficient for M, and M supervenes on P (NRTMC).
- (2). The Nomological-Sufficiency Conception of Causation: A causes B only if A is nomologically sufficient for B (*ibid*).
- (3). The Causal-Realization Principle: if an instance of S occurs by being realized by an instance of Q, then any cause of this instance of S must be a cause of this instance of Q, and *vice versa* (*ibid*).
- (4). The Causal-Inheritance Principle: if mental property M is realized in a system at t in virtue of physical realization base P, the causal powers of this instance of M are identical with the causal powers of P (MRMR: 326).
- (5). The Causal-Closure Principle: any physical event that has a cause at t has a complete physical cause at t (SG: 254).
- (6). The Principle of Causal/Explanatory Exclusion: there is no more than one complete and independent cause (or causal explanation) of any event (MPEE: 250).

The first four principles appear in the essay, 'The Non-Reductivist's Troubles with Mental Causation' (NRTMC). Principle (1), according to

Kim (*ibid*:196), is one of the basic tenets of NRM. The notion of *realization*, in (1), (3), and (4), derives from the idea that the relation of realization links higher-level properties to those lower-level properties on which they *supervene* (see below). For Kim, where a relation of supervenience holds between two properties, instances of the higher are realized by instances of the lower.

A central claim of Davidson's (ME; Chapter 2 above) original formulation of anomalous monism is that mental and physical *predicates* are not connected nomologically. As Kim points out (*ibid*: 5-8), Davidson is deliberately silent about the relation between mental and physical *properties*. However, Davidson adds "almost like an afterthought" (Kim *ibid*: 6) that his account is consistent with there being a supervenience relation between the two. The concept of supervenience as a relation between two sets of properties, in its different forms, has, of course, undergone exhaustive metaphysical analysis in recent years, much of it by Kim himself, but for present purposes all we need is the minimal account he gives (MPW: 9–15) in order to situate the concept in relation to others. Kim is here concerned with mind-body supervenience but, once again, the relation may be taken to apply to relations between higher- and lower-level properties in general.

The relation of *strong* supervenience may be expressed (*ibid*: 10):

Mental properties *supervene* on physical properties, in that necessarily, for any mental property *M*, if anything has *M* at time *t*, there exists a physical base (or subvenient) property *P* such that it has *P* at *t*, and necessarily anything that has *P* at *t* has *M* at *t*.

It is generally agreed that supervenience does not entail reducibility, although it is associated with it in some accounts, including Kim's. The relation is one of necessary covariance, and in addition it is assumed that there is an asymmetric *dependence* of the supervenient property on the base property. As Kim points out, however, (*ibid*: 13) this leaves unaddressed the question of what *grounds* or accounts for the dependence relation. His own answer, as I mentioned, is the functional reductionist one of claiming that; "the mental supervenes on the physical because mental properties are second-order functional properties with physical realizers (and no nonphysical realizers)" (*ibid*: 24). However, although functionalization of mental properties may remove the obstacle, for property reductionists, that mental properties are differently realized in different species, there remain as problems for this theory, first, Davidson's claim that mental and physical properties are heteronomic with respect to each other, and second, the widely-held view that mental content is externally individuated (see 2:III:(iv)). We will see that supervenience has a role in Baker's account, in which relational properties play a large role. A formulation of the supervenience relation that would fit her account would therefore have to be one that allowed properties to supervene *widely* upon, for example, environmental and historical factors.

To return to the principles; (2) – (4) are also, Kim claims (ESC), acceptable to most adherents of NRM. Principle (2), however, seems nearly interchangeable with Davidson's PNCC, which McDowell, for one, rejects, as we saw in chapter 2. As for principle (5), in 'The Myth of Nonreductive Materialism' (MNRM), Kim takes it to be basic to physicalism generally. As he puts it (MPW: 40), it is the claim that "no

causal chain will ever cross the boundary between the physical and the nonphysical”, something that “no serious physicalist” would reject.

Baker presents principle (6), causal/explanatory exclusion, as covering both cause and causal explanation, but in the paper in question, ‘Mechanism, Purpose, and Explanatory Exclusion’ (MPEE: 1993), Kim focuses on explanation, and his claim is that there cannot be more than one complete and independent explanation of any single event. He emphasizes that the argument is not restricted to causal explanations, let alone those concerned only with mental causation, but applies to the more inclusive relation of *dependence* amongst the relevant relata (see EKMD: 182-186). However, causal explanation is the paradigm case and “to have a causal explanation of an event requires that the event specified as its cause be, in reality, a cause of that event” (Kim NPEE: 256). Just as one cannot know that *p* if in reality *-p*, Kim argues, a putative explanation only has that status if the appropriate objective conditions obtain (*ibid*). Thus for two complete and independent causal explanations of an event to exist would require that there were two independent, that is, unconnected or non-interacting, causes for the same effect. The argument is premised on Kim’s assumption of explanatory realism (*ibid*), already mentioned in connection with McDowell (2:I:(iv)). A clear expression of the view is given by Stueber (2005: 245):

Explanations are grounded in and true because of objective and mind-independent relations between events in the real world. A causal explanation of an event E by reference to event C is true only if there exists a real causal relation between C and E that is independent of our explanatory practices. Causal relations between

events C and E hold *in virtue of* certain properties. A causal explanation of event E in terms of C is true only if it cites the properties *in virtue of* which C causes E.

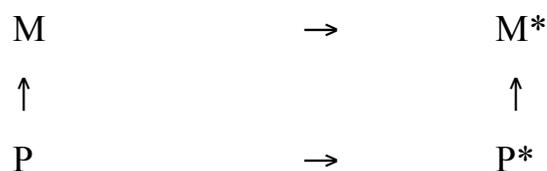
Many would add to principle (6) the proviso ‘except in cases of genuine overdetermination’, to cover the kind of case in which, say, someone is killed by two bullets entering his heart simultaneously. Kim, and many others, have strong intuitions that mental causation cannot involve systematic overdetermination in this way (Harbecke’s premise (NO)). Others disagree; Loewer (2002), for example, argues that, since the supervenience relation between mental and physical properties means that mental properties are not independent from physical, the kind of overdetermination involved is benign. This opens up a number of issues which I return to below (**V:(ii)**), in the context of Kim’s once-held but now rejected notion of supervenient causation. Here, we should note that the combination of this principle with principle (5), causal closure, entails that for any physical event that is caused, its sole cause must be a physical event.

Having assembled these six metaphysical presuppositions, Baker uses them to reconstruct Kim’s ‘key argument’ against NRM (MEL: 102-104). While it is clear enough from Baker’s account of constitution what she herself means by higher and lower ‘levels’ (**II:(iv)** above), as we have noted she takes Kim to understand ‘levels’ differently; firstly, mereologically, that is, that the difference between levels is the difference between wholes and their parts (MEL: 111n41), and secondly as levels of description or explanation rather than of reality. Baker defines irreducibility thus: “a mental property is irreducible if and only if there is no physical property such that *instances* of the mental property

are identical to *instances* of the physical property” (*ibid*:102, emphases added). Note that this definition classifies Davidsonian token identity theory, or at least versions of it couched in terms of properties rather than predicates, as reductive. Baker, like Kim but unlike Davidson, takes an event to be an instantiation of a property in an object at a time, so events can be identical only if the properties instantiated in them are identical. On any view that takes mental and physical properties to be distinct, then, event identity, either type or token, is ruled out (Baker does not consider the Macdonalds’ co-instanting account (see 2:III:(i))).

Baker shows that, based on the six principles, a valid argument can be constructed to arrive at Kim’s conclusion that, if mental properties are irreducible to physical properties, they are causally inefficacious. I will not reproduce her detailed version of this well-known argument (MEL: 102-104; NRM: 115-116). Shortened and informally, one form of the argument is as follows.

Suppose that M and M* are mental property-instances realized by physical property-instances P and P* respectively, but that the *properties* M and M* are distinct from properties P and P*. Suppose P causes P*. According to NRM, M causes M*. The situation can be represented as below, with the horizontal arrows denoting causation and vertical denoting realization



But P is nomologically sufficient for P* (principle (2)), and P*'s instantiation guarantees M*'s instantiation (principle (1)). Therefore P causes M* (principle (3)). Assuming causal closure (principle (5)) and causal exclusion (principle (6)), P is the complete cause of M*. M thus has no causal role of its own and the upper causal arrow is redundant.

On the other hand, if properties M and M* are functionally *identical* with properties P and P*, as Kim claims, M's causal efficacy is assured – it is that of P (principle (4)). For those who insist on the distinctness of mental properties, however, these are unacceptable alternatives.

(ii) Baker's Criticism of Kim's Causal Arguments.

Baker accepts the validity of Kim's argument, given the acceptance of its premises, the six principles. But she finds the conclusion of the argument, and the implications she draws from it, unacceptable, and so is forced to reject some of the premises. The unacceptable implication that she highlights is that the argument generalizes to macrocausation as a whole. That is, it is not just causation by mental property-instances that is invalidated by the argument, but also causation by an enormous class of ID or higher-level property-instances – “e.g., being in debt, being a driver's license, being a delegate” (MEL: 106). Following Kim (MPW: 77ff), let us call this the generalization argument.

Kim himself does not accept that his conclusions about mental causation threaten all macrocausation (MPW: 77-87). That idea, he argues, is based on the mistaken assumption that

the mental-neural relationship is, in all relevant respects, the same relationship that characterizes, say, the chemical-microphysical, biological-physicochemical, or other interlevel cases (*ibid*: 80).

His argument that it is not the same relationship is based on his adoption, described above, of the functional reductive account of mental properties, according to which they are second order properties characterized in terms of their functional roles. Lower, or first order physical properties are the realizers of these functional, or causal, roles. The mental-neural relationship is then one between a second order property (such as being in pain) and its first-order realizer (being in a certain neural state). But this relation “does not track the micro-macro relation” (*ibid*: 82) in the way that interlevel relationships like the chemical-microphysical do. Rather, “a second-order property and its realizers are at the same level in the micro-macro hierarchy; they are properties of the very same objects” (*ibid*), that is, in the mental-neural case, they are properties of the subject.

Chemical and similar higher-*level* properties, on the other hand, are what Kim calls micro-based macroproperties. For example, the (macro) property of being a water molecule is “the property of having two hydrogen atoms and one oxygen atom in such-and-such a bonding relationship” (*ibid*: 84). The former property is completely decomposable into the micro-parts and their properties and relations. As an example of the *causal* properties of such higher-level entities, a table with a mass of ten kilograms has the macro-causal property of making the scale read 10Kg, which is not shared by any of its microconstituents, yet is obviously derived from and decomposable into the sum of those microconstituents’ properties. The causal powers of the table thus go

beyond, and are not reducible to, those of its parts, and the same applies to biological, geological, and other special science properties, and, indeed, to the entire range of objects and properties of the physical domain (*ibid.*: 113-116). The essential point, for Kim, is that while the properties of macrophysical objects supervene mereologically on those of their parts, thus being determined by them, they are not reducible to them (*ibid.*: 116-117). The causal properties of macro-objects are properties at that level, a higher level than the properties of the microconstituents. Kim's view of ordinary objects, then, turns out to be surprisingly like Baker's own on her constitution account. And since these objects are all physical, being composed of physical microconstituents, their having their own causal powers, like those of the table, is perfectly consistent with the causal closure of the physical realm.

As Kim writes, "this means that the case of micro-based properties is not at all parallel to the case of supervenient psychological properties" (*ibid.*: 117). However we might eventually characterize the supervenience of mental on physical properties, we can be sure it is not mereological. As we saw, on Kim's view mental properties and those of their neural realizers are at the same level, possessed by the same entity, the subject, and thus (according to the Causal-Inheritance Principle, Baker's (4)) share *the same* causal powers. Kim's anti-NRM argument then applies; if we insist mental properties are distinct and *cannot* be reduced, the principles of causal exclusion and physical causal closure together dictate that they are epiphenomenal. But on the basis of the distinction we have just arrived at between mental properties and micro-based physical macroproperties, this finding, according to Kim, does not apply cases of macrocausation generally.

Baker, however, challenges neither Kim's argument for his theory of functional reduction, nor his use of the concept of micro-based macroproperties as a defence against the generalization argument, directly. Instead, she objects obliquely by drawing attention to "a huge class" of properties "without which we cannot begin to make sense of the everyday world" (MEL: 106). These are the ID properties that feature in causal explanations in human sciences like economics as well as in ordinary life, and they are not covered by Kim's refutation of the generalization argument. Properties like being in debt are not, Baker argues, readily construed as micro-based – they are not, unlike the property of weighing 10Kg, decomposable into lower-level physical constituents. Nor, she claims, are they amenable to the functional reduction that Kim applies to mental properties. This would involve identifying a physical realizer that fulfills the role of being in debt, and, Baker argues, "it is difficult to think of a *candidate* to be a physical realizer [of this role]" (*ibid*: 108). On Kim's account, then, ID properties are causally powerless, - and thus, if we accept Alexander's dictum, ultimately unreal. But there is, she claims, "overwhelming empirical evidence that ID properties are causally efficacious...without ID properties, we would have no causal explanations of...any historical, economic, social, or legal phenomenon" (*ibid*: 107, 109).

The physical realizer of an ID causal event such as a Savings and Loan (S&L) institution's becoming bankrupt (EA: 126-136) is not the constituting base of the event, Baker points out. The physical constituter of the bankruptcy might be an electronic process that constitutes a transfer of funds, say. But such an event only constitutes a bankruptcy in circumstances where a vastly complex network of social, financial, and

other factors obtain. As we saw in the case of the national flag (**II:(ii)**), these favourable circumstances comprise the wide supervenience base of the S&L bankruptcy, a base that “would include properties instantiated over vast reaches of space and time, extended far enough to include all the physical property instantiations on which economic practices depend – perhaps back...to the Big Bang” (*ibid*:134). Baker takes this conclusion – that the causal underpinning of an ID event like the S&L collapse extends far beyond the narrower set of physical events that *constitute* the collapse, to be a refutation of Kim’s argument, and a vindication of the irreducibility of our commonsense notions of higher-level causation, primarily because of the impossibility of our ever being able to identify the causal realizers.

Baker’s arguments, discussed in (*i*) above, are directed not so much at the substance or validity of what she calls Kim’s key argument against NRM as at the argument’s implications. It is a non-negotiable position, for her, that the very large range of non-mental ID properties – being a driver’s license, being an S&L collapse, etc - are both irreducible and causal. As we saw, the main thrust of her disagreement with Kim is that if we accepted his argument we would be forced to conclude that the properties in this huge group are epiphenomenal, and this conclusion constitutes a *reductio ad absurdum* for Baker, because we cannot do without these ID causal concepts. Her presentation of, for example, the S&L bankruptcy case certainly shows the near impossibility of replacing an ID causal *explanation* in terms appropriate to economics by one framed in physical terms. But Kim’s arguments are premised on the idea that causal explanations rest on one, and only one, real causal process.

It is open to Baker's opponents, then, to reply that the utility, or even indispensability, of concepts are themselves no proof that the concepts reflect reality. While it is convenient and indeed essential for us to say that, say, John's paying his debt *caused* the harassing phone calls he was receiving to stop, this is, the objector could say, a reflection of our explanatory practice. A follower of Davidson could argue that, because of the extensionality of causality, this causal description is as valid as a microphysical one, but as we saw (Chapter 2:III:(i)), that is open to the objection (McLaughlin 1993) that it is valid to ask, in virtue of *what* does one event cause another, and for the reductionist the answer will be that it is in virtue of physical properties. Baker's reason for insisting that ID properties are causally efficacious thus begs the question in favour of the reality of ID properties themselves. The eliminativist or reductionist is free to claim either that there are no such properties or that they reduce to microphysical properties, and that each *instance* of what appears to be the property in question is either real but reducible to a configuration of its microphysical realizers or is nothing other than a configuration of such entities.

Secondly, Baker's deployment of the epistemic argument that we can have no idea what the realizer, or supervenience base, of an ID property like being the collapse of S&L might be, is not itself effective, given that Baker concedes that such a realizer exists, even if we have to envisage it as instantiated over space and time back to the Big Bang. After all, a non-reductive physicalist must believe that *mental* properties like understanding how financial institutions work, or deciding to pay off a debt, are physically realized in the individual who instantiates them. The question of causation in a whole ID system like that of the financial system, with its practices of owning and borrowing, is arguably just the

question of mental causation writ large (cf Macdonald and Macdonald 2010:143). Its realizers include not only objects, ID and otherwise, but the realizers of the mental properties of very many individuals, including many no longer alive. Thus the realizers will be spread out in time as well as space. But the fact that the realization base of the property is so vastly complex that its exact nature will almost certainly never be known is not in itself a reason to doubt that it exists. And recall that it is integral to Baker's constitution account that constituted entities, including persons, are both natural and have nothing but natural components.

Baker confines her arguments here to ID causal properties, but the claim that all constituted objects have irreducible causal powers is part of her account. And it seems that the above considerations also count against similar claims on behalf of, say, biofunctional objects and their causal properties, such as a gene's property of transmitting genetic information, or the Scarlet Gilia's property of adapting to a regularly recurring environmental change. The reductivist or eliminativist can claim that those causal powers are either reducible to those of the gene's realizing microstructure or, again, not real properties but merely predicates reflecting our interests. Like the realizers of being in debt, the realizers of the Gilia's ability include spatially remote and historical factors, but our lack of full knowledge of these is not in itself a reason for denying that a reduction is possible. One can object that Baker is unjustifiably assuming here, as in several other places in her writings (e.g., EA:114 & *passim*), that the non-availability of an epistemological reduction precludes the possibility of a metaphysical reduction. One reply Baker makes to this objection is to put the onus on the objector to "give some reason to think that the strategy [of metaphysical reduction] can succeed...without the "merely epistemological" one has little grounds

for confidence in the loftily metaphysical...if we have no clue about how to find a reduction, we are in no position to claim that it can be carried out in principle” (MEL: 108n33). But, to repeat, there is a distinction between the feasibility of carrying out the reduction and such a reduction base’s actually existing, even if permanently beyond our reach. There is, I think, a basic physicalist intuition that even so complex a property as that of being in debt *must* be somehow rooted in the physical world.

Baker’s arguments on behalf of the irreducibility of macrocausation seem at least partly motivated by the same concerns that lie behind her advocacy of the constitution view – the conviction that “metaphysics should not swing free of the rest of human enquiry..[it]..should be responsive to reflection on successful cognitive practices, scientific and nonscientific” (MEL: 15; see II:(i) above). Thus she writes, “Kim’s view would have us transform a causal connection that we all understand...into a causal connection between totally unknown physical properties” (*ibid*: 108). But mere appeal to such practical considerations, I think, is not enough to counter what appear to be sound and well thought-out metaphysical arguments based on principles that, for many, have stood the test of intuitive acceptability. In the next chapter (4:I:(v) & (vi)) I draw a distinction between two uses of “cause” and its cognates, both indispensable, only one of which, I claim, refers to genuine causation in the manifest physical world, while the other denotes a perfectly legitimate but less rigorous use that occurs in conjunction with causal explanation. I suggest that Baker’s claims that being in debt, etc., are causal properties results from a conflation of these uses.

Whatever the merits of her arguments against Kim, however, Baker needs to deploy her own constructive arguments if she is to refute his position. This is what she attempts with her property-constitution view and account of independent causation, to which I now turn.

IV. Baker's Property-Constitution Account.

(i) Independent Causal Efficacy.

I have already introduced the property-constitution (PC) variant or aspect of Baker's constitution view in section **II:(ii)** above. To recapitulate, while the original formulation of material constitution describes how an object at a higher ontological level, a bearer of novel properties, is constituted by, while not being identical with, an object at a lower level, (PC) is concerned with the conditions for the constitution of *instances*, or token occurrences, in an object at a time, of higher-level properties themselves by instances of lower-level properties. (PC) is thus "a time-indexed relation between property-exemplifications" (NFPP: 209). There is thus a change in emphasis from the constitution of objects bearing properties to the constitution of instances of properties borne by objects. (PC) is a later development of Baker's account, appearing in ARP (2006) and subsequently, and it forms the basis of her argument for the independent causal efficacy of ID properties (MEL:111–116; NRM), and thus of her defence of her version of NRM.

Three aspects of (PC), according to Baker, distinguish it from Kim's picture. First, as I have already discussed, (**II:(ii)** and **III:(i)** above) Baker and Kim differ on the ontological status of levels or layers of reality. Secondly, as we saw (**III:(ii)** above), on Baker's account, the

properties of ID constituted entities do not supervene only on those of their constituters, but more widely; here “supervenience and constitution come apart” (EA: 133). Kim’s concept of realization, as applied to mental properties, is replaced by the “much weaker” (MEL: 113) notion of property constitution. While the necessity which binds mental or other higher-order properties to their realizers is, Kim claims, nomological (MPW: 23-24) rather than logical or metaphysical, it is nevertheless strong enough to entail the kind of tight dependence of higher-level property-instances on lower, embodied in principles (1), (3), and (4) above, that supports Kim’s anti-NRM arguments. On the other hand, on the (PC) account, the occurrence of an instance of the constituting property does not itself entail an instance of the constituted property, but requires in addition the contingent presence of additional favourable circumstances, such as relational ID factors like the existence of certain conventions, as explained in section II.

The third way in which (PC) differs from Kim’s account is that Baker rejects the causal inheritance principle ((4) in her reconstruction of Kim’s argument): ‘If mental property M is realized in a system at t in virtue of physical realization base P, the causal powers of this instance of M are identical with the causal powers of [this instance of] P’. The rejection and replacement of principle (4) marks Baker’s starting position for her arguments against Kim. Her justification for the rejection is her argument, above, that Kim’s argument is sound but its conclusion is unacceptable, hence at least one of its premises must be wrong, and the causal inheritance principle is the one most glaringly at odds with the constitution view. Baker replaces the principle with one of *independent causal efficacy* (*ibid*: 115):

(IC) A property-instance that has an effect *e* has independent causal efficacy if and only if (i) it would have had its effect *e* even if its constituting property-instance had been different, and (ii) it confers causal powers that could not have been conferred by its property-constituting instance alone.

Baker's claim, then, is that instances of ID properties have independent causal efficacy, in the sense that,

the causal powers of higher-level property-instances cannot be reduced to the causal powers of their constituters [constituting property-instances]. Constituted property-instances confer causal powers that are “over and above” those of their constituters. The effect of a vote exceeds the effect of the constituting hand motion alone (*ibid*).

The novel causal powers acquired by constituted higher-level property-instances are, nevertheless, *physical* powers. Baker writes,

On my view, all property-instances are physical in this respect: any property-instance is either identical to or ultimately *constituted by* microphysical property-instances. Higher-level properties – even mental properties - thus are physical properties. So, the causal efficacy of higher-level properties does not violate the causal closure principle (NFPP: 217).

That is, although the new powers are “over and above” the lower-level ones, they are nevertheless of the same basic kind. The claim that constituted properties are physical certainly accords with clause (f) of

Baker's definition (C) of *material* constitution (II:(ii) above), that constituting and constituted entities are of the same "basic kind of stuff". Baker explains (NFPP: 209) that "it is not literally properties themselves that are constituted, but property-instances". Her view is thus similar to the Davidsonian position as developed by, among others, the Macdonalds (see 2:III:(i)), with the difference that for Baker it is the constitution relation, rather than identity, that ensures that constituted property-instances are physical. As we shall eventually see, however (5:I:(iii)), the claim that ID property-instances are physical in this sense can be shown to constitute a fatal flaw in Baker's argument.

The claim (IC), that constituted property-instances have novel and irreducible causal powers that are conferred by constitution, is probably the most important claim of Baker's constitution account, especially when applied to persons, as in III:(i) above. She writes,

My thesis, then, is this: ID properties generally (with mental properties as a special case) are causal properties because their instances have independent causal efficacy (MEL: 115).

Baker uses two examples in support of (IC) (*ibid*: 115-119; ARP: 13-15; NRM: 123-126) which will play a large part in my discussion of her argument here and later. In her first example, her basic strategy is to assume, or stipulate, as a premise that higher-level, constituted causation occurs, and then show that the constraints of (IC) are satisfied. Let

V be Jones' voting against Smith at t.

P be Jones' hand going up at t.

V* be Smith's getting angry at Jones at t'.

P* be Smith's neural state at t'.

C be [vote-favourable] circumstances that obtain at t in which a vote is taken by raising hands.

Suppose V is constituted by P and V* by P*.

Before discussing the example, two points need to be mentioned. First, it is worth noting that one of Baker's initial premises, that P*, Smith's neural state at t', constitutes V*, Smith's getting angry at t', appears to contradict her earlier rejection, notably in EA, of what she calls the Standard View (I:(i) above), that “[propositional] attitudes...are (or are constituted by, or are realized in) particular brain states” (EA: 5). “(A) belief is a global state of a whole person, not of any proper part of the person, such as the brain. Persons have beliefs; brains have neural states” (EA: 153). I think the explanation for this apparent inconsistency lies in the fact that the argument against the Standard View is against the *reification* of propositional attitudes and hence against *material* constitution as it applies to them, whereas this example is concerned with property constitution (PC). It is a property-instance instantiated by Smith, his neural state, that constitutes another (higher-level) property-instance instantiated by Smith, his getting angry at Jones. In view of Baker's above remark, she arguably should have stipulated that P* is Smith's total physical state at t' rather than his neural state. We should note that on Baker's account the relation between token mental and physical states is one of constitution, and hence is weaker, or looser, than on token identity or realization views.

Secondly, it seems reasonable to call these property-instances – V, etc, - events, since on most accounts (see, e.g., Schaffer 2008: 2) events are the bearers of causal efficacy, or the relata of causation. Property-

instances, for Baker, are “property-exemplifications, schematically ‘x’s having F at t’” (NFPP:209), while similarly “causation by properties [is] event-causation...events are...complexes, as, e.g., an object’s having a property at a time” (MEL: 97 n1). She thus follows Kim’s influential property-exemplification account (see EPE), which I discussed in connection with the Macdonalds’ co-instancing account of mental causation (2:III:(i)). Events, however, are usually taken to be *changes* such as an object’s *coming to* exemplify a new property in some category or ‘property space’ (Lombard 1986). P, for example, is an event that is a change in the property space of spatial location, while V is a change in a ‘psephological’ property space, from ‘not having voted’ to ‘having voted’. So what Baker calls property constitution (PC) also, in some cases, merits the name ‘event constitution’. According to Kim’s schema for events, the event V should strictly be designated {VJt}, and V* {V*St’}, but I will stay with Baker’s formulation.

As Baker presents it, then, constitution looks like a promising candidate to be the looser than identity relation between the mental and the physical that was discussed in 2:III:(iv). Hornsby argues in WPEME (2:III:(iii)) against the possibility of pinpointing a neurophysiological event or group of events with which a mental event could be identical, but if the relation is constitution the matching requirement need not be so stringent – all that is needed is the idea that the instantiation of *some* neural property – event or state – constitutes the instantiation of the mental property in favourable circumstances. Further, my claim (2:III:(iii)) that the notion of a mental event is problematic is defused. On the constitution account, the only concept of a mental event that we require is the one deriving from ordinary, subjective experience - we

don't need also to think of it as a metaphysical item that is suitable to be identical with some physical event.

Now, to turn to the example itself, Baker begins from the *assumption* that V causes V*, that is, that V's constitution by P in circumstances C bestows on V, Jones's voting, the causal power of making Smith angry, V* (MEL: 116). If this is granted, then clearly the constituted event V fulfils the conditions of (IC). Firstly, V could have been constituted by something other than a raised hand, such as a spoken 'nay' or an electronic signal, and still had the same effect, through its being a voting against Smith, of causing V*. Secondly, P alone, in the absence of the favourable circumstances – the voting environment - whereby it constituted V, would not have caused V*, but, as a hand-raising, could have constituted a contribution to some other social convention or none.

Baker argues that if (IC) is correct this necessitates the rejection not just of Kim's causal inheritance principle (4), which (IC) replaced, but also principle (1), of physical realization, and principle (3), of causal realization, on the grounds that both of these preclude independent, higher-level causally efficacious properties (MEL: 117). On Kim's account, the relation of realization dictates that instances of mental properties are realized by instances of physical properties (principle (1)) and that the causes of instances of higher-level properties are the causes of their realizing property-instances (principle (3)). On Baker's (PC) account, 'constituted' would replace 'realized' in this description of principle (1), and we saw (III:(ii) above) that where ID properties are concerned, constitution and realization are quite different. Most importantly, the instantiation of a constituting property is not alone sufficient for the constitution of a higher property-instance, because the

presence of favourable circumstances is also required. Also, on Baker's account the causal powers of higher-level property-instances are independent of those of their constituting property-instances, in opposition to principle (3). So the three principles that Baker rejects here, (1), (3), and (4), are just those which deny the independent efficacy of higher-level causes.

If Baker's argument on behalf of (IC), the Principle of Independent Causal Efficacy, is successful, then, Kim's anti-NRM argument is called into question. But the argument, I think, begs the question precisely in that it begins by assuming just what needs to be established in order to show the validity of (IC), which is that V independently causes V^* . As we saw in **III:(iii)**, Baker's rejection of Kim's metaphysical arguments against NRM is based upon the indispensability of the notion of mental and ID causation in everyday life, and this same motivation drives the argument here. If the validity of the argument from indispensability is denied, the stipulation that V causes V^* is inadmissible.

Baker also rejects Principle (6), Causal/Explanatory Exclusion, to which I return in chapter 4. She accepts the remaining two principles – (5) the Causal Closure Principle and (2) the Nomological-Sufficiency Conception of Causation. She does not explicitly discuss the latter principle, but does present an argument, to which I now turn, aimed at showing that her account of independent causation does not violate the widely-accepted Causal-Closure Principle, that any physical event that has a cause at t has a complete physical cause at t .

(ii) Independent Causation and Physical Causal Closure.

Baker first argues that in our example the causal closure principle is not violated because only physical causal properties contribute to and suffice for V^* 's instantiation. This follows directly from her claim that all property-instances are physical. If this claim is granted, it would seem to be trivially true that physical causal closure is not breached in higher-level causation. The problem of mental causation is dissolved by stipulating that the problematic properties are physical after all. I return to my reasons for believing Baker's account of independent causation is not compatible with this view in my final review of the account in 5:I:(iii), after I have introduced my account of ordinary ("manifest") physical causation in chapter 4.

In the meantime, however, anticipating objections to the claim that ID properties are physical, Baker argues further that even if this claim is rejected, and only micro-based properties and property-instances and their aggregates are taken to be physical, causal closure is still not violated. To illustrate this, she introduces the second of the examples which will feature prominently in our discussion (MEL:118-119), a simple case of basic action involving only the subject's own body. Jane is asked to raise her arms for an airport security check. Let

M be Jane's willing (choosing, intending, etc) to raise her arms

M^* be Jane's raising her arms

MP be the microphysical constituter of Jane's willing to raise her arms

MP^* be the microphysical constituter of Jane's raising her arms

Suppose that Jane's willing to raise her arms causes her to raise them.

As with P and P* in the previous example, we should take MP and MP* to be events which are exemplifications of properties in Jane. Again, note that M's being the cause of M* is a premise of the argument. The example is designed to show that although MP is not the complete cause of (is not nomologically sufficient for) MP*, nevertheless M's causing M* does not violate physical causal closure. To show that MP is not nomologically sufficient for MP*, Baker asks us to consider a world with the same laws as ours in which Jane's brain is in a vat. In that world, MP does not cause MP* because Jane has no arms to raise (*ibid.*: 118). Of course (Baker does not add), in that world M would not cause M* either, so if all causal efficacy requires nomological sufficiency, and if causal efficacy in the vat-world is a condition of nomological sufficiency, then we would be forced to conclude that M does not cause M*. Yet, *ex hypothesi*, M *does* cause M*, and since, also *ex hypothesi*, M and M* are physical property-instances, the causal relation between them is bound by the principle of nomological sufficiency (Kim's Principle (2), which Baker does not reject), it appears that Baker cannot rely on this argument to establish that MP does not cause MP*.

Nevertheless, the point Baker is making is correct, as the voting example shows more plainly. It is easy to see that P, Jones's hand-raising, is not the *complete* cause of P*, Smith's neural state. V's causing V*, then, is not just a consequence of their micro-constituters causing each other. In email correspondence (21.8. 2012) Baker writes "I don't think that V's causing V* depends on any microphysical relation between P and P*", which, of course, is implied in principle (IC), of Independent Causation. Nevertheless, even if it is granted that (IC) is correct and the ID causal relation is independent and irreducible, I don't think it can be denied (nor does Baker deny it) that a causal link between P and P* *exists*, in the

form of a physical causal chain or mechanism. At that level, P* is caused by P together with a range of other factors, most obviously the light patterns caused by P reaching Smith's retinas and activating rod and cone cells, action potentials occurring in Smith's optic nerves, physical property-instances being instantiated which constitute his recognizing the meaning of P, and so on. I will call the complete cause of P* 'P&ae' (ae for 'additional events'). Similarly, Baker says of the arm-raising case, "MP is only a proper part of a larger collection of microproperties [whose instantiation] *is* nomologically sufficient for MP*" (MEL: 118, emphasis added). We can call this range of property-instances MP&ae.

It is worth noting that while in the voting example, Baker calls P* the *neural* constituter of P, in the Jane example MP and MP* are, respectively, the *microphysical* constituters of M and M*. Perhaps this difference is merely an oversight, but I think it might reflect an ambiguity in the term 'microphysical'. In one use, 'microphysical' can refer to the most fundamental level that physics can (currently) reach, that of subatomic entities such as quarks and leptons, at which our common sense notions of physical categories, including that of causality, may no longer apply. In another use, it can simply mean 'visible only by microscopy', in which case it includes the reasonably well-understood and (with appropriate scientific techniques) *observable* behaviour of neurons and their connections, at a level of organization many orders of magnitude higher than the fundamental, and at which we *can* apply notions such as causality. I suggest, then, that 'neural' is the term that most accurately describes the constituters of Jane's willing and of her arm-raising. If we were seeking to explain how nervous systems work, it seems appropriate to begin looking at the level at which the unique properties of neurons and systems of neurons that mark them out from

other tissues – basically, the ability to code, store, and transmit ‘information’ in the form of patterns of variations of electrical potentials across neuronal membranes - come into view, rather than, say, the level of intracellular metabolism, which neurons share with many other kinds of cell, let alone the microphysical level which they share with all matter. Nervous systems contain identifiable *mechanisms* that enable them to do this. So MP&ae, in my characterization, is intended to refer to causal mechanisms (see chapter 4:II:(iii)) involving neural and other well-understood phenomena, rather than events described at the level of fundamental particles. Some of the events picked out by ‘P&ae’, in the voting example, are also neural, while others, such as the transmission of light waves, are not.

To return to the example, causal closure is not violated, Baker now argues, because MP* *has* a complete physical cause, even though MP is not that complete cause. She writes,

There is no difficulty for the property-constitution view in saying: (i) Jane’s willing to raise her arms is constituted by MP; (ii) Jane’s raising her arms is constituted by MP*; (iii) Jane’s willing to raise her arms causes her to raise her arms; but (iv) MP does not cause MP*. If the microphysical state of one sizable spatiotemporal region that ends at the time of Jane’s willing caused the microphysical state of a slightly later sizable region that begins at the time of Jane’s raising her arms, then the Causal-Closure Principle is honoured (MEL: 118-119).

I think we can take the microphysical state of the first “spatiotemporal region” to consist in the instantiation of what I have called MP&ae (to be

precise, M immediately followed by ae), and the second in that of MP*. So MP&ae is nomologically sufficient for MP*, and the relation honours the causal closure principle – unsurprisingly, since its relata are indisputably physical events. But what Baker needs to show, and what she has not shown by the above argument, is that the purported ID causal relation, M’s causing M*, given that it is, *ex hypothesi*, an independent, novel, and irreducible causal relation, does not violate the principle. Not only has it not been shown that this higher causal relation does not violate physical causal closure but, as in the voting example, there is nothing in support of the stipulation that there even *is* such a causal relation as M’s causing M*. Baker appears tacitly to assume that the fact that the ID events, willing and raising, are constituted by events, MP and MP*, that are *parts* of a complete microphysical causal relation between two “sizable spatiotemporal regions” is sufficient to show this. I argue that this does not follow. Even though MP bears to M, and MP* to M*, the close relation of constitution, or unity without identity, we cannot assume that what can be said about the lower-level causal relation also applies to the upper. First, Baker stipulates that the ID causal relation is independent of any lower-level relation, and second, nothing at the ID level corresponds to the additional events ae, which is an essential component of the lower-level causal chain and which does not constitute anything at the ID level.

We can better understand this by looking at the complex interplay of the relations of causation, constitution, and supervenience in Baker’s account. Recall (I:(i) above) that, for Baker, for *x* to have a *relational* property entails, minimally, that there is some *y* distinct from *x*. Baker defines *supervenience on local microstructure*:

A property P supervenes on local microstructure if and only if: Necessarily, if x has P and y lacks P , there is a microphysical difference between x and y (EA: 63).

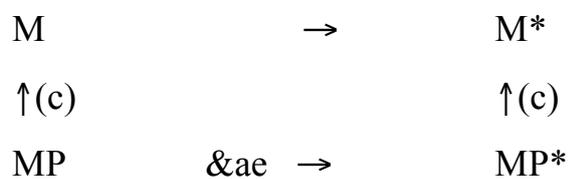
Baker insists that relational properties (see II:(i)) do not satisfy the biconditional. It is possible for there to be two microphysically identical structures, one of which is a planet, with the relational property of orbiting a star, and the other not. A lump of rock that is not a planet could exist in a world in which nothing else existed, but a planet could not. Relational properties, then, do not supervene on local microstructure, and Baker claims that they include *causal* properties, agreeing with Fodor that, “(i)t is not in dispute that [relational] properties like being a meteor or being a planet ...constitute causal powers” (*ibid*: 64; Fodor 1991). In the Jane-at-the-airport example, then, the causal property M does not supervene on MP , its constituter, or local microstructure, because the properties instantiated in M , an ID property-instance, include, as well as its ID causal property, other relational properties such as Jane’s being in environmental circumstances in which she is asked to raise her arms and understands the appropriateness of being asked, and, more generally, her being a bearer of propositional attitudes, properties she could not have if she was the only object in existence (see *ibid*: 63). It is because of these relational factors that Jane’s mental state is describable as a willing to raise her arms to be searched.

There appear to be three distinct ways in which underlying microproperties (or neural properties) figure in an account of M ’s causing M^* : (1) property-instance M is *constituted* by property-instance MP , its local microstructure, in favourable circumstances. Here the

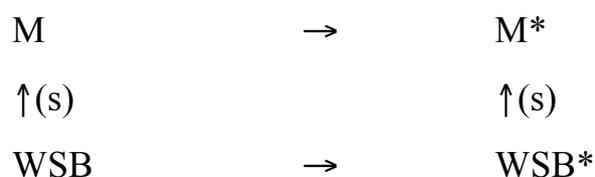
microproperties in question are neural, and intrinsic to MP. But as we have just seen, property M, being relational, does not *supervene* on property MP alone. (2) A causal chain or pathway of neural and other physiological property-instances MP&ae, also internal to Jane's body, leads, as we saw, from the instantiation of MP to the instantiation of MP*. But M, although also a causal property, does not supervene on these causal microproperties either. Here, 'ae' denotes such property-instantiations as transmission of neural activity from other parts of Jane's brain to her motor cortex, and thence transmission of peripheral nerve action potentials to her arm muscles. The supervenience base of M cannot include these additional components 'ae' of MP&ae's causing MP*, since the instantiation of ae is *caused* by the instantiation of MP, and hence occurs at a time later than the constitution of M by MP (which, given the favourable circumstances, and given that M just *is* MP in the constitution sense of 'is', is instantaneous). (3) Rather, if M supervenes on some physical property or properties, that role must be filled by the *wide* range of microproperties of its constituting micro-property together with the supervenience base of "all the circumstances in which the instance of the constitution relation obtains" (MEL: 119) – that is, of the 'willing to raise arms for search'-favourable circumstances. These are the factors that form the physical base upon which social practices such as airport security checks supervene. As I will argue in chapter 4, these properties, most of which are not neural, need not necessarily be thought of as fundamental or ultimately basic. It seems that it is *their* instantiation, and the fact that their instantiation causes the instantiation a moment later of a slightly different set which makes up the wide supervenience base of M*, that ensures that when M causes M* physical causal closure is not violated. The two successive regions that

form the supervenience bases of M and M* are very much larger than those of MP and MP*, though the former include the latter.

I think we can now better understand the source of difficulty in Baker's conclusions from the Jane example. We have two relations connecting the lower- and ID levels. Firstly, there is *constitution* of M and M* by MP and MP* respectively, while we have established that MP* is caused by MP. Using horizontal arrows for causation and vertical for constitution (c), and understanding that the relata are all property-*instances*, this may be represented:



The lower-level causal process here clearly does not violate physical causal closure. But the stipulated causing of M* by M is, *ex hypothesi*, independent of the lower-level causal relation, and nothing in the constitution relation dictates that, if this relation indeed exists, *it* does or does not violate the closure principle. There is also a second relation – here the vertical arrows denote instances of a supervenience (s) relation between *properties*, and WSB stands for ‘wide supervenience base’:



It is the causal relations linking the spatiotemporal regions WSB and WSB* on which M and M* supervene that ensures that M's causing of

M* does not violate causal closure in the situation Baker depicts, rather than any relation between the supervenience bases of their constituters. Thus, if M's causing M* honours the closure principle its doing so is not directly connected with their constitution relations. The wide supervenience bases of M and M* include, for example, the physical bases upon which the whole history, sociology, and politics of air travel and security checks, represented here by WSB and WSB*, supervene. And while Baker claims that M and M*, as constituted property-instances, are irreducible to their constituting property-instances – this is central to her constitution account as set out in section II above –, she makes no explicit claim about the relation of property-instances such as M and M* to the *supervenience* bases of the properties they instantiate. Her conclusion therefore seems vulnerable to a version of Kim's well-known argument (III:(i) above; see, for example Kim, MPW: 38-47) that M, in this example, must either be reducible to its supervenience base WSB or else be epiphenomenal, not a real cause at all. I say a *version* of Kim's argument because Kim takes putative mental causes like M to supervene on *local* microstructure, but I think it is clear enough that if M supervenes *widely* on WSB, and supervenience implies reducibility, then the supposed causing of M* by M either reduces to or is preempted by the causing of WSB* by WSB, as in Kim's argument. Since the supervenience relation dictates that the instantiations of M and M* are *necessitated* by the instantiations of WSB and WSB*, the upper horizontal arrow in the second diagram is redundant, according to Kim – M* will be instantiated regardless of whether or not it is a causal effect of M. On this argument, then, Jane's willing cannot be an independent cause of her arm-raising, and so the question of whether it violates the causal closure principle does not arise.

I think we can conclude that Baker's attempt to show, with the Jane example, that causal relations amongst ID entities honour the principle of physical causal closure does not succeed. I have argued that it is the causal relationship between the wide supervenience bases of M and M*, and not the existence of the causal mechanism M's causing of M*, that is relevant to the ID causal relation's honouring causal closure. But Kim's argument against NRM is directed precisely at the claim that irreducible mental property-instances, supervening on (i.e., realized by) physical property-instances, are independently causally efficacious and causally relevant. Baker's defence here is the epistemic one that we are unlikely ever to be able to identify the wide supervenience bases of M and M*, even though she acknowledges that they exist (MEL: 119; NFPP: 219-220).

Serious doubt is thus cast on Baker's conclusion (MEL:120) that her account of property constitution "vindicates irreducible intentional causation with mental causation as a special case". We have no reason to accept the validity of Baker's principle of independent causal efficacy (IC) except our common sense-based intuitions about mental and ID causation, but these intuitions, are, or should be, what the arguments were designed to confirm and bolster. The existence of the causal connection between M and M*, which Baker acknowledges but discounts as irrelevant to M's causing M*, poses, I believe, a further threat to the higher causal claim (see 5:II:(i) below).

The fact that Baker does not acknowledge that there might be a problem with her claim to have vindicated ID causation is, I believe, related in a deep way to her adherence to a difference-making account of causality, in which the notion is explicated mainly by counterfactual analysis. On

this view, all that is needed to establish that, say, Jane's willing causes her raising of her arms, is a simple claim to the effect that in the closest possible worlds to the actual one, if Jane had not willed to raise her arms she would not have raised them. Because she takes explanation to be prior to cause (e.g., in MMC) and takes counterfactual analysis to tell us all that can or need be said about causation, I think, Baker misses important differences between ID and lower-level causation that emerge only when we address the question of the grounding of the counterfactuals, so that the issue of how content could be a cause of a physical event need not arise for her. But as I will argue in the next chapter, we have deeply-held and, I think, empirically well-founded intuitions about what is required to uphold the claim that something is physical, or that something is a physical cause, not (necessarily) in a sense appropriate to an ideal or fundamental physics, but in an ordinary, or manifest image (Sellars: PSIM) sense. These include that there should be a discernable mechanism (in however inclusive a sense of "mechanism"; see Glennan (2009) and 4:II:(iv) below) whereby that thing is able to interact causally with the rest of the physical world – that we should in principle be able to see at least how we could set about understanding *how it works*. We tend to use the terms 'magical' or 'miraculous' for purported causal interactions that do not measure up to this test, and the mere claim that willing (say) *qua* willing, is physical is not enough to establish that it would not belong in these categories.

My conclusion is that not only does Baker not establish that independent causation, as she conceives it, does not violate physical causal closure, but that she fails to construct, from the starting point of the metaphysics of events and physical causality, a viable argument that commonsense independent ID causation even exists. One way of putting this is that

while Baker is, I think, on the right lines with her formulation of independent causation (IC; (i) above), she does not take her own notion of causal independence far enough, but tries to portray it as in some ways independent of, but in others just the same as, lower-level causation. I argue that meaningful content *as such* is not the kind of thing that can be part of physical causal relations, and in chapter 5 I will try to show how the constitution account can provide an alternative explanation of how content can figure in causal relations.

V. The Distinguishing Features of ID Causation.

(i) *Constitution and Causal Levels.*

Before going on to a discussion of causation in chapter 4, I want to highlight the ways in which Baker's own constitution account already contains the resources to overcome the problems I have identified with Baker's causal argument against Kim. The second condition of the independent causal efficacy principle (IC) is that a property-instance has independent causal efficacy if and only if it confers causal powers that could not have been conferred by its property-constituting instance *alone*. This way of putting it seems at first sight to suggest that the property-constituting instance alone does indeed make *some* causal contribution to those of the causally independent constituted property-instance, but that this is insufficient for the effect. Applied to our voting example, then, one interpretation of Baker's claim is that P, Jones' hand-raising, alone makes some *direct* causal contribution to V's effect, V*. The idea would be that P is a necessary but not a sufficient cause of V*, suggesting that P's instantiation alone was not quite not enough to bring

about V*'s instantiation, and that V was required as well to complete the causal task.

Baker herself makes it clear, however, that this is not how she interprets the contribution of P. She writes (MEL: 115-116), “the causal powers conferred by the constituted property-instance...are independent of the causal powers conferred by the constituter...[t]he contribution of Jones’s hand going up to Smith’s anger was exhausted by the fact that the hand’s going up constituted a vote against Smith”. Baker thus accepts that there is no *upward* causation of V* by P. My strategy in chapter 5 will be to go beyond what Baker is prepared to countenance and argue that there is no downward causation either – that V has no causal effects at the lower, nomological level, *qua* lower level. I mentioned above that Baker’s claim that all property-instances are physical is a barrier to adopting this position.

By analogy with an argument that Kim deploys in the context of ‘supervenient causation’ (see (ii) below), we can say that P’s constituting V guarantees the instantiation of V regardless of any causal relation involving P. *Ex hypothesi*, V is an independent and sufficient cause of V* – independent, that is, of P’s properties. P’s contribution to the causal powers of V is not through any exercise of P’s own independent causal powers, as a hand-raising. The independent powers of P include the ability to displace air molecules, or to block light waves, and, as we saw, to act as a necessary but insufficient cause of P*. P’s contribution to V, and thus indirectly to the instantiation of V*, is constitutional, not causal.

So if, as Baker claims and as our common sense understanding of causality dictates, V indeed causes V*, these causal powers of V must

not only be irreducible to but, as Baker says, “over and above” those nonderivatively possessed by P. What is meant by this is brought out by her claim, about material constitution (PB: 33; see **II**:(iii) above): “The identity of the constituting thing is submerged in the identity of what it constitutes. As long as x constitutes y , y encompasses or subsumes x ”. “As long as x constitutes y , x has no independent existence” (PB: 46). The exact meaning of these claims is admittedly obscure (Baker concedes they are metaphorical; MEL: 166), particularly as Baker also insists that, despite being subsumed, constituting entities have their own nonderivative properties, like the lower-level causal powers of P just mentioned. I think we can take the claims to mean that constituted entities, or at any rate those aspects of them designated by their primary kind properties (**II**:(i) above), have a kind of ontological priority, perhaps like the priority accorded to base properties in the supervenience relation, except that here the priority devolves upwards rather than downwards. So, just as on many accounts supervenient properties are taken to be *reducible* to their base, subvenient properties, the analogous idea would be that constituting entities are *subsumable* by the entities they constitute. The notion of ontological priority is also rather vague, but one possible gloss on it is that what is ontologically prior is what figures in what we should accept as the true *explanation* of an event or state of affairs.

The relations of causation and constitution on Baker’s account, then, are quite distinct, with different roles. As Baker spells out the difference, the constitutional relation between person and body “is a material one, of a kind that obtains throughout the material world: and the relation is noncausal” (MEL:178). Causation and constitution are independent, and in no sense in competition with each other - it may well indeed be the

case that everything that is caused is also constituted, and *vice versa*. Both causation and constitution are asymmetric relations among particulars, but while constitution is a synchronic ‘vertical’ relation among entities (objects, property-instances, events) at different ontological levels, causation is a ‘horizontal’ relation in which, at least in the ordinary, macroscopic world, causes temporally precede their effects. In contrast, the constitution relation is such that when a particular piece of marble is in favourable circumstances it just *is thereby* a statue; it does not become one after a time interval, no matter how short. Gillett (2012: 84), spells out other differences; relations of constitution (“composition” in Gillett’s usage) do not relate wholly distinct entities, while causal relations do, and, unlike causal relations (as I will argue in chapter 4), constitution relations do not involve the transfer of energy or mediation of force. To paraphrase a remark by Tyler Burge, “Constitution is a matter of how things hang together. Causation is a matter of how the things that hang together come about” (Burge 2007: 371, substituting ‘constitution’ for ‘supervenience’).

My suggestion, then, is that causation should be understood as only obtaining between or among entities (such as, typically, events) at the same ontological level. The constitution relation would seem to remove the need to postulate upward or downward causation. Baker’s attitude to the latter, however, does not consistently reflect this conclusion, as I will show in chapter 5:II:(i). On an account like Kim’s, principles like (4), the Causal Inheritance Principle, render the question of whether or not there is causation across levels comparatively unimportant, since a mental property-instance’s efficacy at the physical level is guaranteed by its physical realization. But the question has real significance for Baker’s account, on which higher level causes are claimed to be independent and

distinct. That causation may be a strictly intra-level relation is not something that Baker considers, but the idea that lower-level and ID causes have only lower-level and ID effects, respectively, will be a major theme of the discussion to follow.

This generalizes, I argue, to all cases of ID causation, and beyond to other non-ID, but relational, cases. When a bad investment causes the collapse of S&L, when Jane's willing to raise her arms causes her to raise them, or when adaptation to a regular change of pollinator causes *Scarlet Gilia* to change the colour of its petals each midsummer, then, on the constitution account, a distinct, independent causal relation is instantiated whose relata supervene widely on the microphysical properties of an extended space-time region, rather than on those of its local microstructure. This contrasts with, say, a case in which a cubical block of ice made up of eight identical smaller blocks is put on a scale, causing it to read 8 Kg. Here, the large block is constituted by the aggregate of the small blocks, and the causal relational property instantiated by the large block's making the scale read 8 Kg supervenes precisely and narrowly on the causal relational properties whereby each small block causes the scale to move by 1 Kg, that is, on the *intrinsic* properties, or local microstructure, of the constituter. So, because relational properties are not involved, the constitution and supervenience bases are co-extensive. In this case Baker's conditions for independent causation (IC) do not hold, in that all the causal powers of the constituted property-instance, such as its making the scale read 8 Kg, *are* conferred by its constituting property-instance(s) alone and are not multiply realizable.

(ii) *Kim and Supervenient Causation.*

Baker's account of independent ID causation bears some resemblance to a view once held by Kim, amongst others. In the 1970s Kim (see ESC, 1976) developed an account of *supervenient causation* that he claimed was applicable to macrocausation generally, including mental causation (ESC: 106-107). In that paper, Kim does not distinguish between properties and property-instances, but his meaning at each mention is clear from the context. In the mental case, the idea is that, for mental states [i.e., instantiations of mental properties] M1 and M2 and physical states [i.e., instantiations of physical properties] P1 and P2, M1 superveniently causes M2 if [property] M1 supervenes on [property] P1, [property] M2 supervenes on [property] P2, and [property-instance] P1 causes [property-instance] P2 (see Kim, PSNE: 63). As we saw, Baker rejects the idea of higher-level properties' supervening on those of their local microstructure that is assumed here, because of her insistence on the reality of relational properties, and on the need for *their* supervenience relations to be taken into account. Assuming that Kim envisages properties P1 and P2 as intrinsic to a person or her brain, comprising local microstructure, then on Baker's account the assumed supervenience relation between them and M1 and M2 does not hold. Rather, mental properties supervene on a much wider base of physical properties. As I will argue, however, the structure of Kim's concept of supervenient causation can be shown to fit a *constitutional* account of higher- and lower-level causation.

To take Kim's concept of supervenient causation at face value for the moment, in ESC Kim argued that while this form of causation is reducible to lower-level causation, it is nevertheless 'real' because it supervenes on the real, microphysical causal processes that form its

reduction base (*ibid*:107). Kim wrote (*ibid*), “It would be foolish to pretend that the proposed account accords to the mental the full causal potency we accord to fundamental physical processes.. [but] ..(m)ental causation does take place; it is only that it is epiphenomenal causation, that is, a causal relation that is reducible to, or explainable by, the causal processes taking place at a more basic physical level”. Furthermore, “supervenient epiphenomenal causation does not place the supervenient events at the level of the underlying causal processes to which it is reduced” (*ibid*), that is, higher-level causes are not envisaged as having lower-level effects. Kim claimed (e.g., SMK: 358) that the notion of supervenient causation applies independently of whatever particular metaphysical account is held of causation at the lower level.

Kim himself subsequently concluded that supervenient causation was not ‘real’, but “an empty verbal ploy ...a gimmick” that “neither adds new facts nor reveals any hitherto unnoticed relationships ...embracing this approach would lead us back to the overdetermination/exclusion problem” (PSNE:62). In rejecting supervenient causation, Kim refers instructively to what he calls ‘Edwards’s Dictum’, which he derives from the eighteenth century theologian-philosopher Jonathan Edwards:

There is a tension between “vertical” determination and “horizontal” causation. In fact, vertical determination excludes horizontal causation (PSNE: 36).

An everyday example of ‘vertical determination’ (*ibid*) is the synchronic determination of the macroproperties of a piece of bronze, such as its yellow colour, at time t , by its molecular make-up. We have two possible answers to why the piece of bronze is yellow at t ; (a) the synchronic,

mereological answer; its surface has a particular microstructure at t , and (b) the diachronic, causal answer; it is yellow at t because it was yellow an instant before, at $t - \Delta t$, and certain boundary conditions obtained during this period. Kim claims that the first explanation preempts or excludes the second. As long as the piece has the appropriate microstructure at t it will be yellow *regardless* of what obtained at $t - \Delta t$. Edwards himself argued that, just as successive images in a mirror are not causally related to each other, but are renewed at each instant by the mirror's reflection of new rays of light (*ibid*: 37-38), so God creates the world itself *ex nihilo* at each instant. Kim, however, applies Edwards' insights to the problem of mental causation.

Mind-body supervenience, or the idea that the mental is physically "realized" – in fact, any serious doctrine of mind-body dependence will do – plays the role of vertical determination or dependence, and mental causation, or any "higher-level" causation, is the horizontal causation at issue (*ibid*: 38).

The notion of continuous *ex nihilo* creation is replaced, on Kim's account, by microphysical causation, which ensures, *via* the operation of the vertical determination relation at each instant, that higher-level causation is *redundant*, an empty verbal ploy. In terms of Kim's example, above, instance P1's causing instance P2, together with property M2's supervenience on property P2, completely determines M2's instantiation.

My proposal will be that if we accept Baker's account of independent causation, this result can be avoided, since constitution relations share some, but crucially not all, of these features of vertical determination. As

I have already discussed, the supervenience bases of relational, particularly ID, properties do not track only their local microstructure, that is, the physical properties of their constituters. The relation of vertical dependence that Kim describes therefore does not hold between constituting and constituted properties. On the constitution account it is *not* the case that V^* would have been instantiated, through its vertical determination by P^* , regardless of whether V had been instantiated, for two related reasons. Firstly, the ID causal relation of V 's causing V^* is independent and is not merely a reflection, or redescription, of a lower-level causal relation between P and P^* , and secondly, the relations of vertical, constitutional, dependence connecting P with V and P^* with V^* are not necessary, like those connecting the macroproperties of a piece of bronze to its microstructure, but contingent, depending on the presence of favourable circumstances, as we have discussed.

CHAPTER 4. MANIFEST PHYSICAL AND INTENTION-DEPENDENT CAUSATION.

I. Causation and the Manifest Image.

(i) Introduction.

In this chapter I make several claims about the ordinary physical world - the world of Sellars's manifest image, the world that we, as possessors of second nature, inhabit, and the world that belongs in the space of reasons, as McDowell claims (chapter 1). I claim, (i) that this world (the manifest physical world) is real, (ii) that there are objective causal facts about the manifest physical world, one of which is that causes *produce* their effects in a sense to be explained, and (iii) that the manifest physical world is causally closed. If this picture of the manifest physical world is right, then the problem of causation by mental content, with the threat of epiphenomenalism, will be shown to belong firmly within the manifest image, and to be central to the very notion of the space of reasons. This conclusion is at odds both with Hornsby's argument (chapter 2), endorsed by McDowell, that the problem of mental causation arises from an inappropriate conflation of the standpoints of science and of common sense, and with Baker's view (chapter 3) that the problem of mental causation dissolves if we take the availability of causal explanation, and susceptibility to counterfactual analysis, to be the hallmarks of causality itself.

(ii) The Threat of Causal Eliminativism.

My conclusion in chapter 3:IV was that Baker's argument, aimed at countering Kim's attack on NRM and at establishing the reality and independence of ID causation, falls short of its aims. Baker claims that, in her airport example, Jane's willing to raise her arms to be searched, M, independently causes M*, Jane's raising her arms, and that this does not violate the causal closure principle. I argued that it is this causal relation's supervening on a wide base that guarantees non-violation of the principle, and that this same supervenience makes the claimed independence of the relation vulnerable to the Kimian anti-NRM argument. I also argued that Baker's constitution account of mental causation is subject to the same difficulty that besets token identity theories, in that on Baker's account semantic content, *qua* content, does not seem to be causally relevant.

However, Kim's anti-NRM argument itself depends on further assumptions. The Principle of Causal/Explanatory Exclusion states that there is no more than one complete and independent cause (or causal explanation) of any event, thus stipulating that mental causation does not involve widespread and systematic overdetermination, as would be the case if each effect of a mental cause also has a physical cause. If it could be shown that this constraint did not apply, Baker could still maintain that M independently causes and explains M* despite the existence of a causal explanation at the microphysical level that makes up the wide supervenience bases WSB and WSB* of M and M*.

Baker rejects the causal/explanatory exclusion principle (MEL: 101-102), exploiting the possibility that, at a sufficiently fundamental level, it will turn out that either the notion of cause has no application, or that there is an infinite downward regress of causal levels. The argument for

the principle assumes that, if there *was* one single complete and independent cause of an event, its locus would be at the lowest, most fundamental physical level. Doubts about the fundamental basis of causality take two related forms, both of which raise the possibility of causal eliminativism, which might be thought threatening especially to metaphysically substantial, ‘production’ accounts of causality such as Kim’s. Such a development, it is argued, would open the field for an ‘anything goes’ view of causality which, as we shall see, would suit Baker’s argument.

The first form of the threat to the fundamental basis of causality is raised by Schaffer (2008: 24-26) and others (see Price and Corry, eds, 2007), who draw attention to a tension between the place of causation in science and its role in everyday life, in that a number of philosophers from Bertrand Russell (1912) onwards have claimed that since fundamental physics does not require the notion of causality, philosophers should also eschew it. As for the second threat, Schaffer (2003) has also suggested that there may be an infinite regress of levels of causal dependence, and Block (2003), in a challenge to Kim, has called this the problem of “causal drainage”. If there is no fundamental, or bottom, causal level, Block argues, there is “endless subvenience”, and *nothing* has any causal powers, at least in the sense that Kim wants. Kim (PSNE: 57ff; CMC: 251) argues, along with others (e.g., Mumford and Anjum 2011:16-17) that the ‘no-bottom-level’ possibility does not rule out the appearance, or “emergence”, of what we know as causation in the world at some level higher than the fundamental. Ross, Ladyman and Spurrett (Ladyman and Ross, eds, 2010: 257-297) argue in detail that, despite the fact that, as they claim, fundamental physics has no use for causality, in special sciences like chemistry or sociology the identification of causal relations

is an indispensable heuristic device for the identification of the “real patterns” that constitute those sciences’ subject matter. I argue below (v) that causality is a real feature of the world of the manifest image. For the moment, though, I want to concentrate on how Baker exploits the no-bottom-level argument in support of her own views.

Baker (MEL:101-102) first argues that the causal exclusion principle, if true, would be a *necessary* truth. Therefore, she claims, the existence of the *mere possibility* that there is no bottom level is enough to refute the exclusion principle, since that would mean the principle *cannot* be a necessary truth, and thus, on Baker’s initial assumption, cannot be a truth at all. The argument is valid, but the premise that the principle would be a necessary truth is, I think, open to question. It certainly seems right that if there is *actually* no bottom level, the *prima facie* persuasiveness of the causal exclusion argument would be considerably undermined, since exclusion would then appear to have to continue *ad infinitum*. Baker, however, is basing her refutation on the argument that, even if there *is* a bottom level, it might (for all we know) have been the case that there was not. But whether or not there is a fundamental causal level would seem to be an empirical, or nomological, fact about our world - as Schaffer (2003: 502) says, it seems unlikely that its truth or falsity will be settled by philosophical argument. Suppose that physicists conclude that there *is* a fundamental causal level – perhaps, as Kim argues, the level above which something recognizable as causation emerges. In that case, it would seem dubious to deny the causal exclusion principle on the basis of something – the possibility that things might have turned out differently – which has been shown not to be the case. On the contrary, that finding would seem to provide us with good grounds for *accepting*

the exclusion principle, and this, I suggest, counts against Baker's claim that the principle would embody a necessary truth.

Having rejected the exclusion principle, Baker enlists its rejection in support of her claim, *contra* Kim, that not only mental causation but also ID causation at multiple levels, that is, what Baker calls commonsense causation (MEL: 97-120), is irreducibly real. The latter claim, however, is not primarily *based* upon rejection of the exclusion principle but, as we saw, on the indispensability, for us, of the notion of commonsense causation (3:I:(ii)) – that is, on what Baker sees as the catastrophic effect that its rejection would have on our everyday lives. Baker also takes the no-bottom-level possibility as a reason to embrace the possibility of overdetermination (MEL: 101); “if it turns out that there is no fundamental microphysical level, then we cannot deny overdetermination, lest all the causal powers drain away”. Finally, she takes these conclusions to support the rejection of a metaphysical, productive concept of causality in favour of her epistemological one, based on dependence or difference-making (MMC: 92-95).

However, there seems to be a tension between these second and last conclusions. What Baker seems to argue is, first, that commonsense causation is real because there may be no fundamental level to which it might be considered reducible. Putting it slightly differently, because there is no basic, *privileged* form or level of causation, there is no reason not to consider commonsense causation as ontologically robust as any other kind. As Schaffer (*op.cit.*: 513) puts it, on the no-bottom-level assumption, “Because there can be no privileged locus for the causal powers, and because they must be somewhere, they are everywhere”. But, secondly, because the no-bottom-level possibility rules out a

fundamental metaphysical characterization of causality, commonsense causality cannot be given a metaphysical characterization at all. Block's picture of an infinite regression of causes could give grounds for thinking that causation, far from being widely overdetermined, is not real at *any* level – that causal powers really do “drain away”. That is, rather than acting as support, *via* rejection of the exclusion principle, for Baker's independently-reached ontological position on commonsense causation, the no-bottom-level possibility could be argued to militate directly *against* that position. If causation is not real at any level, then it seems that all we are left with are our causal explanatory practices, and since Baker insists on the priority of these practices this might seem congenial to her position, but it is not. It is, as we saw in chapter 3:II, part of Baker's account that causality, including the causality of constituted objects and properties, is real and irreducible, despite its ontological dependence on explanatory practice. So we are brought back to Kim's position; in virtue of what are these causes real and not just reflections of our explanatory practice?

Jessica Wilson writes:

causal powers are grounded in specific fundamental forces or interactions...it is plausible, even obvious [that] the causal powers of being positively charged are grounded in the electromagnetic force; the causal powers of quantum colour properties are grounded in the strong force; the causal power of being able to sit on a chair without falling through it is grounded (at least) in the gravitational and electromagnetic forces; and so on (2005: 432).

Wilson is discussing what I have been calling lower-level causation, and her argument seems intuitively correct. On the no-bottom-level assumption, of course, the fundamental forces she mentions are not fundamental at all, but merely stages in an infinite regression. Adherents of Block's causal drainage argument may object that the claim of grounding here is illusory, since the grounding entities are not themselves grounded. So, would the absence of ultimate grounding for lower-level causation rule out the possibility of a metaphysically substantial, "thick", productive account of it? If it does, and if that means the only account available is a counterfactual dependence-based one, my attempt to metaphysically distinguish lower-level from ID causation would be threatened, since the distinction could reflect only our causal talk and not anything more substantial.

My concern, however, is not with causation as ultimately interpreted by fundamental physics, but as it features in our ordinary, manifest image of the world. What is important for my purposes is not how this causation relates to the fundamental, but how it is distinguished from ID causation. The possibility raised by Block, Baker, and causal eliminativists is that the theory of lower-level causation is not closed because it is, so to speak, open at the bottom, that is, at a sub-atomic level. I will argue, however, that causality in the manifest image closed at the *top* – that mental content cannot be counted amongst its relata. If I am right that the manifest image is the appropriate setting in which to address the problem of intentional causation, then questions about the nature of causality at fundamental levels, while obviously important for other reasons, are not crucial to my argument. I am not claiming that the distinction I am making between fundamental and 'ordinary' physical causation marks a sharp division in reality, but only agreeing with Kim (CMC: 232) that

above some organizational level of reality the physical causation with which we are familiar appears.

(iii) Causation and the Manifest Image.

My claims, then, are about causality as it concerns the physical entities – objects, substances, their properties and relations - that make up the ordinary, macroscopic world in which we live, together with some of its well-understood extensions into the microscopic - the world that corresponds to what Sellars (PSIM: 1-40) envisaged as corresponding to the manifest image of man in the world. I will take the relata of this kind of causation to be events, where an event is construed as an instantiation of a property in an object at a time (see 3:IV:(i)), although the exact nature of the relata is not central to the argument. I will call this *manifest physical* causation.

The first claim is that, no matter how problematic the notion of causation may be at a fundamental level, there exist objectively real causal relations among these observable physical entities, where ‘objectively’ is to be understood at least in Baker’s and McDowell’s ‘minimal’ sense (see 1:II:(iii)) of recognition-independence; that is, facts about causal relations within this nexus obtain (or do not) independently of the thought contents of any individual or community. The very idea of causation has its natural home in the manifest image, and it is plausible that these causal relations reflect factors that constrain the behaviour of these entities in ways that people have been familiar with, with varying degrees of understanding, since earliest times. No doubt this familiarity is largely a reflection of our having successfully evolved in a world with just these constraints. In more recent centuries much of the behaviour of

this physical world has been codified in laws backed up by scientific theory, and the extent to which it is understood, including an understanding of how it fits together into a unified whole, has greatly increased. My point, however, is that we have a deep and intuitive understanding of what is and is not possible within the causal nexus of the manifest world. We know, for example, that macroscopic objects cannot change their spatial location from *a* to *b* without passing through space between *a* and *b*, and that objects in general do not spontaneously appear or disappear. As de Muijnck puts it (2003: 46), if we cannot find any physical influences connecting alleged cause and effect, we would sooner suspect coincidence than ‘action at a distance’ – that is, than some kind of magical cause-like process. These constraints operate regardless of what may or may not characterize causality at a fundamental level. Probably the most profound basis for our deep and intuitive understanding of manifest physical causation, however, is that it is just the kind of causation that we experience ourselves as embedded in. Leaving aside for the moment the question of causation by content, it is clear that idea of manifest physical causation captures the nature of the action of our bodies upon the world, and *vice versa*.

Kim himself, as we saw (3:III:(i)), fully endorses commonsense causation involving everyday objects, while at the same time relating their causal powers to their microstructure. When a baseball breaks a window the shattering is caused by the ball and not by the individual particles composing it (Kim PSNE: 56). This is an example of what he calls micro-based macrocausation (MPW: 82ff). The baseball’s properties supervene mereologically on those of its parts, and its causal powers are therefore *determined* by the causal powers of the properties

and relations that figure in the construction of the ball's causal powers, but the latter are not *identical* with the former;

Micro-reductively explainable causal powers may be new causal powers, net additions to the causal structure of the world. None of this is in conflict with the basic commitments of physicalism. Physicalism need not be, and should not be, identified with micro-physicalism (Kim MPW: 117).

I explained in Chapter 3:III how Kim does not regard mind-body relations as examples of micro-based macrocausation, hence his anti-NRM stance. While his views on ordinary physical macrocausation are much in line with the claim I am making here, I will suggest below (II:(iv)) that it is the notion of *mechanism*, rather than mereological structure, that provides the most satisfactory account of causation involving objects such as baseballs.

My second claim is that there is more to causal relations amongst everyday, manifest physical entities than just their support of counterfactual statements. These relations are characterized, not just by spatio-temporal contiguity, but also by production or connection, or the presence and persistence of some kind of physical signal, such that causes actively *produce* their effects. This notion of physical connectedness is the basis of a number of theories of causation, such as that it is energy transference (Fair 1979), the manifestation of causal powers or dispositions (Harre and Madden 1975; Mumford 2008; Mumford and Anjum 2011), mark transmission (Salmon 1984), or persistence of a conserved quantity (Dowe 2009). I will not try to enlarge upon or compare these. It suffices that they all agree, as does

Kim (CMC), that physical connectedness is required. Following Hall (2004) I will call this the “production” view of causality, and the contrasting one the “dependence” view, in the discussion in **II** below.

A third claim is that, as one would expect on this production construal of causality, causal relations in the physical world of the manifest image are bound by the principle of physical causal closure; any (manifest) physical event that has a cause at *t* has a complete (manifest) physical cause at *t*. Most versions of this principle assume that it derives from physical theory (e.g., see Vicente 2002: 76) and take it to be about closure primarily at a hypothetical fundamental microphysical level. I am claiming, however, that the causal structure of the *manifest*-level physical world is closed to the non-physical and that our intuition that it is closed derives not, or not only, from scientific physics but from our observations of the everyday world. So the claim is that it is a recognition-independent fact that the manifest physical world is causally closed. Thus, when someone hammers a tack into a wall, the tack’s movement into the wall is caused exclusively by the force and direction of the impact of the hammer and the orientation of the tack, and not by his desire to hang a picture.

Thus, as I argued in chapter 3:**IV**:(*iii*), the principle rules out the physical causal efficacy, *qua* mental or ID, of mental or ID property-instances. So it would be redundant to add to the statement of causal closure the usual stipulation “assuming that there is no systematic overdetermination of effects by both mental (or ID) and physical causes”. The requirement of a physical signal or mechanism rules out, I will argue, the possibility that a relational property-instance, such as someone’s wanting to hang a picture, could contribute at all in the sense of being causally efficacious

for any physical event. So, as I claimed in 2:II:(v), our intuitions about the nature of physical causal relations clash with another deeply held intuition, that our beliefs, desires, etc, make a *direct* causal difference in the physical world.

(iv) Reality and the World of the Manifest Image.

Adopting this standpoint on causal relations among the middle-sized objects of the everyday world is obviously only possible if the *reality* of the everyday world is accepted. Baker and McDowell, of course, both strongly affirm its reality in their different ways, but it has been challenged by many, notably by Sellars with his distinction between the manifest and scientific images of man. In Sellars's formulation, the manifest image is a sophisticated framework that encompasses the space of reasons and much of the realm of law. It includes persons and their activities, but also "lower forms of life and 'merely material' things, like rivers and stones" (PSIM: 9) – familiar things whose behaviour we normally interpret in the realm of law way, as constrained only by natural (physical) law. As de Vries (2005:10) points out, the manifest image has been refined over time through reflection and what Sellars (PSIM: 7) calls correlational induction, so that it encompasses the findings of many kinds of scientific study of the *observable* natural world such as, for example, the Boyle-Charles law correlating changes in the pressure, temperature, and volume of gases.

What is excluded from the manifest image is the type of scientific reasoning that "involves the postulation of imperceptible entities, and principles pertaining to them, to explain the behaviour of perceptible things" (*ibid*). These postulations, and the theories surrounding them,

have led to the rise of the scientific image and have enabled it to constitute a rival alternative to the manifest image in a way that can threaten the very existence of most of the latter's contents. In EPM §41 Sellars famously pledges his allegiance to science and affirms the unreality of the commonsense world of physical objects. In PSIM he claims that these "manifest objects are 'appearances' to human minds of a reality which is constituted by systems of imperceptible particles" (PSIM: 26), rejecting two alternative views, first, that manifest objects do exist and are *identical*, as mereological sums, with the particles of the scientific image, and second, that only manifest objects *really* exist, and all the postulates of the scientific image are abstract representations of them. Kim's view, above, that manifest objects have causal powers that are mereologically supervenient upon those of their micro-constituents but are yet distinct, appears to fit somewhere between these two positions, and we may assume Sellars would reject Kim's position also. Sellars can thus be seen, in this context, as an early model for Baker's reductionist opponents or McDowell's bald naturalist. He does, however, concede the indispensability of what de Vries (2005: 275-279) calls the "practical reality" of persons as beings whose full characterization requires more than the kind of description of "what is the case" (PSIM: 40) that the scientific image can provide, while acknowledging that the latter image can also provide "a postulational image [of a person] in which he is a complex of physical particles" (*ibid*: 29). As de Vries says (*ibid*: 277), "persons *per se* simply do not show up in the pictures produced by science". Thus while much of what is contained in the manifest image is fated to be superseded by the scientific, the "conceptual framework of persons" is, rather, something that will need to be *joined* to the scientific image (PSIM: 40) in the final synoptic picture that will reconcile the two images.

And yet persons are, as Sellars agrees, the “primordial category” (de Vries, *ibid*: 276) of the manifest image, without whom there would be no images of any kind. One might well wonder how a final picture could contain persons but not the world with which they interact (or at least with which they unavoidably take themselves to be interacting), which is the world of the manifest image. De Vries (2012) points out that Sellars’s talk of “joining” the conceptual framework of persons to a complete scientific image suggests that the latter image could somehow exist apart from that framework. But of course “any conceptual framework within which a scientific image of the world is possible must *already* be a conceptual framework containing persons and the language of individual and community intentions...science is primarily a human activity” (*ibid*: 12,13). Sellars acknowledges a distinction between methodological and ontological aspects of the scientific image, and that the scientific is “*methodologically* dependent on the world of sophisticated common sense” (PSIM: 20), while insisting that *ontologically* the scientific supersedes the manifest. Van Brakel (1996), however, argues that the dependence of the scientific image on the manifest goes much further than the purely methodological. Sellars’s “*scientia mensura*” – his claim that science provides the ultimate criterion for the existence of any entity (EPM: §41) - is itself made from the standpoint of the manifest image, van Brakel points out (*ibid*: 149), as is the claim that the images are ‘rivals’. Not just the methodology, but also all the values surrounding science – its goals as well as the criteria for judging its success – are grounded in the manifest. Van Brakel concludes that, “It is not a matter of whether we should grant priority to one or the other. There is *no choice* but to start from the world of daily life” (*ibid*: 154). As Danielle Macbeth (2010: 183) puts the same point,

“the perspective of the scientist, the view from nowhere, would not be a view at all were it not grounded in the view from here as its evidential base” (see also Rosenberg 2009). Sellars himself allows that, even though the manifest image is “but one way in which reality appears to the human mind” (PSIM: 15) (a way, nonetheless, that must in *some* way reflect “the intelligible structure of the world” (*ibid*)), there are objective – that is, intersubjectively verifiable - *truths* about contents of the manifest image.

Bill Brewer (2011) presents one form of an argument for the reality of familiar macroscopic objects, as opposed to those of “fundamental scientific-physics”, as the explanantia of our perceptual experiences. “Commonsense” physical explanations, he writes, are “robust”, in that they “maximize modal correlation with the perceptual experiences they explain” (Brewer 2011: 78; see also van Gulick 1993: 247). He presents a well-known example from Putnam, that the best explanation of why a square peg will pass through an appropriately sized square hole but not a round one of the same area is given by the peg’s size and shape. This remains so for pegs of very different microphysical structures, and, as Brewer puts it (*ibid*), what unifies respects in which those microstructures might differ in such a way as to alter the peg’s ability to pass through the hole is that the respects are precisely those that would alter its commonsense explanatory properties, that is, its size and shape. Brewer concludes by rejecting the view that “the best explanation of everything going on in the physical world is ultimately to be given in terms of fundamental scientific-physics” (*ibid*: 79). Thus a complete descriptive ontology must include the familiar macroscopic objects.

However, I think the strongest and most cogent argument for the reality of the world of the manifest image comes from McDowell's 'transcendental empiricism' (1:II:(i); 3:II:(v); RN: 6), according to which it is a *condition of possibility* of our having conceptual capacities and of our being entitled to the notion of an objective world at all, that we engage perceptually with that world,

it must be intelligible that conceptual activity has a *subject matter*...the very idea of a conceptual repertoire provides for conceptual states or episodes in which a subject matter for conceptual activity is sensibly present" (WL: 37).

But this subject matter is made up of the persisting objects of the world of the manifest image, with their properties and relations, not the postulated entities of the world of fundamental physics. Our perceptual experiences, as McDowell insists, are irreducibly conceptually structured, and thus take the form of logically structured intuitions, in McDowell's Kantian sense (see WL: 23-43), the components of which are objects, figuring as *de re Sinne*, and their properties and relations. In his Woodbridge Lectures (WL: 15-43), McDowell attributes Sellars's insistence that perceptual experience requires a non-conceptual sensory manifest to his belief that the ordinary objects of perception do not exist, so that what appear to be our perceptions of them must *really* be of sense impressions that we somehow construct from 'ultimate reality', which, for Sellars, means not Kantian things-in-themselves but the entities postulated in the scientific image. But it is, of course, central to the McDowellian account that there are no such non-conceptual items (see chapter 1:II), and that our experiences are, instead, instances of direct openness to the world. On McDowell's account, then, the reality of the

world of the manifest image is a condition of the possibility of our having conceptual capacities at all. I have already (3:II:(v)) mentioned how this aspect of McDowell's account provides strong support for Baker's claims of the reality of constituted objects. Baker herself does not acknowledge the scientific/manifest distinction – she moves freely between manifest and scientific (“microphysical”) characterizations in, for example, her causal arguments ((ii) above). Although the concepts of supervenience and realization are usually assumed, perhaps tacitly, to belong to the scientific image – in the literature higher properties are often taken to be supervenient on and/or realized by those of ‘microphysical’ entities - I see no reason why these terms shouldn't be adapted to apply within the manifest image; mental properties could be seen as supervening on and/or being realized by a combination of observable environmental/historical properties and neural properties, for example.

Acceptance of these arguments, however, certainly does not commit us to the view that the fundamental entities of the scientific image do *not* exist. What it seems to lead to is a liberal view of what constitutes an existence claim, such that, perhaps, existence claims, rather than being absolute, are relative to a framework (see Sacks 1989; Thomasson 2007: 110-129). It is undeniable that the objects of the manifest image, with their properties and relations, exist in *our* framework, in the world as it is for us (the ‘lifeworld’), as opposed to, say, the world as it might appear in the perspective-neutral image that science strives to attain. And, as Sellars agrees, there are observer-independent facts about the manifest image - including, I am suggesting, causal facts.

(v) *Uses of “Cause”.*

Our analysis of Baker's defence against Kim's rejection of mental and other higher-level causation in Chapter 3: **IV** shows, I have claimed, that her arguments do not succeed in establishing the independent reality of ID causation. Baker takes our pre-theoretical causal intuitions, such as that Jones' voting caused Smith's anger, as premises in her arguments, while the validity of those intuitions is precisely what Kim questions. As a prelude to suggesting a solution to the problem, I now want to look at what I think lies behind this clash of intuition with metaphysics.

There are, Putnam claims (quoting John Haldane), "as many kinds of cause as there are senses of 'because'" (Putnam 1999: 149). That is, the term 'cause' and the plethora of causal verbs like 'break', 'make', 'drop', etc, that stand in for it have many different uses in our ordinary discourse, uses which are intimately bound up with our explanatory practices and are linked by the use of the explanatory conjunction 'because' and its cognates. In particular, in our causal explanatory practice we routinely use agential or other ID causal descriptions alongside, and mixed with, norm-free causal descriptions of the kind that apply, in McDowell's by now familiar phrase (FAM: 328) to "the way things generally tend to happen". That is, our everyday discourse is shot through with both upward and downward causal descriptions, from ID to lower-level and *vice versa*, such as:

He purposely threw the ball that smashed the window.

The earthquake caused economic upheaval.

The gardener's failure to water the plants caused their death.

Human exploitation of nature causes climate change.

In conversation we usually do not distinguish either between causation and causal explanation (Beebe 2004: 293), or among events, states, objects, facts, or negative facts as causal relata. But when we unpack our causal discourse, I will argue, we can distinguish two meanings, or uses, of ‘cause’, neither of which is dispensable, but only the second of which denotes real manifest physical causation.

The *first* meaning is the everyday, pre-theoretical, generic one depicted in examples like the ones above and many more, in which the notion of cause generally depends on, or is secondary to, that of causal explanation. The above are cases in which ID property-instances, or events, cause lower-level property-instances, or *vice versa*, but of course there are countless causal statements and explanations in which both of these relata are ID:

Excessive sub-prime mortgage lending caused the recession.

The division of powers built into the American constitution makes it difficult for a minority party President to introduce new legislation.

As declarer in contract bridge, Mary’s perception of the situation caused her to realize that she should lose an early spade trick while she still had first round control in the other suits, in order to make the rest of dummy’s spades.

Jones' voting against him made Smith angry,

The arrival of a different pollinator in midsummer causes Scarlet Gilia to change the colour of its petals.

and so on (the last example involves biofunctional, rather than ID, causation, but is nevertheless not straightforwardly physical (see 2:III:(ii))). Precisely because the use of 'cause' in statements such as these does not imply a commitment to any metaphysically rigorous meaning of the term, the causal relata in them can switch freely among events, facts, states, etc, and causal explanations like these routinely and unproblematically include both negative causes, or things that happened because something did *not* happen, and preventions, or things that did not happen because something else intervened, as well as more complex cases like overdeterminations and pre-emptions. Demonstration of counterfactual dependency is often enough to satisfy us that this kind of causation is occurring. Most of the examples that Baker uses in support of her claims on behalf of independent higher level causation, such as a bad debt's causing the collapse of a financial institution (3:III:(ii)) are like this, but I reject Baker's assumption that *all* causal statements are in this category.

These uses of 'cause' are part and parcel of our normal explanatory practice just as much as are positive causal claims. A common feature of these causal statements is that they furnish answers to questions about

why the world is as it is in a certain respect...which appeal either to how it came about that it is as it is in the first place or to what maintains it as it is (Steward 1997: 183).

They typically do so by subsuming the individual instance under a *ceteris paribus* generalization appropriate to some ID domain, which might be technical, or scientific, such as psychological or economic, or cultural, for example from sport, the law, and so on. When causal statements or claims are used as explanations in this way, there is a general assumption that the asker of the ‘why’ question will understand the generalization and hence the explanation. There are, however, a multitude of explanations that are typically formulated using the sentential connective ‘because’ but are *not* causal in this sense. As Eric Marcus (2012: 160ff) points out, these explanations pick out dependence relations that are constitutive (though not (necessarily) in Baker’s technical sense) rather than causal. To use an example of Marcus’s, although we might say that a child is an orphan because she has no living parents, the latter is not the cause of her being an orphan but rather that *in virtue of* which she is an orphan; the *cause* is that which explains the death of her parents (*ibid*: 161). ‘Because’, then, has an even wider spectrum of ordinary use than ‘cause’, covering both causal and non-causal dependencies.

The first meaning of ‘cause’ here, then, reflects universal and multifaceted practice within the manifest image, and this meaning is, of course, perfectly legitimate and indispensable. However, it does not ordinarily attribute to the concept anything more than that causation is a dependency relation which, through its general regularity, supports causal explanations. These regularities may in some cases be strong enough to be called laws, but seldom if ever in a strict sense. This use of ‘cause’ is essentially tied to our epistemic interests; “[i]t is convenient to represent the world in terms of cause and effect because a causal

representation, if correct, enables us to make successful causal inferences” (Williamson 2007:107). When we are using ‘cause’ in this way, the associated sense of ‘explanation’ is not the objective one of Kim’s explanatory realism, but an epistemic sense – the explanation is simply that which provides epistemic satisfaction for the ‘why?’ questioner, who in the successful case is enabled to mesh the explanation with his existing set of concepts. In Chapter 5 I will argue that true ID causal statements describe genuine causal processes that are *constituted*, in favourable circumstances, by lower-level causal relations.

Williamson (*ibid*) adds, “Humans think in terms of cause and effect because of this convenience, not because there is something physical corresponding to cause which humans experience”. I argue, however, that there is a *second* sense in which ‘cause’ is used, one that presupposes that there *is* “something physical corresponding to cause”. This is the sense I am calling manifest physical. In everyday discourse it routinely occurs alongside, and seamlessly mixed in with, the first sense, and again there is no suggestion that there is anything inappropriate about this. This second sense of ‘cause’ is the one I described in (ii) above, which applies solely to explanations in terms of what we have been calling lower-level, norm-free causation. To repeat, I do *not* mean here a notion of causation at a hypothetical fundamental level at which anything recognizable as answering to our concept of causation may not be apparent. Rather, I mean an objectively real relation of lower-level, or physical, causation that in some way *produces*, or generates, effects in the inanimate and biological world around us (and within our bodies), is characterizable in a non-normative, context-free way, and whose relata are generally taken to be events rather than, say, facts. Thus, unlike the first sense, this sense of cause is not only epistemic or explanatory but is

also made true or false by the manifest physical world. This kind of causation also originates in the manifest image, but as I mentioned above, since the scientific revolution, the range of phenomena included in this image has grown far beyond what is immediately apparent to our unaided observational powers.

Van Brakel (1996:148) writes; “The scientific image is concerned with things like neurons, DNA, quarks, and the Schrödinger equation, again including sophisticated reflection, and a promise of more to come”. But, I suggest, scientific progress has been such that, rather than being superseded, the scope of the manifest image has expanded. It is surprising that van Brakel, a philosopher of chemistry and a leading opponent of Sellars’s prioritization of the scientific image (see van Brakel 2000), includes the first two items on his list as belonging to the latter. The existence and nature of neurons, as well as such things as the arrangement of the elements in the periodic table, the hyperosmotic filtering mechanism of the mammalian kidney, or quasars, are now well enough established, I think, that we no longer need to call them ‘postulates’, and the causal relations amongst them clearly belong in the same all-embracing nexus as those of everyday, non-ID nature. As pointed out earlier (3:IV:(ii)), neurons, like many other kinds of cell, can be made visible through an ordinary light microscope using simple staining techniques, and they and their functions have been studied, manipulated, and catalogued in exhaustive detail. So while they were not part of the manifest image of, say, the seventeenth century, there is a strong case for saying they are part of it in the twenty-first. Their mere imperceptibility to the naked eye is no reason to classify them along with subatomic entities such as quarks – small though they are, they are much closer to the human scale of magnitude than they are to the subatomic. It

might indeed be the case that most, if not all, of the subject matter of the special sciences – those other than fundamental physics – belongs in the manifest image. Statements such as

A lightning strike caused the forest fire.

Heavy rain caused the river to overflow.

Local electrical depolarization of the axonal membrane causes opening of voltage-gated sodium ion channels.

The presence of a low local pH causes enhanced release of oxygen from oxyhaemoglobin

are all examples of manifest physical causal claims. We should note that these manifest physical causal *statements* also support counterfactuals (if there had not been heavy rain, the river would not have overflowed, etc), but I argue below (section II) that, in addition, these counterfactuals are made true by causation in the sense of production. Mere counterfactual dependency is not sufficient to establish manifest physical causation.

The ubiquity of the first sense of ‘cause’ in ordinary language tends to obscure the fact that uses of ‘cause’ in the manifest physical sense occur embedded within it. As the examples show, causal descriptions and explanations involving the different uses usually share the same grammatical form. In addition, we often attribute the same kinds of effects to agents’ purposeful action and the forces of inanimate nature - the child or the wind can blow out the candles, the careless camper or the lightning strike can start the forest fire. We experience even in our own

bodies the effects of both physical forces and our own volition, often at the same time. Nevertheless, I argue below that we can isolate manifest physical causation as a distinct category.

(vi) Manifest Physical Causation.

In our ordinary causal discourse, then, cause and effect statements in which both relata are ID, in which one is ID and the other is lower-level, and in which both are lower-level, are intermingled, reflecting the first, naïve, meaning of ‘cause’. But this use disguises the fact that, as I claim, the manifest physical causal statements that are embedded in ordinary language, in which both, or all, causal relata are lower-level property-instances, have important features that are not shared by the other kinds of causal statements.

Part of what makes manifest physical causal claims distinct is captured by the principles of physical causal closure and causal exclusion. Together, these principles stipulate not only that every physical event (property-instancings in an object at a time) that has a cause has a physical cause, but also that no non-physical, such as mental or ID, cause is efficacious in the physical domain. As Kim puts it:

If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the nonphysical (MPW: 40).

Davidson reaches a similar conclusion in a different way, making the contrast between the mental and physical in terms, not of causes, but of laws, or rather, statements of laws:

Nomological statements bring together predicates that we know a priori are made for each other...mental and physical predicates are not made for one another (ME: 218).

As we saw in the discussion of anomalous monism (2:II:(iii)), Davidson goes on to show how physical predicate statements have an internal consistency that suggests that they form part of a single comprehensive system. Mental predicates, he argues, are not part of this system. We know this, he claims, not from empirical evidence, but *a priori*, from the irreducibly anomalous nature of propositional attitude attributions and the criteria we use to attribute them:

Beliefs and desires issue in behaviour only as modified and mediated by further beliefs and desires, attitudes and attendings, without limit (*ibid*: 217).

Davidson, of course, exploits the notion of the extensionality of causality and the theory of token identity of mental and physical events to build his account of mental causation according to which mental causal efficacy is ensured *via* this identity. For him, mental predicates are only predicates, not attributions of properties. Kim's approach (e.g., CMC: 261-262), in contrast, involves the functional reduction of mental properties. In chapter 2 I discussed reasons for rejecting token identity, and in Chapter 3:V:(ii) I suggested that Kim's conclusion that mental events must be reducible if they are not epiphenomenal can be avoided

by developing Baker's constitution account. But I am with Kim in accepting physical causal closure, or the causal closure and completeness of the manifest physical causal nexus, and my argument for the distinctness of ID causation depends on establishing that manifest physical causation is a real feature of the world of the manifest image.

The first way, then, in which manifest physical causal relations are distinct is that, despite the looseness of our ordinary causal discourse, these relations are causally closed. This closedness is closely related to their being the kind of causes that are *productive* of their effects, in the sense I discuss in **II** below. A second way they are distinct, I claim, is that they are uniquely related to reality. It is not just that causal claims that relate only manifest physical property-instances are *objective* in the sense, shared by McDowell and Baker, of being recognition-independently true or false. ID causal statements like 'excessive sub-prime mortgage lending caused the recession' are also objective in this sense – it could be that, despite the opinion of most or even all economic experts, the statement is false.

The second distinguishing feature of manifest physical causal relations, rather, is just the one that McDowell identifies as the distinguishing feature of realm of law explanations (FAM: 328) - that they have complete descriptions that do not allude to normativity or related properties that are connected with our interests, such as meaningfulness or goal-directedness. As explained in chapter 1:**II**:(iv) in the discussion of McDowell's arguments in TSN, he regards realm of law facts as those we obtain by adopting the meaning-free, 'view from nowhere' perspective of science. The kind of causation we are discussing clearly has a central position in this mode of explanation and intelligibility,

especially in the special sciences. Put thus, this is an epistemic rather than a metaphysical point. But we have seen that for McDowell it is a transcendental necessity that the structures and relations that are characteristic of the realm of law mode of intelligibility are reflected in the make up of the empirical world itself. Both the events that comprise the relata of manifest physical causal events, and the relation between them, then, are devoid of normative or meaningful features, which is why we can imagine their occurrence in a situation in which bearers of intentionality had never existed. This does not, however, mean that *explanations* of manifest physical causal events need necessarily be framed in an interest-free way.

Manifest physical causal explanation and McDowell's realm of law mode of explanation are not quite co-extensive, since McDowell envisages the realm of law as encompassing all non-normatively explainable phenomena, and this includes those of fundamental physics. But all manifest physical causal explanations are of the realm of law kind. The precise quantitative relations and qualitative structure characterizing individual instances of manifest physical causation can be elaborated in terms of the regularities that we call laws of nature – mechanical, electromagnetic, and so on, even though at the manifest rather than fundamental level the laws may be rather less than strict in Davidson's sense, and the descriptions are to that extent less than complete. So we can say that something like the PNCC constrains this form of causality. But unlike Davidson's, my account does not claim that *all* causal statements are redescriptions of causal statements that in principle can be formulated in terms of laws. In Davidson's case his monism requires that mental events have physical descriptions, under which their interactions with other events are law-governed. On the

account I am developing, rather than identity, it is the more complex, but less strict, relation of constitution that links the mental and physical.

In chapter 2:II:(i) I discussed how Hornsby (and McDowell, who follows her thinking in this area) claims that worries about mental causation result from a misplaced adoption of the scientific standpoint in an area – mental causation and “intuitive physics” – which should be viewed from the standpoint of commonsense psychology. I argued to the contrary that the problem of mental causation, or more generally the ‘mind-body problem’, arises from *within* commonsense psychology, once a certain minimal level of sophisticated thinking has been attained. In the terms I am using here, the problem is one that arises within the manifest image, and it stems from the features, notably objectivity, physical causal closure, and production, that I have attributed to manifest physical causation. So I reiterate that it is mistaken to claim that it is a clash of standpoints that has led to the appearance of a problem.

Nevertheless, I believe that the picture presented in rather different ways by Hornsby, Baker, and McDowell is correct. Our actions and perceptions are indeed directly world-involving, and, as McDowell says (MW: 27), no ontological gap separates mind from world. But in turning aside from, rather than confronting, the metaphysical points that Kim and others raise, I think these authors miss an opportunity to maximize the persuasive virtues of their view. It is a common theme in the writings of Hornsby and McDowell that the first-personal, rational and normative way of experiencing and acting in the world that Hornsby (SMH) calls naïve naturalism and McDowell second nature are not even discernible from an impersonal standpoint, such as either the fully scientific or the ordinary impersonal standpoint that reveals the manifest physical world.

But, once again, the opposite is clearly *not* the case. On the contrary, it is our openness to the space of reasons that enables us to adopt these other standpoints. Natural science in particular is, one might say, a paradigmatic space of reasons activity (see 1:I:(ii)). And everything that the adoption of the scientific standpoint (leaving fundamental physics aside) has achieved over the centuries points to the conclusion that that world is causally self-sufficient, reinforcing what our encounters with the everyday, manifest physical world tells us.

So, I conclude, the picture of the physical world that gives rise to Kim's and others' metaphysical claims originates from within the space of reasons itself. I believe the problems arising from the causal closure of the physical and similar intuitions are not pseudo-problems that can be sidelined – no account of the place of mind in the world can be complete until it has addressed them. McDowell is surely right that the scientific is only one approach to the real, but it is so spectacularly successful within its own sphere that philosophical positions tend to be regarded as suspect to the extent that they do not take science's view of the world into account. I would maintain, then, that it is incumbent on upholders of views such as Hornsby's and McDowell's to provide an account of the place of both science and the manifest physical within their world picture. As well as an account of causation by content, we need an explanation of how the space of reasons can so much as exist in a world in which physical causal relations otherwise seem so all-pervasive. I hope to show how the resources of the constitution account might point us in the direction of a solution. But first, my claim that manifest physical causation is distinct and ubiquitous needs further defence.

II. Manifest Physical Causation as Production.

(i) Production versus Dependence Accounts.

In I:(iii) and (iv) above I distinguished two forms of our *use* of causal (and causal-explanatory) terms – a broad category of use in which causal statements express dependency, as shown by their support of counterfactuals, and a narrower category contained within the broad one which, in addition to sharing these features, is an expression of manifest physical, recognition-independent, productive causation.

I have been analyzing the disagreement between Baker and Kim mainly as one over causal exclusion and overdetermination. But, as I have mentioned several times, these differences have a deeper origin in opposing views of the nature of causality itself, which centre around the difference between production and dependence accounts. Baker's construal of causality as dependence, characterized by regularity, nomologicality, and support of counterfactuals, is essential to her claims for independent ID causation. But counterfactual analysis, for example, is just as applicable to ID causation as to lower-level (manifest physical) (see Marcus 2012: 178-179), and this is unsurprising since, as Kim points out (CMC: 254-255), counterfactual dependence and regularity are part of our broad concept of causality. Counterfactual analysis, an adherent of the productive account can say, clarifies the structure of causal relations in a metaphysically neutral way, but does not itself address the question of what grounds the counterfactuals. On this view the obtaining of counterfactual relations is a consequence of the obtaining of causal relations in some deeper sense, rather than being definitive of them. So, if there is a distinction between ID and manifest

physical causation, construing causality in terms of counterfactual analysis will fail to mark it. But, as I have been arguing, on a production construal ID and manifest physical causation emerge as quite distinct.

Wim de Muijnck (2003) and Ned Hall (2004) acknowledge the differences between the dependence and production accounts and believe that, as de Muijnck says (*ibid*: 41-50), they mark an unavoidable duality in our concept of causality. Each of these authors independently claims that we need both concepts because there are some imaginable causal scenarios, such as pre-emptions, which resist analysis in terms of counterfactuals, and others, notably ones that include causation by omission, that resist analysis in terms of production; thus neither can provide a univocal account.

There is an enormous literature on various rival analyses of causality. Here I will concentrate only on features of the production and dependence views and will try to show that, despite Hall's and de Muijnck's insistence that we cannot do without the dependence view, manifest physical causation can be analyzed entirely in terms of production.

Hall claims that five theses are true of causation (Hall 2004: 225-226):

Transitivity: If event c is a cause of d , and d is a cause of e , then c is a cause of e .

Locality. Causes are connected to their effects via spatiotemporally continuous sequences of causal intermediaries.

Intrinsicness. The causal structure of a process is determined by its intrinsic, non-causal character (together with the laws).

Dependence. Counterfactual dependence between wholly distinct events is sufficient for causation.

Omissions. Omissions – failures of events to occur – can both cause and be caused.

Hall then proceeds to show, with a series of examples and ‘neuron’ diagrams, that the five theses cannot all be true. In particular, if *Transitivity* and/or *Locality* and/or *Intrinsicness* are true, then *Dependence* and/or *Omissions* must be false, and *vice versa*. Hall concludes (*ibid*: 253) that there are two senses of cause, the first three theses on the list being characteristic of the first sense, production, and the other two being characteristic of causation in the sense of dependence.

As my purpose is to show the full adequacy of the production account for manifest physical causation, I’ll comment only briefly on the generally agreed difficulties of counterfactual analysis in accounting for pre-emptive cases, using de Muijnck’s example (*op.cit*: 66-68). A man sets out across the desert. One enemy has poisoned his reserve drinking water, but another enemy, not knowing this, pierces the bottom of the can. The water leaks out and the traveler dies of thirst. It is widely recognized among counterfactual theorists that the standard counterfactual analysis, which bases the claim that *c* causes *e* on the fact that if *c* had not happened, *e* would not have happened, runs into difficulties when faced by this kind of scenario. It fails here because of

the back-up poisoning, and so will not deliver the result that it was the piercing of the can that caused the death. To get this result, de Muijnck claims, we have to step out of the counterfactualist framework and resort to our intuitions “shaped by the structure of the physical interactions among puncturing instrument, can, water, poison, and man” (*ibid*: 67-68) – that is, to our intuition that enemy number two *produced* the effect by puncturing the can.

Kim (CMC: 254-260) also supports the indispensability of the productive view of causation, especially with respect to *agent* causation, because,

“agency requires the productive/generative conception of causation... mere counterfactual dependence is not enough to sustain the causal relation involved in our idea of acting upon the normal course of events... These causal processes all involve *real connectedness* between cause and effect, and the connection is constituted by phenomena such as energy flow and momentum transfer, an actual movement of some (conserved) physical quantity” (*ibid*: 256-257).

This picture of the production account ties it firmly to Hall’s notions of *Locality* and *Intrinsicness*. As de Muijnck points out (*op.cit*: 47), to deny spatiotemporal contiguity as a feature of causes and effects would be to allow everything to be causally related to everything else – “the very concept of causation becomes hollow”. Here we have, in a nutshell, the explanation of why Baker believes there are no metaphysical barriers to the claim that, say, Jane’s willing, M, causes Jane’s arm-raising, M*, and Kim thinks there are. As a counterfactual theorist, all Baker needs in

order to establish this as a *bona fide* causal relation is that it be true that if M had not occurred, neither would M* (see 4:IV:(iii)). I agree with Kim that something more is required, though not (as I will argue below) with his further claim that the causal relationship reduces to one between the neural/physical realizers of such mental events (CMC: 260-262; PSNE: chapter 4).

(ii) *The Problem of Causation by Omission.*

The biggest barrier to acceptance of the productive account has been the problems of causation by omission (or disconnection) and causation of omission (or prevention) (see Dowe 2000: 123-145; Schaffer 2000; de Muijnck 2003: 80-82; Hall 2004: 248-252; Beebe 2004; Armstrong 2004; Kim CMC: 258-260). Schaffer (2000: 289) argues that “causation by disconnection is causation full force”. Its existence, he claims, refutes the programme of “address[ing] the Humean problematic by seeking a connection in the objects via physical processes such as energy flows”. The production approach cannot accommodate causation by disconnection, he claims, since the latter “involves no persistence line between disconnector and effect, but rather the severing of one” (*ibid*: 291).

Schaffer’s example, however, turns out to be instructive (*ibid*: 286-288). An assassin shoots a victim through the heart. Causation by disconnection is involved in this event, Schaffer argues, in three ways: (a) when the assassin contracts his trigger finger, nerve signals cause muscle contraction by *preventing* tropomyosin from blocking the actin-myosin binding that constitutes the contraction; (b) pulling the trigger *disconnects* the sear, releasing the spring so that the striker hits the

cartridge; (c) the heart-piercing causes death by *preventing* oxygen from reaching the brain.

I think that, despite the apparent force of this argument, it can be shown that manifest physical causation as I have characterized it does not require omission or prevention. In Schaffer's argument, (a), (b), and (c) are all contextual, interest-bound descriptions of manifest physical causal processes in the form of mechanisms (although (a) is an action, it is redescribed in intention-free terms in Schaffer's analysis), framed in ways that make them *explanatory* of their role in the wider context of the assassin's killing his victim. This is most obvious in (c); if we analyze the process, not in the context of an assassination, but at a lower, or simpler, level of description – if, that is, we bracket our natural tendency to think of the life-death contrast as the all-important explanandum - we can describe the process without alluding to disconnections or omissions at all. We do not have to go anything like as deep as a fundamental microphysical level to do this. In the victim's brain, at the intracellular level, metabolic mechanisms driven by high-energy phosphate bonds that depend on oxidative phosphorylation change in a multitude of ways in the altered environment. Sodium and chloride ions flow in through the cell membrane and potassium ions out, altering its electrochemical properties. The cell switches briefly from aerobic to anaerobic (lactate-based) metabolism, then to a state close to equilibrium. These last descriptions could themselves be broken down further, but the point is that we can quite easily identify a level of description that is entirely in terms of productive, manifest physical causation. What from our standpoint is a catastrophic loss of normal function is, from the norm-free 'standpoint' of the manifest physical, simply an instancing of productive causation. In (a) also, one can redescribe the mechanism in

terms of what *does* happen, rather than what is prevented from happening. The nerve impulse releases calcium ions that bind to troponin, forming a complex that alters the binding characteristics of the tropomyosin in such a way that the potential energy stored in myosin is converted to kinetic energy. And one could show that something analogous happens in (b).

Two other examples, due to Dowe (2000: 123ff), are

- (d) the father's grabbing the child prevented the accident (prevention)
- (e) the father's inattention was the cause of the child's accident (omission).

Dowe treats these statements just as they are presented, as challenges to his own version of the productive account of causality. His account (*ibid*: 132-133), which I will not enlarge upon, involves analyzing (d), for example, "not as genuine causation but as a counterfactual truth about genuine causation", such that the father causally interacted with a causal process (an approaching car) that *would have* caused the child's accident (genuine causation) had the father not acted, and there is a comparable analysis of (e). But if one redescribes what happened in (d), say, in the more fine-grained terms of basic manifest physical processes and without reference to our interests, one gets an admittedly very long-winded account involving physical descriptions of the movements of the father and the child on one hand, and the car on the other. Since there was no contact between car and child, there is at this level of description no interaction between these two sets of events. The account need not

refer to the prevention of anything and counterfactual statements need not appear in it at all.

I conclude that, at least in cases of manifest physical causation, omissions and disconnections feature in causal descriptions like Schaffer's and Dowe's examples for the excellent reason that they contribute greatly to the description's explanatory power at an appropriate level of interest, but that their use is part of a different, higher explanatory pattern than we need to use for a purely manifest physical causal explanation. As Helen Beebe (2004) argues, references to omissions are perfectly legitimate in causal *explanations*, but not in strictly causal statements. If the relata of causation are events, she argues (*ibid.*: 291), there cannot be causation by absence, since it is a plausible assumption that "there are no events whose essence is the absence of a property or particular".

Beebe also argues that if a case of apparent causation by omission such as

the gardener's failure to water the plants caused their death,

is taken to be a description of a genuine causal relation, on the basis that it supports counterfactuals like 'if the gardener had watered the plants they would not have died', we end up with too many causes, since the plants' death was, arguably, caused by *everyone's* failure to water them. I will not discuss this last argument; my discussion of the Schaffer and Dowe examples, together with Beebe's point about the nature of events, are enough, I think, to show that omissions and preventions are not, and

cannot be, “causation full force”, as Schaffer (2000: 289) claims, in the sense of figuring in the basic ontology of manifest physical causation.

However, while even if these arguments show that a description of a manifest physical causal process does not require reference to omissions and preventions, this does not apply to a description of a *mechanism*, such as the ones Schaffer describes in his assassin example. I now turn to a brief discussion of the philosophy of mechanism, which has been a subject of growing interest in the last couple of decades. I include it because a consideration of the ubiquity of mechanisms in physical causation provides further evidence against the possibility that mental content causally affects the physical world.

(iii) Mechanism and Manifest Physical Causation.

I will discuss mechanism only to the extent of its relevance to the above aim (see Machamer, Darden, and Craver 2000; Glennan 1996, 2002, 2009, 2011; Craver and Bechtel 2007; Craver 2007; Illari and Williamson 2012; Gillett 2012; Craver and Darden 2013). Stuart Glennan (2009: 315) defines a mechanism as a set of “systems or processes that produce phenomena in virtue of the arrangement and interaction of a number of parts”, while Illari and Williamson’s (2012: 119) definition is broader:

A mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon.

The philosophers working in this area are generally engaged in a search for a model of scientific explanation that will be applicable, in particular, to the physical special sciences, such as neuroscience (Craver 2007), and will be an improvement on, for example, the covering law model, which is widely thought unable to deal adequately with the special sciences (see Gillett 2012: 76). Special scientific investigation tends to be interested less in the discovery of laws – which, in the special sciences, may almost always be expected to be *ceteris paribus* – than on finding out how things, such as the processes of genetic transmission or continental drift, work.

I think it is clear from Schaffer's assassin example ((ii) above) that the notion of mechanism dovetails neatly with that of manifest physical causation. The operation of a mechanism requires, as well as instantings of causal processes ("activities"), the presence of a set of what one might call boundary conditions, that is, a precise arrangement of objects and substances ("entities") in a particular set of spatiotemporal relations, quantitative and qualitative, to each other and their environment. Glennan suggests (2009: 323-324) that the notion of a mechanism should include not just comparatively stable systems like, say, a handgun or a neuronal membrane, in which processes are regular and repeatable, but also singular causal processes, such as a baseball's breaking a window, in which the parts and their relations come together in an ephemeral configuration that may only form once – in this case the velocity and mass of the ball, the fragility of the glass, and their relative positions. Another such instance might be de Muijnck's example ((i) above) of "the structure of the physical interactions among puncturing instrument, can, water, poison, and man"

I suggest that, in Schaffer's examples, it is the combination of the operation of straightforward *productive*, manifest physical causality with the presence of the structure and the relations, spatial and temporal, amongst the relevant parts, that constitutes the mechanism. It is worth noting that the case of Scarlet Gilia discussed in chapter 2:III:(iv) also fits this model. The microbiological causal events that constitute Gilia's regular colour change, which can themselves be regarded as mechanisms with simpler causal components, take place in wider circumstances such that together they constitute a still higher *adaptive* mechanism. It is when we view mechanisms as a whole that they can give the appearance of necessarily involving, for example, causation by omission. But in fact omissions are not part of what is going on at the basic causal level (basic, of course, in the manifest physical rather than a fundamental sense – we might even *define* this basic level as that at which there is no need to invoke such things as omissions or preventions). Reference to omissions becomes necessary when we describe the mechanism as a whole.

Mechanisms, whether artifactual or natural, are clearly part of the manifest image of man in the world, whatever their ultimate ontological status. "Discovering a mechanism", writes Glennan, "is the gold standard for establishing and explaining causal connections" (2009: 315). He acknowledges (Glennan 1996: 50, 56) that, just as we have noted for the case of causation itself (I:(ii) above), the notion of mechanism would seem to have no application at the fundamental level, and on this basis he suggests there is a dichotomy between fundamental physics and the rest of science. Kutach (2013) makes a similar distinction, on a similar basis, between fundamental and "derivative" aspects of causation, and his further division of derivative into "non-

metaphysical” and “metaphysical” kinds corresponds to the distinction I have made between ID and manifest physical causal statements. Glennan’s dichotomy, and Kutach’s first distinction, approximate the distinction I have suggested in I:(v) above between the scientific and manifest images. What I am calling manifest physical causation appears to be well suited to fill the role of the kind of causal connections Glennan means - we can say that when manifest physical causes are instanced in a suitably structured setting, we have a mechanism. What mechanisms in general have in common seems to be that a source of energy of some kind – in a cell, typically high-energy phosphate bonds; in an internal combustion engine, petroleum ignition – *drives* a process involving mechanical, chemical, electromagnetic or other changes, all of which are causal in the productive sense. One feature of mechanisms, I suggest, that distinguishes them from the theoretical entities of fundamental physics and establishes them as part of the manifest image is that they are such that we, or at least suitably trained people, are often able to *picture* them, and thus see, or grasp, how they work. This is in contrast with the entities and theories of fundamental physics that are usually only graspable mathematically. Ross, Ladyman, and Spurrett (2010), defending a radically scientific approach to metaphysics, warn that “there is no reason to imagine that our habitual intuitions and inferential responses are well designed for science or for metaphysics” (*ibid*: 2). They decry attempts to understand the world in terms of “microbangings” amongst “little things”, denying that there are any such ultimate constituents; “indeed, it is no longer helpful to conceive of ...the world ...as ‘made of’ anything at all” (*ibid*: 4). But whether or not this is a true account of ultimate reality, I hope to have put forward enough argument in this chapter, based especially on McDowell’s transcendental empiricist and other arguments, to establish the

indispensability of the manifest physical world in *our* reality, the world that rationally constrains us (1:I:(ii)). In this reality, mechanisms have a central role, both in everyday life and in its extensions into biology and engineering. The design of mechanisms pre-dates the scientific revolution by millennia, as in for example the water clocks of the ancient world, whose existence presupposes designers whose knowledge of how the manifest physical world works enabled them to envisage their structure and workings in advance. Understanding of biological mechanisms, on the other hand, proceeds by reverse engineering by investigators. It seems likely that to some extent it was early investigators' knowledge of the artifactual world that enabled them to recognize, say, the heart as a pump or the lungs as a bellows. But none of this, of course, establishes that there cannot be mechanisms that would be too complex for us to grasp in their entirety; this may be true of the most complex mechanism we know of, the human brain.

Finally, it is worth noting the similarity between the above picture of mechanisms and Baker's account of material constitution, which I don't think is at all coincidental. If we take what I have called boundary conditions – the structural arrangements of parts – to be favourable circumstances, we have the result that a mechanism *m* is constituted when aggregates of materials are in *m*-favourable circumstances (see Baker, NFPP: 225-226). Craver and Bechtel (2007) specifically invoke constitution: “Our account of mechanistically mediated effects... allows us to retain a univocal conception of causation as intralevel and to account for relations between levels in terms of constitution” (*ibid*: 554). “Constitution” here is not meant in Baker's precise technical sense, but is nevertheless similar. These authors draw attention to how, even in science, “there is a temptation to say [wrongly] that the behaviour of the

whole mechanism is caused by the activity of its parts” (*ibid*: 555), and *vice versa*. For instance, taking one of Schaffer’s examples, it would be a mistake to say that the anoxic change in neuronal metabolism that leads to the cell’s reaching an equilibrium state causes the death of the cell. Rather, it *is* the death of the cell; in Baker’s terms, this is the “is” of constitution. Craver and Bechtel rely on the notion of *levels* of mechanism (*ibid*: 551), which play a similar role to Baker’s constitutional levels. So a series of causal events in a mechanism at one level constitutes (in favourable circumstances, such as that the parts of the mechanism are organized in such a way that they are components of a neuron which is a component of a nervous system) the death of the cell. Similarly, though in the other direction, when Schaffer’s gunman pulls the trigger he does not cause the unblocking of actin-myosin binding. The unblocking, rather, is caused by another part of the contraction mechanism, and the operation of this mechanism, as a whole, constitutes his trigger-pulling.

It seems that mechanisms provide a model of how causation and constitution might interact. However, Craver and Bechtel emphasize that there is nothing mysterious about the interaction in the scientific and other examples they give, in the sense that no processes or properties are involved beyond those that are already understood within the manifest image of the physical world. As they write “higher levels of mechanisms are, by definition, mechanistically explicable” (*ibid*: 550). The account I am presenting, while based on the same structure of interlevel constitution and intralevel causation, is concerned with these relations as they apply between the level of the manifest physical world and that of mental and ID causes and relations, which, I am claiming, has its own distinct causal nexus. It thus has no mechanistic explanation in Craver

and Bechtel's sense. The causal constraints on ID 'mechanisms' involve the instantiations of semantic and normative properties that have no counterparts in the manifest physical causal nexus.

If the conclusions of these authors who have analyzed the relation between physical causation and mechanism are correct, they present a powerful argument against Baker's view that all property-instances are physical. Her claim that Jones's voting's causing Smith's anger is a causal relation between two physical property-instances that is independent of any lower-level causal processes founders, I think, when one realizes that the way in which the claim is framed precludes there being any physical mechanism involved. Rather, the claim is that Jones's action *directly* causes a change in Smith's mental state. I will argue that this is indeed what happens, but that the causal relation is not part of the manifest physical nexus.

(iv) Manifest Physical Causation and Mental Content.

To repeat, my point in developing this account of manifest physical causation and its role in mechanisms is to argue that there can be no causal role for mental content in the manifest physical causal nexus. The question of whether mental content can have a causal role has, of course, featured prominently in chapters 2 and 3 and section I of this chapter. The Davidsonian and the Kimian approaches have different emphases, although their conclusions are similar. Davidson, as we saw (chapter 2), appeals to the holistic and open-ended character of mental predications to argue for mental-physical anomaly. Kim's approach (chapter 3) concentrates on properties as exemplified in events. He argues from the principles of causal-explanatory exclusion and the causal closure of the

physical to the conclusion that unless mental properties are reducible to physical properties they can have no causal role.

My argument here appeals to both conclusions, but where it differs is by drawing attention not to a distinction between mental predications or properties and physical or microphysical theory, but to a distinction between these and the ordinary manifest physical world. My claim, again, is that the *manifest* physical is causally closed – that anything (say, any event) that has a manifest physical cause at t has a complete and solely manifest physical cause at t (I:(vi) above). The key to understanding the complexity of the manifest physical world is the realization that manifest physical causes are involved in mechanisms.

I think the empirical evidence is overwhelming that manifest physical causation, including, of course, that involving our bodies, is causally closed. The principle that the physical is causally closed is accepted by virtually all parties in the debate, including Hornsby (ACE), McDowell (email communication, March 2012), and Baker (MEL: 117-120). When applied to the *manifest* physical the principle has the added endorsement of common sense, at least among modern educated people, who do not in general believe there are gaps in physical causal chains. If there were, not only would physics, as Beckermann claims (chapter 2:II:(i)), be a “weird science”, but the everyday world would also be weird. This is why we assume we are being tricked when confronted with, say, the showman Uri Geller’s purported telekinetic bending of spoons (Marcus 2012: 235-237).

Generally speaking, however, it seems that only philosophers make the connection between the weirdness and impossibility of telekinetic spoon-

bending and the problem of how an externally-individuated mental content, like Jane's desire and intention to comply with airport security procedures, can have physical effects, such as a change in the physical location of Jane's arms. From what we know of neurophysiology, we can assume the physical mechanisms that subserve Jane's arm-raising initially include multiple arrays of spatiotemporally organized changes in the membrane potentials of neurons in Jane's motor cortex. But these changes belong, like any others, in the physical causal structure of the world. We know that they have physical causes – specifically, changes in the configurations of trans-membrane ion channels brought about by the locking of neurotransmitter molecules released by other 'upstream' neurons on to membrane receptors on the target neuron. These well-understood mechanisms do not belong to the world-view of Hornsby's "atomic physicist" (CPCP; see 2:II:(i)), but, I have been claiming, are part of our manifest image, even though they have only been so for a comparatively short time, and even though they are so only for some specially trained people. And I do not think there can be any doubt that these manifest physical processes are causally sufficient for their effects. We know, for example, that if we trace the causal chain back in time, what we will encounter will be more neural activity. But there does not seem any need to invoke Jane's mental content in the description of this causal network.

A number of philosophers (e.g., Loewer 2002, Thomasson 1998, and, as we saw, Baker MEL: 100-102) have argued that the non-contingent nature of mental-physical relations would allow that there could plausibly be widespread overdetermination of physical events by both mental and physical causes. Several counterfactual theorists also argue for the causal efficacy of the mental. Menzies (2013), for example,

presents a sophisticated argument using a modification of the causal exclusion principle framed in terms of causation as difference-making (that is, counterfactual dependence-supporting) rather than causal sufficiency. His analysis, using possible world semantics, identifies situations in which, he claims, it is appropriate to cite the supervenient (such as mental), rather than the subvenient, of two or more candidates as being the real cause of an effect – that is, that causal exclusion can operate downwards as well as upwards.

But when we consider that neurons and neural connections are mechanisms, which involve causation as production, these sophisticated arguments are seen to miss the point. We have every reason to believe that manifest physical mechanisms are causally influenced only by physical forces – mechanical, electromagnetic, or chemical. In the last chapter I concluded that Baker fails to refute Kim’s argument (3:III:(i)) that if mental properties are not reducible to physical they have to be regarded as epiphenomenal, and in my discussion of Jane’s arm-raising (3:IV:(iii)) I concluded that the threat of epiphenomenality is also a problem for the constitution account, despite Baker’s claim that mental and ID property-instances are themselves physical. Even if it is conceded that Jane’s willing, M, independently causes her arm-raising, M*, I argued, M’s physical properties, and hence its physical causal powers, do not differ from those of its constituting physical property-instance, MP. The relational properties of being in the appropriate environment, and so on (i.e., the favourable circumstances), which enable M to qualify as Jane’s willing to raise her arms for a security check, can make no physical difference to this local and intrinsic causal process. Therefore it seems that the only way in which semantic content could possibly causally influence a neural mechanism *qua* semantic content would be if

there was a relation of identity between semantic properties and the properties instantiated by these physical causal forces – that is, if semantic content actually *is* (the ‘is’ of identity) the causal force embodied in the neural mechanism - and nearly all the philosophers under discussion here rule this possibility out. So long as mental *properties* are thought of as distinct from physical – even if they are token- identical with, co-instanced with, or constituted by physical properties – they are debarred from contributing causal efficacy to neural mechanisms. If this is right, theoretical counterfactual analysis, as in Menzies’ account, cannot deliver physical causation by content if, as I believe is the case, we have no account of *how*, in the sense of by what mechanism, such causation could operate.

We are forced to conclude, I think, that proposed solutions to the problem of mental causation which involve claiming some kind of tight relation, such as constitution or token identity, between mental and physical property-instances fail to achieve the intuitively desired result, which would be that mental properties and property-instances cause physical changes *in virtue of* their semantic content. They fail, that is, to establish the causal relevance of mental content.

This is certainly not a new conclusion. As Kim (PSNE: 73) and others have pointed out, the question of how the mental and physical, thought of as distinct substances, could possibly interact was raised by Princess Elisabeth of Bohemia in a letter to Descartes in 1643. But what I have tried to show in this chapter that the problem of mental causation does not arise from an inappropriate incursion of an alien, impersonal scientific standpoint into our common sense world. Baker uses the word ‘microphysical’ to describe the constituters of Jane’s willing and arm-

raising. I have already suggested (3:IV:(ii)) that ‘neural’ would be a more appropriate term, because I think the level of description that best captures the physical process or processes that constitute these ID properties is that of neural mechanisms. When we consider that for mental content to cause physical change would require that the content be a component of a neural mechanism, rather than being a property-instance that is distinct from but connected by some relation such as co-instantiating or constitution to some property-instance broadly characterized as microphysical, suggestions such as that there might be overdetermination (perhaps of Baker’s “harmless kind” (MEL: 119 n66)) of physical effects by mental properties, come to look incoherent. We see that for an externally-individuated property of Jane, such as her believing she is in an appropriate environment for a security check, to have any causal influence on the workings of a neural mechanism is no more plausible than telekinetic spoon-bending. As Malcolm wrote almost 50 years ago,

If we bear in mind the comprehensive aspects of the neurophysiological theory – that is, the fact that it provides sufficient causal conditions for all movements – we shall see that desires and intentions could not be causes of movements (Malcolm 1968: 58).

To say that the mental cause overdetermines the already sufficient physical cause does not make the picture any more attractive, and arguments about the ontological extravagance of widespread overdetermination are superfluous. Tyler Burge writes that thinking of mental causes “on a physical model – as providing an extra ‘bump’ on the effect” (Burge 1993: 115; see also Gibbons 2006) is likely to be the

source of the incoherence here. It seems we need a different model, and the next chapter will be concerned with working out the implications of this idea.

Neural mechanisms, I have argued, form part of the manifest physical causal nexus. A purported fact such as, say, ‘NMDA receptor activation in Jane’s pre-motor cortex occurs when Jane wills to raise her arms’ is thus in the same category as “spring has begun” (McDowell, MW: 27) – a true (or false) thinkable about the “layout of reality” (*ibid*). Like all facts, it is subject to normative assessment, and so is part of the space of reasons. So, I suggest, the picture of neural processes as mechanisms composed of manifest physical causes reinforces the claim (2:II:(i); I:(v) above) that the problem of mental or intentional causation requires resolution in the philosophies of McDowell, Hornsby, and Baker.

I also think it is clear that that the physical causal relations of the manifest image exhibit an evident continuity or commensurability with each other in Davidson’s sense (ME: 219ff; I:(iv) above). Davidson writes of laws that we can say that the relation of ordinary lower-level causal statements to putatively more precise statements at the micro-level is *homonomic*. Homonomic laws, according to Davidson, draw their concepts from the same “comprehensive closed theory” (*ibid*). We can, however, identify ID causes, as in some of my earlier examples;

The division of powers built into the American constitution makes it difficult for a minority party President to introduce new legislation.

As declarer in contract bridge, Mary's perception of the situation caused her to realize that she should lose an early spade trick while she still had first round control in the other suits, in order to make the rest of dummy's spades.

Jones's voting against him made Smith angry,

in which, firstly, mechanical, chemical, or electromagnetic factors do not figure at all, and secondly, which are, again in Davidson's (ME: 219) terms, heteronomic. To paraphrase him, we have every reason to believe that each time ID causal relations are instantiated, so are lower-level causal relations, but the latter can be stated only by shifting to a different vocabulary. The causal relation linking ID events has quite distinct features which make it independent both of the middle-range physical causal concepts that ground our ordinary concept of lower-level causation *and* of whatever fundamental properties, if any, the latter may turn out ultimately to be grounded on.

(v) *Conclusion.*

I believe I have established that there is a category of manifest physical causation which applies to the norm-free, realm of law world of the manifest image, and which is the basis, through the operations of mechanism, of a nexus of causes and effects from which mental/ID causes and effects are excluded. If this is right, ID property-instances do not, *qua* ID property-instances, causally affect the physical world directly. I have also claimed, first, that manifest physical causes are *real*, or objective, just as the manifest image to which they belong is real, and second, that they are *productive* of their effects in Hall's (2004) sense.

If this is correct, what are the consequences for the reality and status of ID causal claims? It might seem that we are forced to adopt a position similar to Kim's. My motivation for arguing for the category of non-ID causation at the manifest level was to establish that, however things may be at the fundamental microphysical level, the principles of physical causal closure and causal exclusion, which are central to Kim's case, remain unrefuted by Baker's (and others') arguments based on the problematic nature of causation at that level. As we saw (3:III:(i)), Kim concludes that we are forced either to accept that mental (and other ID) properties are identical with physical properties, or to concede that these properties, if they are distinct, must be epiphenomenal – that is, that there is no mental or ID causation. If I have succeeded in removing an objection to this reasoning, it might seem that all I have done is strengthen Kim's argument.

Kim's own response to his conclusion, that mental properties must be functionally reducible to physical properties, is shared in broad outline by many functionalists and type identity theorists (see Melnyk 2003; Bermudez 2005). The epiphenomenalist option - the view that mental property-instances are caused by neural events but themselves cause no physical effects - is regarded by most philosophers as to be avoided at all costs – most would agree with Burge's remark that “[t]he metaphysical grounds that support the [epiphenomenalist] worries are vastly less strong than the more ordinary grounds we already have for rejecting them” (Burge 1993: 97). Alternatively, one might follow Alexander's dictum to the conclusion that mental and ID property-instances should not be included in our ontology at all. This leads to eliminativism, the view that the commonsense propositional attitude ascriptions that

characterize ID causal explanations do not refer to any real entities (P.M.Churchland 1984; P.S. Churchland 1988).

The position I want to argue for is that mental and ID causation are indeed real, but of a different kind from manifest physical causation. They are, however, *constitutionally* related to manifest physical causes. This suggests a different gloss on Alexander's dictum – to have a certain kind of reality, the ID kind, is to possess causal powers of a kind appropriate to that reality. As Tim Crane puts it, the idea is that mental and physical causation are not “homogeneous” (Crane 1995: 17ff). In the final chapter I develop this idea further.

CHAPTER 5: THE NATURE OF INTENTION-DEPENDENT CAUSATION.

I. ID Causation and Causal Nexuses.

(i) Summary.

In this final chapter I bring together the significant arguments of the previous four. The position we have reached is this. McDowell (chapter 1) defends a philosophy of the relation between human persons and the rest of nature that is nonreductive and, in my view, one that could help to take philosophy out from under the shadow of science and restore it to a central place in our culture. But (chapter 2) McDowell leaves some important questions about the relation between the space of reasons and the realm of law in need of answers, and in my view his account contains inconsistencies, especially in the area of the relation between agential and nomological causation. Baker (chapter 3) presents an account of a constitution relation which exists throughout nature, and which holds out the promise of a way of linking McDowell's two logical spaces. But, I have argued, Baker's account also encounters problems when it comes to accounting for mental and ID causation. I have tried (chapter 4) to consolidate my position by developing an account of causation within the manifest image, and arguing that mental and ID causation are not part of the manifest physical causal nexus. I begin this chapter by completing, in the light of my claims in chapter 4, my argument against Baker's constitutional account of independent causation and NRM. I discuss the consequences of its rejection and show how the resources of the constitution account can provide a basis for a distinct ID causal nexus.

(ii) Rejection of Baker's Account of Independent Causal Efficacy.

Baker's account of property-constitution and independent causation is aimed at countering Kim's arguments against non-reductive physicalism, and is presented in MEL using the examples of Jones's voting making Smith angry and of Jane at the airport security check. The first example is intended to show how Jones's angering Smith exhibits the features of Baker's principle of independent causation; first, that Jones's angering Smith – V's causing V* - is independent of the nature of its lower-level constituter, since Smith would still have been angered if the vote was cast by some means other than hand-raising, and second, that V has causal powers "above and beyond" those of its lower-level constituter, P, Jones's hand-raising. I argued (3:IV:(i)) that while this conclusion is valid, the argument's presupposition, that V indeed independently causes V* is question-begging, since it is central to Kim's anti-NRM stance that it does not.

The second example, of Jane at the airport, is designed to show that independent causation does not violate the principle of physical causal closure. I argued (3:IV:(ii)) that what Baker shows is only that the microphysical causal relation MP's causing MP, between the constituters, MP and MP*, of Jane's willing to raise her arms M and her raising her arms M*, conforms to the principle. Further, I argued, the conformity of M's causing M* to the principle is due to the fact that the causal relation between the wide supervenience bases of M and M* conforms to it, and hence, again, the very existence of the higher causal relation is threatened by Kim's reductive arguments.

In chapter 4:II:(iv) I claimed that content can have no direct influence on the manifest physical world. In the light of this and of some of Baker's own presuppositions, I'll now argue, using the voting example, that, *as Baker presents the case*, V cannot be the independent and irreducible cause of V*. We saw that Baker claims that ID property-instances are physical, and thus are part of the same causal nexus as lower-level property-instances. If so, two further features of the account follow. First, Baker allows that V's causing V* involves "a harmless kind of overdetermination" (MEL: 119n67). This is because, as we saw, V and V* have wide supervenience bases (see III:(ii) above), consisting of their constituters together with the supervenience bases of the favourable circumstances of their constitution, and a causal link also obtains between these bases. Second, as I discuss below, Baker insists (NFPP: 231-233) that there is downward causation of lower-level by mental and other ID property-instances. 'Downward causation' may be taken to mean the direct causation of instances of lower-level properties by instances of irreducibly higher-level properties, where 'direct' means that the causation does not operate *via* any lower-level realizers of the higher properties. It is established, then, that Baker holds that independent, irreducible ID causal relations like this one are of the same kind as, that is, able to interact with, other physical causes in a single, but seemingly multi-level, causal nexus.

As I described in 3:IV:(ii), and as Baker agrees (MEL: 118-119), there undoubtedly exists a lower-level causal pathway linking the ID property-instances V, Jones's voting, and V*, Smith's getting angry, *via* their constituting instances and the additional events that form a chain of mechanisms which I labeled 'ae'. And, as I noted (3:IV:(ii)), according to the principle of independent causation "V's causing V* [does not]

depend on any microphysical relation between P and P*” (Baker, email 21.8.12).

But, given these points, it becomes clear how it is not possible that V could cause V* without any dependence on the lower-level relationship of P&ae’s causing P*. Constitution, recall, is a relation of unity without identity. V*, then, *is* P* (the “is” of constitution) in the presence of V*-favourable circumstances. So if V causes V* directly, without any lower-level intermediary, P* is thereby instantiated. But P* is a lower-level state of Smith’s nervous system, and while it is possible (absent the right favourable circumstances) for P* to be instantiated but not V* (see (PC), 3:III:(ii)), it is not possible that V* could be instantiated without P* (or, perhaps, without P* or one of a number of similar states {P*'1, P*'2...P*'n} that are also capable of playing the role of constituting V*). Baker’s claim, then, that V causes V* (both, *ex hypothesi*, physical property-instances) in a way that is independent of any lower-level connection between them, must include the claim that when V* is caused, P* (or equivalent) is instantaneously instantiated, which is tantamount to the claim that V *causes* P* - that Jones’s voting causes, independently and without any manifest physical intermediary, the instantiation of a particular state of Smith’s nervous system. The fact that, *ex hypothesi*, V and P* are on different ontological levels does not preclude this result, since on Baker’s account V and P* are both physical property-instances belonging to the same multi-level causal nexus.

There are several reasons, I believe, why we should find this conclusion implausible. As I argued in chapter 4, we have every reason to think that, in the world of the manifest image – the everyday world that we inhabit – there is real physical causality which is productive in character,

causally closed, and proceeds according to familiar and generally well-understood physical principles. This is supported by Hall's (2004; see 4:II:(iii)) account of productive causality as involving *Intrinsicness* and *Locality*, as well as by the idea, which I have argued is plausible, that all causality at non-fundamental levels involves the operation of mechanisms. Baker's picture of independent causal efficacy, as exemplified in V's causing V*, plainly has none of these properties. As a claim about causation in the physical world, it seems to be a case of just the kind of inexplicable action at a distance that de Muijnck (2003; see 4:II:(ii)) compares with magic. As he says (*ibid.*: 47), without these constraints, "the very concept of causation becomes hollow", since there would be no reason not to allow everything to be causally related to everything else.

We should also note that, if it *was* the case that V was a cause of P*, P* would be causally overdetermined. When P is instantiated and followed by the lower-level causal chain ae, the instantiation of P* is ensured, so it seems there is no *need* to invoke V as a cause of P*. In 3:III:(i) and 3:V:(ii) I touched on Kim's argument in which he claims that if mental properties are irreducible to physical properties, they are causally inefficacious, and there are similarities between that argument and mine here, as I pointed out in the discussion of Kim's abandoned 'supervenient causation' (3:V:(ii)). There is an important difference, in that on the present account the distinction between higher and lower levels is envisaged as being due to the constitution, in favourable circumstances, of the higher by the lower, hence there is no question of the higher being reducible to the lower, while for Kim the relation is one of supervenience or realization.

But a crucial question is, even if it is conceded that V is constituted by, and irreducible to, P, and that V* is similarly related to P*, and, further, that P causes P*, does this entail that V *causes* V*? As I discussed in connection with causal closure in 3:IV:(ii), I think it does not; P's causing P* is a manifest physical causal relation, as described in chapter 4. *Ex hypothesi*, V's causing V* is also a physical relation, belonging in the same causal nexus as the lower-level relation, yet independent of it. Nothing in V's causing V* corresponds to 'ae', the chain of neural and other physical events leading from P to P* - in fact *no* manifest physical connection links Jones's voting, *qua* intentional action, with Smith's getting angry, *qua* mental event. That V causes V* is not deducible from the argument. In my view, our intuition that V causes V* is an example of the less strict, explanatory sense of 'cause' that I contrasted with manifest physical causation (4:I:(v)).

There remains Baker's other defining feature of independent causation – that an independent cause would have had its effect even if its constituting property-instance had been different. But suppose V had been constituted by, say, a telephone conference vote, in which each voter in turn gave his name and his vote. When Jones said “Jones, nay” a different lower-level causal mechanism, say P' causes P*, this time involving Smith's auditory system, would have caused P*, his neural state, to be instantiated, and V* would have been constituted just the same. So the multiple realizability of V establishes its independence from any *particular* lower-level constituter, but V's (purported) power to cause V* remain dependent upon its having *some* appropriate constituter – one that is *functionally* equivalent to P.

(iii) *The Extent of the Physical.*

Earlier I suggested that one major reason why Baker does not acknowledge the problems I find with her account (IC) is her counterfactual-based, dependence view of causation (3:V:(ii); 4:I:(i)). I think, however, that her view, introduced at 3.IV:(i), that all property-instances are physical also leads to serious problems, central though it is to her philosophy. If the account of the manifest physical and its causal properties that I developed in chapter 4 is right, Baker's view cannot be upheld.

In the voting example, P, Jones's hand-raising, a physical property-instance by definition, constitutes V, Jones's voting, in favourable circumstances, and according to Baker V is also a physical property-instance. V has, through its constitution, irreducible causal powers that are independent of P's – powers that, on Baker's account, we must also think of as physical. P has only (manifest) physical causal powers – mechanical and electrochemical. But in the constitution sense, V *is* P – it is just P with the added relational property of being instantiated in an environment where hand-raising counts as casting a vote in a shared social milieu. On this view, then, the property-instances 'being a hand-raising' and 'being a hand-raising in an environment where hand-raising constitutes voting' are instances of different physical properties, with quite different causal powers.

There are parallels here with the widely held objection to Davidson's anomalous monism according to which his account fails to establish the causal and explanatory relevance of mental properties (2:III:(i)). On a productive understanding of physical causation, no mechanism exists through which an addition of semantic, or normative property-instance such as 'being a voting against Smith' could contribute causally to P's (V's) causal powers, and labeling such a property-instance as physical

will not confer this ability. In fact, Baker's account seems to have even greater problems in this regard than Davidson's. While for Davidson, mental events are physical because they are token identical with physical events, and are causally efficacious *just as* those events, for Baker, constituted mental property-instances, constituted as they are by physical property-instances with their own causal powers, acquire their causal efficacy in the form of *additional* physical causal powers that they accrue simply by being in a particular environment. One cannot help feeling that physical causal powers are being conjured out of nowhere.

Baker justifies the assertion that all property-instances are physical as following from her view that all non-fundamental entities are ultimately constituted by physical entities. One way in which this could perhaps be disputed is by arguing that, while constituted entities generally, in Baker's schema, share derivatively the non-derivative properties of their constituters, a property like physicality could be considered to be *essential* to its bearer, hence a modal property, and that as such it would belong with the group of excluded properties (II:(i) above) that, according to Baker, need *not* be conferred upon constituted entities by their constituters.

Jessica Wilson suggests the following characterization of what it is to be physical:

an entity is physical just in case it is (approximately accurately) treated by current or future (at the end of inquiry, ideal) physics, and is not fundamentally mental (Wilson 2005: 428).

This last condition would seem to rule out, for example, Smith's getting angry as a physical property-instance. And there seems a reasonable case for adding "...or a property that depends for its instantiation on the existence of creatures capable of bearing propositional attitudes" to Wilson's definition, thereby excluding all ID property-instances from being physical. This need not be incompatible with such property-instances' bearing a relation – constitution and/or supervenience, say, to physical property-instances. I won't discuss the possibility that there are other non-physical property-instances – modal or mathematical, for instance, that the definition does not account for (II:(ii) below).

There is, of course, controversy over the exact nature and limits of the physical (see, e.g., Stoljar 2010: 28-50), and, despite arguments to the contrary like Davidson's in ME, maybe one cannot rule out *a priori* the possibility that some future physics will be able to incorporate the mental. But on the understanding of physical that I introduced in the last chapter, and which I contend supports an account of manifest physical causation as an objective phenomenon, the difference between the property of being, say, a chemical gradient across a neuronal membrane and that of being an intention or willing to raise one's arms marks a fundamental distinction. The claim that all instances are physical, then, not only robs the term of the ability to do any useful discriminatory work, but also threatens to obscure this vital distinction. Again, merely stipulating that ID property-instances like being a willing or having value are physical is not enough to make it unmysterious that they could be causally efficacious in the physical world, in the same way as, say, a falling rock or an electric current is efficacious. ID property-instances also seem to be subject to quite different constraints when contrasted with lower-level causes characterized in terms of these mechanical,

electrical, or similar properties – the same contrast that appears in McDowell’s central epistemological distinction between (to repeat yet again), “explanations in which things are made intelligible by being revealed to be...as they rationally ought to be [and] ...explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen” (FAM: 328); that is, the distinction between the space of reasons and the realm of law. My suggestion, then, is that Baker’s ontological claim that all property-instances are physical is rendered irrelevant by the much more important distinction between the kinds of causal powers possessed by, and the kinds of explanatory constraints governing, instantiations of lower-level and ID properties

(iv) Downward Causation.

As I explained (4:II:(i)), I think the reason Baker believes her independent causal efficacy account vindicates ID causation is that she assumes a view of causality on which the determinants of whether a causal relation exists in a particular case are, first, whether there is a causal explanation, and second, whether counterfactual analysis can show that if the (putative) cause had been absent, so would the effect. Further, Baker’s claim that all property-instances are physical allows her to assume that virtually any property can be causal at any level and, seemingly, across levels.

I say “seemingly” because Baker’s position on causation between levels is to some extent ambiguous. Recall (3:V:(i)) that, in the voting example, she claims that “[t]he contribution of Jones’s hand going up to Smith’s anger was exhausted by the fact that the hand’s going up constituted a

vote against Smith” (MEL: 116) – that is, there is no upward causal relation between Jones’s hand-raising and Smith’s anger. Elsewhere (NFPP: 225-227), she attributes the (causal) power of a fire escape to save a life – an ID causal power - not to the “sums of particles” of the fire escape but to its having been designed for its purpose, that is, to its properties as a constituted object. So an aggregate of materials, in the fire escape-favourable circumstances of structural arrangement and purposeful design, constitutes a fire escape, an ID object which thereby acquires ID causal powers. Here, then, Baker’s account seems to be in line with Craver and Bechtel’s (2007; see 4:II:(iv)) claim that causal relations are intra-level, constitution relations inter-level.

One would think that the same considerations would rule out direct *downward* causation. But in NFPP (226-233), Baker is explicitly committed to downward causation from ID to physical levels. She presents empirical findings, for example, evidence that hippocampal size in London taxi drivers increases proportionately with the duration of on the job navigational experience (Maguire et al 2000), as proof of “the causal efficacy of upper-level properties” (*ibid*: 231-233). Baker writes,

To deny that these examples are cases of downward causation by giving a reductive interpretation of upper-level properties seems like a “work-around” especially in light of the fact that no one has an inkling of what a reduction of, say, learning our way around London might be (*ibid*).

But it seems that Baker is once again retreating to an epistemic defence (see 3:III:(ii)) – appealing to our ignorance in order to reject a metaphysical argument. It is true that the details of a reduction of the

learning are not accessible, but even on Baker's own account, the taxi drivers' learning must be constituted, in favourable circumstances, by neural events or states. Even if we remain forever ignorant of the exact nature of these, we at least know that they must be finely enough individuated to distinguish between, say, knowledge of two different routes from Paddington to Victoria. I don't think it can be in doubt that the hippocampal enlargement is caused by these other neurophysiological changes. To think otherwise would be, once again, to endorse a kind of magical, mind over matter process. If this appears to have things the wrong way round it is because we naturally, and correctly, say that the learning rather than the neural changes provides the explanation. Our common sense inclination to do this is vindicated if, as I suggest in the next section, constituted, ID causal properties subsume the lower-level ones that constitute them.

Baker's epistemology-based interpretation of causal relations and her claim that all property-instances are physical combine with her denial, based on the "no bottom level" possibility (see 4:I:(i)), that any causal level is privileged, to enable her to reject or ignore some of Kim's stringent metaphysical arguments, notably his causal exclusion argument and his argument against systematic overdetermination of physical effects by mental and physical causes. I have argued against each of these three claims of Baker's individually in chapters 3:IV and V and 4:I. In chapter 4 I argued at length, first, that the appropriate physical context for discussion of mental causation is the physical world of the manifest image, and second, that within our understanding of causation in the manifest physical world, causation by semantic or relational property-instances is incoherent, given the plausible assumption that these properties are not identical with physical properties. I concluded

that if mental and other ID properties are causal, they do not cause things in the manifest physical world.

However, despite Baker's reluctance to reject downward causation, I am sympathetic to her non-reductive approach to philosophy of mind and mental causation. My response to these criticisms, then, is to extend the constitution account beyond what Baker would accept or think necessary and argue for an independent ID causal nexus, linked by the constitution relation to the manifest physical causal nexus.

II. Constituted Causation.

(i) The ID Causal Nexus

Recall once more, from 2:II:(ii), Harbecke's inconsistent tetrad of premises about mental causation:

(MC) Mental events cause physical events.

(CP) The realm of the physical is causally complete.

(NI) Mental events are not identical with physical events.

(NO) Physical events are not pervasively, or systematically, causally overdetermined.

Baker's position, as I have set it out, is difficult to map precisely on to the premises. She clearly accepts (MC) and also professes to accept (CP). The difficulty, once again, stems from her insistence that mental events, like all property-instances, are physical. So her version of (NI) is the claim that mental events are not identical to the events that *constitute*

them, but because they are nevertheless physical, this does not debar her from claiming that her account is consistent with (CP). But since mental events and their constituters are both physical on her account, Baker must allow that they are *both* causally efficacious, which is why she rejects (NO), claiming that the overdetermination involved is “harmless” (MEL: 119n66).

The position I now want to argue for is based upon Baker’s constitution account, but differs from her account in ways that are succinctly captured by its treatment of the inconsistency expressed in Harbecke’s quadrilemma. My position accepts (CP), (NI), and (NO), but rejects (MC) as it is usually understood. I argue that mental and ID events do not cause physical effects directly, but are related to them indirectly through the relation of constitution. They thus are distinct from physical effects and neither compete with them nor causally overdetermine them, while the causal completeness of the physical realm remains intact. I contend that if mental and ID events are understood in this way, the problems highlighted in I:(ii) and (iii) above are resolved.

I also want to argue that ID causes and effects interact in a causal nexus that is distinct from that in which manifest physical causes have their effects. In other words, the position is that the mind does not directly causally affect the physical world, or *vice versa*. Positions like this are usually labeled ‘epiphenomenalism’ and rejected out of hand by most writers on this subject as flying in the face of common sense. For example, I have already mentioned (4:II:(vi)) Burge’s comment that the grounds for rejecting epiphenomenalism far outweigh any metaphysical arguments in its favour. But the position I defend is not epiphenomenalist in this sense. It is true that ID causal properties do not

have effects in the manifest physical causal nexus, but this, I believe, is more than compensated for by their causal efficacy in their own sphere. Further, ID properties are *constitutionally* connected to the manifest physical order, and as we have seen, it is an essential (although, it has to be said, largely unsupported by argument) feature of Baker's account that constituted entities ontologically subsume –trump, one might say - their constituters. Far from being consigned to a second class existence as 'causal danglers', or mere shadows of the entities that do the real work, then, ID causal properties are restored by the constitution account to their rightful place as real causes and real explanations in a way that accords with our intuitions. So when we say that Jones's voting against him angered Smith, we are describing the *real* cause and providing the real explanation, rather than these roles being devolved to some lower, non-intentional level. The account thus allows us to avoid concluding, with Kim (4:V:(ii); PSNE: 62), that it is "an empty verbal ploy" to attribute causal efficacy to supervenient properties.

In the following sections I elaborate further on the concept of an ID causal nexus. I begin by discussing the differences between manifest physical and ID causation and the necessary constraints that ID causation must be subject to. I then develop my suggestion that extending the constitution account to include constituted higher causes – Constituted Causation. I go on to compare my concept of ID causation with other attempts to formulate non-reductive accounts of mental and ID causation - notably that of Eric Marcus (2012), whose position is the closest to mine that I have found in the literature - and to discuss what the ontological status and relata of a distinct ID causal nexus might be. Finally, I address the question of the similarities between, on the one hand, the concept, developed from Baker's constitution account, of the

ID causal nexus, and on the other the Sellarsian /McDowellian concept of the logical space of reasons.

(ii) *Constraints on ID Causation.*

In her Principle of Independent Causal Efficacy (IC),

A property-instance that has an effect *e* has *independent causal efficacy* if and only if (i) it would have had its effect *e* even if its constituting property-instance had been different, and (ii) it confers causal powers that could not have been conferred by its property-constituting instance alone (MEL: 115),

Baker encapsulates how Property Constitution enables the constitution, in favourable circumstances, of novel, irreducible, and independent causal powers. I have been suggesting that a very large category of higher causes – the mental and ID ones – have the further property of acting in their own distinct causal nexus. Our understanding of manifest physical causation, both in ordinary use and in the physical sciences other than fundamental physics, includes, I claimed in chapter 4, the assumptions that it is physically grounded, causally closed, and productive in Hall's (2004) sense, and that manifest physical causal mechanisms involve instantiations of physical laws of various kinds. But no such grounding underpins ID causal relations, which appear to be subject to quite different sets of constraints. If this is right, we need a separate and distinct account of ID causation.

In the light of Baker's account of how ID causal property-instances are constituted, and of my discussion, I suggest a reformulation of the principles she attributes to Kim as they might apply to ID causation:

- (1'). An ID causal property-instance is *constituted*, in favourable circumstances, by a manifest physical property-instance.
- (2'). ID causation, *per se*, either does *not* require nomological sufficiency, or perhaps (if it should turn out that there are, say, psychological laws or laws of economics) requires nomological sufficiency of a distinct, ID kind. When an ID causal event is instantiated, no laws of the kind governing manifest physical causal events are applicable.
- (3'). When an ID causal event is constituted by a manifest physical causal event, the cause of the ID event is not the cause of the manifest physical event.
- (4'). There is no causal inheritance between instances of ID causation and their constituting instances; the causal inheritance principle is replaced by (IC), the principle of independent causation.
- (5'). It is true that any manifest physical event that has a cause has a complete physical cause (physical causal closure). But ID causal events, as such, are not subject to this constraint.
- (6') Because manifest physical and ID causation operate in distinct nexuses, there is more than one cause and causal explanation of any event that has both a manifest physical and an ID characterization.

These revised principles almost entirely concern the constraints that ID causation is *not* bound by – they give us very little positive information

about ID causation. I think that there is one vital constraint, deriving from the manifest physical, that does apply to ID causation. This is that it should be naturalistic in the sense that it should be compatible with, that is, not violate, physical law. De Caro and Voltolini (2010: 71) propose a constraint they call the *constitutive claim* of contemporary naturalism:

No entity or explanation should be accepted whose existence or truth could contradict the laws of nature, insofar as we know them.

The authors emphasize the differences between this constraint and the narrower ones which they see as binding *scientific* naturalism (*ibid.*: 72-73), which are first, “that ontology should be shaped by the natural sciences alone and that, in principle, the natural sciences can account for reality in all its aspects” and second, that philosophy is, or should be, continuous with science. They regard the first constraint, the constitutive claim, as compatible with a ‘liberal’ naturalism of the kind of which McDowell is a prominent advocate (1:I:(*ii*)).

They point out (*ibid.*: 76-78) that the constitutive claim is compatible with the admission to our ontology of irreducible entities whose existence and nature lie outside the scope of natural scientific characterization, and further, that it is compatible with the properties of these entities supervening on those of entities that *can* be so characterized. They include among the entities in question values, abstract entities such as numbers, modal properties, free agents, and conscious phenomena (*ibid.*: 72). A consequence of the constraint, however, is that these entities “make no difference in the causal order of

the world” (*ibid*: 78) – if they did, they would “contradict the laws of nature”. Understanding “the causal order of the world” to include the order of what I am calling manifest physical causation, we can see that ID causal properties, on the account I have been developing, conform to these criteria. They supervene widely upon the physical causal order, but do not, *qua* ID entities, interact causally with it. In NFPP, Baker expresses sympathy with these views but rejects the authors’ claim that the non-scientific entities countenanced by liberal naturalism must lack causal powers (NFPP: 12-17). In email correspondence (21.8.12) and in NFPP: 15n11 Baker reports that de Caro and Voltolini have retracted this claim. De Caro confirms (email 12.6.13) that he now wants to leave open the possibility that the physical is not causally closed. However, I remain committed to these entities’ lacking manifest physical causal powers. Indeed, if the term “laws of nature” is interpreted broadly enough to include not only the laws of physics but those of the physical special sciences as well as well-understood constraints on the behaviour of everyday objects, de Caro and Voltolini’s constitutive claim seems to me to encapsulate precisely the constraint that, I am claiming, both McDowell and Baker fail to observe in their views of causality. A *prima facie* difficulty with accepting the claim is that it appears directly to contradict Alexander’s Dictum, that everything that exists has causal powers (see 3:III:(i)). As Harbecke (2008: 81) points out, “if such things as causally non-eficacious entities existed, there would be no way of knowing that they did”. I will argue, however, that they interact causally with other ID entities, and that we, as inhabitants of the space of reasons, do thereby have a way, unique to us or creatures like us, of knowing they exist. Like McDowell, (RGM: 238), however, I think we must be careful not to think of these ‘entities’ as immaterial substances, but perhaps as properties or aspects of reality (see V:(iv) below).

A further constraint that should accompany acceptance of the constitutive claim is that the Nomological-Sufficiency and Causal-Closure Principles are inviolate as they apply at the manifest physical level. Thus, to take Baker's voting example, no account that allows Jones' voting (V) to cause Smith's getting angry (V*) – a case of one ID event's causing another - can be admitted if it does so by requiring that any lower-level entities violate the principles that constrain the behaviour of manifest physical entities. In particular, as I explained above (I:(iv)), ID causes cannot work *downwards* (de Caro and Voltolini op.cit: 78) – if they did, they would causally affect physical mechanisms, which I have argued is incoherent.

(iii) *Constituted Causes.*

I concluded in I:(ii) that, in Baker's voting argument, it is not deducible, from the obtaining of the constitution relations plus the lower-level causal relation of P&ae's causing P*, that V causes V*. I want now to suggest that V's causing V* is ensured if the higher causal relation *itself*, in addition to its relata, V and V*, is *constituted*. This would mean that not just the cause, V, and the effect, V*, are constituted, but also that the causal link between them is constituted, in favourable circumstances, by the indirect causal link between their constituters. The suggestion goes beyond Baker's claims (PC, IC) that a constituted ID property-instance, such as V, has causal powers that are independent of the nature of its constituting property-instance and that are greater than those that that constituting property-instance possesses alone. I take the step of claiming that the *relation* between ID cause and ID effect is constituted because, unlike Baker, I claim on the basis of the above arguments that

such relations form a causal nexus distinct from the manifest physical. We have already seen that Baker's original schema of material constitution for objects can be applied to other kinds of particulars – events, states, and property-instances. An instance of a causal relation, like that holding between $P \&ae$ and P^* , say, would seem to be a particular, and hence to be of the right metaphysical kind to fill the role of constituter of a new causal relation, in this case V 's causing V^* .

Adapting the schemata developed by Baker for material constitution (C) and property constitution (PC), I suggest the following schema for *causal constitution* (CC), that is, for a relation whereby not just two ID property-instances, but the causal relation between two ID property-instances, or events, like V and V^* in our example, would be constituted by a manifest physical causal relation. Suppose we have a purported ID causal event whose constituters are causally connected *via* additional events ae . Let manifest physical events f and g constitute ID events F and G at times t and t' , respectively. Let the instantiation of a complex event or chain of events $f \&ae$ that includes f , together with a large grouping of other nomologically necessary manifest physical events ae , beginning at t , cause the instantiation of g , at t' . Adapting Baker's established format;

(CC) $f \&ae$'s causing g constitutes F 's causing G =(df):

- (1) $f \&ae$'s causing g is in $\{F$'s causing $G\}$ -favourable circumstances during $t - t'$: &
- (2) it is necessary that if $f \&ae$ causes g during $t - t'$ and $f \&ae$'s causing g is in $\{F$'s causing $G\}$ -favourable circumstances during $t - t'$, F causes G during $t - t'$; &

(3) it is possible that $f \&ae$ causes g during $t - t'$ and F does not cause G during $t - t'$.

In the voting example, for $P \&ae$'s causing P^* to be in $\{V$'s causing $V^*\}$ -favourable circumstances would be for the right psychological, social, and environmental factors to be present to make it rationally and emotionally appropriate, from Smith's first-person point of view, for him to become angry on realizing that Jones has voted against him. The possibility contained in (CC)(3) that, in the example, V does not cause V^* even though $P \&ae$ causes P^* , would arise if the circumstances were not $\{V$'s causing $V^*\}$ -favourable. For example, if, unknown to either Jones or Smith, or both, the rules that would mean that a hand-raising constituted a vote did not currently apply, then what might appear to either of them as Jones's voting making Smith angry would, given externalist assumptions shared by Baker and McDowell, *in fact* not be this causal relation (see 2:III:(iv); 3:IV:(iv)). To assert or believe that it was would be *incorrect* in the recognition-independent sense that they also share (1:II:(iii)).

Like other constituted entities, V 's causing V^* , as a constituted causal relation, derives properties from its constituting lower-level causal relation. The properties of the causal relation $P \&ae$'s causing P^* include being bound by the nomological sufficiency principle and the causal closure principle, which apply to lower-level causal relations. Does this then imply that the ID causal relation, V 's causing V^* , derivatively shares these latter properties? I think it is consistent with Baker's account to conclude that it does not. Recall (3:II:(iii)) that a group of properties, the excluded class, cannot be held derivatively, and that these include alethic properties that are characterized in modal terms. A

property like being bound by the principle of nomological sufficiency, expressed in terms of necessity, would seem to belong in this category. We can agree that ID causal properties, like other ID properties, *supervene* widely on manifest physical causal properties, thus ensuring that de Caro and Voltolini's constitutive claim of contemporary naturalism is respected. We can concede this, however, while maintaining that ID causal relations, *qua constituted* causal relations, are not bound by the nomological sufficiency or causal closure principles – that is, their adherence to the constitutive claim is quite independent of their constitution relations.

By analogy with Baker's account (C) of material constitution, we can call V's property of causing V* a *primary kind* property (see 4:II:(ii)). I do not intend this to mean that V has the property of causing V* essentially (Baker, email 21.8.12) but rather that, once constituted, the causal *relation* of V's causing V* has the property of being that relation essentially, and thus could not exist without being that relation. The relation, P&ae's causing P*, on the other hand, has the properties of the relation V's causing V* only derivatively. By the same analogy, we can say that while P&ae's causing P* continues to constitute V's causing V*, that is, while the manifest physical causal relation and the appropriate favourable circumstances obtain, P&ae's causing P* has no independent existence as a causal relation (see PB:46; 4:V:(i) above). As Baker writes of material constitution (PB:33), "The identity of the constituting thing is submerged in the identity of what it constitutes. As long as *x* constitutes *y*, *y* encompasses or subsumes *x*". As I suggested (3:V:(i)), it seems that the *higher*-level causal relation, once constituted, has the ontological priority. In our example, V's causing V* is, so to speak, the *real* causal relation, in a reversal of the order of priorities

assumed by reductionist accounts which assume that the real causal work is done at the microlevel. And this, in turn, delivers what is perhaps a more important result, that Jones' voting is the *real explanation* of Smith's getting angry.

In claiming this we need not deny that the relation P&ae's causing P* has its own non-derivative properties – electromagnetic, chemical, and so on. What is subsumed is the relation's primary kind property of being the manifest physical causal relation it is, just as, when Piece constitutes *David*, Piece's primary kind property of being a piece of marble is subsumed, and it acquires, derivatively, the ID property of being a statue. Similarly, we can say that in favourable circumstances the causal relation P&ae's causing P* acquires, derivatively, the ID property of being V's causing V*, while retaining its other physical properties.

One of my objections to Baker's account of ID causation based on principle (IC) ((i) above) was that according to that account a constituted, ID property-instance such as Jones's voting, V, causes, in a way that does not depend on any physical connection, another ID property-instance, V*, despite the fact that V*'s instantiation necessitates the simultaneous instantiation of the physical property P*. I objected that as well as violating the causal closure of the manifest physical, this would be incompatible with the realist, productive account of causality that I defend. To what extent does (CC) answer this objection?

On my account, the constituted ID causal relation V's causing V* is outside the manifest physical causal nexus. Thus it does not violate the causal closure of that nexus, and for the same reason V's causing V*

does not compete with P&ae's causing P*, so that the causal exclusion principle does not apply. Nor is either P* or V* causally overdetermined. It remains true that the instantiation of V*, through its being caused by V, is necessarily accompanied by the instantiation of P*, its constituter. But on this account this need not imply that there is any mysterious action at a distance. Constitution is a synchronic relation, so, on the assumption that there is a very short time interval $t - t'$ during which the causal processes take place, we can say that P* is instantiated through being caused by P&ae, at the same instant, t' , at which V* is instantiated through being caused by V. This is plausible because V's causing V* just is P&ae's causing P* in the presence of favourable circumstances. When the ID causal relation is constituted, no direct physical connection between V and V* is needed, since neither of these has any nonderivative physical properties. Everything physical, so to speak, happens at the lower level.

This account has the virtue that it preserves not only the causal efficacy but also the causal and explanatory relevance of content in a way that accords fully with our common sense intuitions. It leaves us free to formulate accounts of mental and ID causation without the need to tailor these accounts to lower-level causal constraints. If there is a disadvantage, it is that it does so at what some will regard as the unacceptable ontological cost of introducing a dualism of physical and ID causal properties. I will return to ontological aspects of the account after enlarging discussion of the nature of distinct ID causation.

Having formulated constituted causation, we can now see that ID causal statements, such as those I used as examples in chapter 4:I:(v), for example

Excessive sub-prime mortgage lending caused the recession

are, after all, genuine causal statements, but statements of constituted, rather than manifest physical, causal relations. We can say this, I think, even though the constituting bases of the excessive mortgage lending and of the recession, let alone the causal relations amongst these bases, are no doubt too complex and multifarious ever to be clearly identified (see 3:III:(iii)). A hybrid physical/ID causal statement like

The earthquake caused economic upheaval,

can, I suggest, also be regarded as a genuine causal statement if it is understood as shorthand for “the earthquake [a complex manifest physical event] caused a variety of manifest physical effects [tsunamis, floods, wrecked infrastructure, and so on] which *constituted* ID effects which taken together fit the description ‘economic upheaval’”.

Constituted causation provides an explanation for empirical findings such as those in Baker’s taxi driver example. Prolonged and repeated instances of particular brain states or events, we can say, associated with gaining navigational experience, cause increased hippocampal size. The fact that our ignorance of the details of these states or events means that our only way of characterizing them is in ID terms need not, I would argue, affect the argument, which only requires that we can assume that they exist. This lower-level causal mechanism, we can then say, constitutes the ID causal relation, “the learning associated with navigational experience causes increased knowledge of, or memory of, the layout of London”.

I turn, in the next two sections, to some approaches to the problem of mental causation that are similar to the one I am developing.

III. Causal Pluralism.

(i) *The Dual Explanandum Strategy.*

The account of ID causation I am proposing bears some resemblance to what has become known, in a term attributable to G.H. von Wright (1971), and given wider currency by Kim (DREB: 292-297; MPEE: 241: EEPMC: 133; see also Vicente 2002; Stueber 2005), as the dual explanandum strategy. But my account differs from this strategy in ways that would not meet Kim's approval. The strategy is based, as Stueber (*ibid.*: 255) puts it, on "the claim that psychological explanations do not *in general* attempt to explain the same phenomenon or aspects of the same phenomenon as physical or neurobiological explanations". It is motivated by the need to respond to the claim of Kim and others that NRM collapses into either mental epiphenomenalism or reductive physicalism (3:III:(i)). In general, supporters of the strategy accept all four premises of Harbecke's tetrad, but claim that while mental events have physical effects, these effects have non-physical descriptions.

In his (MPEE) discussion of the causal exclusion principle, Kim makes a convincing case that where two putatively complete causal explanations of the same event *are* genuinely independent (in that one is neither identical to nor reducible to the other), then either at least one of them is not complete or at least one cannot be admitted as a genuine explanation. But, as he acknowledges (*ibid.*: 242n10) the principle only applies to

explanations of the same event. So if, say, an ID and a physical cause and causal explanation could be shown to cause and causally explain *different* events or explananda, the principle would not apply. Thus the dual explanandum strategy seems to offer a way of circumventing the exclusion principle.

But, according to Kim, attempts to do this soon run into difficulties. One well-known application of the strategy, which I have already touched on (2:III:(ii)), is that of Dretske (1988). Very briefly, Dretske distinguishes between triggering and structuring causes of an event such as the plant Scarlet Gilia's annual colour change (Dretske 2004), that is, between the immediate (local, physicochemical) causal explanation of the change and the wider (environmental, historical) explanation of how the plant has acquired its regular behaviour of responding in this particular way. More generally, if an internal state C of an organism causes a motor output M, the triggering cause explains why a token occurrence of C's causing M happened just when it did, while the structuring cause explains why C is causally linked to M, rather than something else, or nothing. Dretske (1988) describes how such structuring causal mechanisms are plausibly formed through adaptation in simple systems, and suggests that the triggering / structuring distinction could provide the basis for an account of intentional action, distinguishing psychological explanations of actions from those of "events simpliciter" (Kim EEPMC: 133). Kim, however, writes of Dretske's account:

Are these [structuring] causes physical entities or are they not? If they are not, we have an overt dualism... If they are, then these explananda, special though they might be, cannot serve to separate

psychology from physical theory, and the exclusion problem arises again (*ibid*).

That is, if the structuring causal relation shares the same causal mechanism, and is part of the same causal nexus as, the local or triggering relation, then not more than one of them can be a complete and independent cause. Kim's point, then, is that what is presented as a distinct causal process is actually a part of the physical causal mechanism, and as such is not a distinct cause at all. The explanation of how an internal state of an organism has come to be 'hooked up' to a particular behaviour involves a complex, but straightforwardly causal in the manifest physical, nomological sense, account in which environmental and genetic factors interact. The cause of a particular movement of the organism on an occasion is just one small and localized event in that much larger set of events, and its explanation is thus part of the larger, single, more complete explanation. It is not that the structuring mechanism is epiphenomenal, but rather, that there *is no* distinct structuring cause, but instead a series of local, lower-level causes. The explanation based upon the structuring cause is therefore, given explanatory realism, not a true explanation, even though it has clarificatory value for us. As a causal explanation, it is another example of the first, epistemic or conceptual use of 'cause' that I identified in chapter 4:I:(*iii*). What Kim's criticism of the dual explanandum strategy shows, then, is that attempts to use it to get around the causal/explanatory exclusion principle are doomed to failure if the two explananda are participants in the same causal nexus.

Attempts to apply the dual explanandum strategy to the analysis of ordinary intentional actions run into the same difficulty. Consider the

claim that John's going to the fridge to get a beer has one explanation when described as an action, in terms of John's beliefs and desires, and another when described as a series of neurophysiological and other physical events. According to Kim (MPEE: 242), however, "there is an evident sense in which they "describe" one and the same event, the same concrete happening" – a physical event. For the two explanations both to be correct, complete, and independent, then, would be either for the event to be causally overdetermined or for the two explanations to be in terms of non-overlapping dependency relations. On the first alternative, if an event is determined by two complete and independent causes, the implication is that each of them would cause the event in the absence of the other, and as Kim writes (*ibid.*: 247), it is highly implausible that the psychological cause would be effective in the absence of the neurophysiological, acting "telekinetically" on John's muscles. And on the second alternative, for there to be two non-overlapping relations, the psychological cause would have to be part of a causal nexus that did not interact with the physical causal nexus.

Kim, of course, is a reductionist with respect to the psychological (see Chapter 3: III:(i)), so for him the two explanations are not complete and independent. On the account I am proposing, however, the manifest physical causal process that begins with John's initial movements and culminates with the bottle of beer's removal from the fridge *constitutes*, in favourable circumstances, John's purposeful action, an instance of ID causation. These are not, then, different descriptions or explanations of the same thing, but instances of non-overlapping dependencies. In the terms of Kim's dilemma for Dretske, the relata of this ID causal relation are not "physical entities". Rather, we can say that the causal relatum is John's self-conscious appreciation of what was necessary to satisfy his

desire for a beer, and the effect is his performing the action. These are both, I argue, *contentful* entities, and the causal relation involved is not physical in that it does not belong in the same causal nexus as that to which causal relations in the ordinary manifest physical world belong.

The lesson of Kim's argument against Dretske's dual explanandism seems to be that, to carry weight, a dual explanandum strategy needs also to be a dual cause strategy, which is what I am advocating. Of course, this implies what Kim calls "an overt dualism"; the challenge is to show that it is an unobjectionable one.

(ii) Distinct ID Causation in Context.

For each of the four premises of Harbecke's inconsistent tetrad (**I:(ii)** above), there is at least one theory of mental causation that denies that premise. However, accounts like the present one, which deny premise (MC), that mental events (or properties or states; see Harbecke *op.cit.*: 93-94; 101) cause physical events, are in a small minority. Any non-epiphenomenalist account that denies (MC) while accepting (NI), the distinctness of mental and physical properties, is committed to causal pluralism.

As Tim Crane (1995) writes, both reductive and non-reductive physicalism are attempts to solve the problem of how to accommodate the causal efficacy of mental properties in the face of the causal completeness of physics, in the assumed absence of systematic mental/physical overdetermination. But the very idea that there *is* a problem of mental causation is premised, Crane continues, on the [hidden] assumption that "the notion of causation is the same notion

applied to the physical and the mental alike...[t]here is no conflict – and thus no need for an identity thesis – if the notions of causation employed are so different” (*ibid*: 8). Crane himself does not enlarge on the prospects for what he calls the “denial of homogeneity” of mental and physical causation, and prefers as a solution either a modification of the principle of the completeness of physics or a rejection of the denial of overdetermination (*ibid*: 18-22). The point he is making about causal homogeneity is that its denial, far from being a physicalist solution, would be not just a rejection of physicalism but the removal of what was its motivation in the first place – the need to account for mental causation in physical terms: “[i]t’s only insofar as mental states have effects *in the very same sense that physical states have effects* that we need to think of them as physical states” (*ibid*: 22, emphasis in original). The point of my arguments in the past two chapters, however, has been to show that mental and ID properties, as such, *cannot* have physical effects - that it is a mistake to think that they could, as Burge (1993: 115) puts it, “provide an extra ‘bump’ on the effect”. Eric Marcus (2012: 235; see **IV** below), like Kim ((i) above) calls this the “telekinetic” view of mental causation, as “some kind of *spiritual pushing*” which “involves the causation of [otherwise] inexplicable physical motion”, and claims this is a philosophical distortion of the commonsense view of mental causation as *rational*. Marcus includes Baker (MMC) among holders of the telekinetic view. If this is not how mental causation works, the challenge is to show how it does work.

William Jaworski (2006, 2011) presents an account of mental causation that bears some similarity to mine. It is based on *hylomorphism*, originally the Aristotelian idea that substance consists of matter and

form. Jaworski's explicitly contemporary (Jaworski 2011: 314⁵) hylomorphic theory of mind is based on a more general metaphysics, the "hylomorphic worldview" (269-305), which in many respects resembles Baker's account of material constitution. As Jaworski emphasizes (282-288), the hylomorphic view of the material world also has much in common with the view of mechanisms that I discussed in chapter 4. Its central tenet is that a full description of the world must include not only the fundamental materials of which it is composed but also their organization, the way in which the materials are structured. The best explanation for the empirical descriptive and explanatory success of, especially, the biological sciences, which depend upon such notions as organization and structure, Jaworski argues (297), is that organization and structure are real, irreducible ontological and explanatory principles, over and above those of fundamental physics.

Jaworski insists that the behaviour of organized systems never violates any lower-level laws (275, 288), but that this does not imply that everything is *determined* by the lower-level, physical laws. Fundamental laws, rather, are the enabling conditions for the existence of organized systems that exhibit "causal properties and relations that do not fit the mould set by physics" (291). Structure thus allows the appearance of novel causal features in both the natural world of biology and the artifactual world of engineering and technology. These new causal phenomena do not "influence the behaviour of...individuals in the way that forces do" (302) – that is, they are not causes in the productive sense I discussed in chapter 4. Like Baker's, Jaworski's view of cause is that it is dependent upon explanation; "there are as many causal relations as there are explanatory relations" (296).

⁵ Page numbers are from Jaworski 2011.

Jaworski's idea that systems acquire higher-level causal powers through organization approximates, in a way that applies to the special sciences, of Baker's account of how objects are endowed with novel causal powers through their constitution. The appropriate structuring of a particular grouping of basic materials would make up part of the favourable circumstances for the existence of, say, an organism. Jaworski argues that the reality of structure is the best explanation for the success of the special sciences.

From this general account Jaworski develops a hylomorphic theory of mind (306-357), whose central claim is that the biological notion of structure carries over into the psychological domain. Here, the structures are formed by, or consist in, "patterns of social and environmental interaction" (307ff). Accordingly, the theory is strongly externalist and anti-Cartesian; Jaworski firmly rejects the notion that sensations, beliefs, desires, etc, are private, inner occurrences. His view of mind is similar to McDowell's in this respect and in several related others - he is an advocate of disjunctivism, both of perception and action, and of direct realism in perception, while he also argues that other peoples' mental content is not private but open to view.

My interest here, however, is in Jaworski's views on mental causation (344-352; Jaworski 2006). He claims that because mental phenomena – thoughts, perceptions, actions – are "ways in which lower-level neural processes can be structured or organized" that form "patterns of social and environmental interaction" (309), they admit of different kinds of explanation depending on which aspects of them we are interested in explaining. So, for actions, "explanations that appeal to reasons and

explanations that appeal to physiological mechanisms do not...compete to occupy a single explanatory role” (308). But as I noted, for Jaworski, where there is a causal explanation there is a cause (296). He thus advocates a causal pluralism that is based on an explanatory pluralism. The fact that we can give different answers to a question such as, “Why is Madeleine reading?”, according to whether we are seeking a neurophysiological or a rational explanation, is enough, according to Jaworski (346), to establish that these explanations appeal to different causal relations. Actions, as intentional items, are *rationally* caused by mental states, he claims, while the physical components of actions, such as muscular contractions, are *triggered* by neurophysiological events (345).

But if my arguments in chapter 4 are on the right lines, only a minority of the explanations that we normally think of as causal pick out relations in the physical causal order. I argued (4:I:(v), (vi)) that our ordinary causal understanding of, and causal-explanatory talk about, the world of the manifest image makes use of two different meanings of “cause” – the loose epistemic sense and the objectively real, productive, manifest physical sense. On my account, then, the physical causal relation in Jaworski’s example has a privileged status, and a claim to being identified as the real cause.

Like Baker, Jaworski takes the two kinds of cause that, as he claims, are involved in an action, triggering and rationalizing, to have equal epistemic status, the choice between them being just a matter of what kind of explanation one is seeking. He defends his view against the possible objection that the two kinds cause are in competition by arguing that they have different explananda (347). But his argument is then

vulnerable to Kim's objection to the dual explanandum strategy (III:(i) above)

If I am right, the rationalizing *cause* has *only* an epistemic status in the physical causal order, while the triggering cause, if we identify it with the manifest physical cause has, in addition, a robust recognition-independent reality. To return to Baker's Jane-at-the-airport example, while the action and the physical movement may well have different explanations, it can't be denied that the physical arm-raising is a *component* – the physical component - of the action. On this understanding, the causal exclusion argument *does* apply, so if we agree that the manifest physical – neurophysiological – cause of Jane's arm-raising is real, this seems to force the conclusion that the triggering, or manifest physical, is the sole cause of the arm-raising. So if there *is* a rationalizing cause, it appears that it does no physical causal work in contributing to Jane's arm-raising.

The conclusion seems to be that if the productive account of manifest physical causation described in chapter 4 is correct, higher-level causes like the rationalizing cause of an action that Jaworski proposes, if they are real causes, must belong to a distinct causal nexus, separate from that in which manifest physical causes work. This, of course, is what I have been arguing in the case of ID causes, which, recall (3:II:(i)), are causal relations whose instances presuppose that there are beings who are bearers of propositional attitudes.

As I mentioned, in Jaworski's general account of hylomorphism, he writes that causes that derive from the way materials are organized, for example in biology, do not act on objects "in the way that forces do" (302) – that is, causal-nomologically, or manifest physically, in the terms

I have been using. Presumably he is referring to causes of phenomena like Scarlet Gilia's regular colour change, which is attributable to the species' evolutionary history, together with that of its environment. As discussed above (III:(i)), however, the account of Gilia's colour change can be broken down into a complex arrangement of lower-level causal processes that *are* causal-nomological. That is to say, the two kinds of cause share a common causal nexus.

But if I am right, ID (or rationalizing) causal processes *cannot* be broken down in this way. So what Jaworski should be claiming about rationalizing causes is, to put it simplistically, not just that they do not affect things "in the same way that forces do", but that they do not affect the *same* things *at all*. Admittedly, the notion of a separate psychological or ID causal nexus does not fit well with Jaworski's insistence on the essential embodiment of thoughts, perceptions, and actions, which means, as he writes (307), both that these phenomena are not non-physical and that the physical processes that enable them are not non-psychological. But he is also an externalist, and I would repeat my insistence that *content* that is extrinsically – historically and environmentally – individuated does not have physical effects. Jaworski presents a picture of mental phenomena as patterns of interaction among items that are individuated in just this extrinsic way. In chapter 2:III:(i) I discussed the Macdonalds' (1995c: 104) concept of rational patterns as autonomous with respect to the causal-nomological patterns exhibited by physical properties, and as irreducible to them. Both Jaworski's picture and my account of constituted causation offer metaphysical accounts of how such patterns could exist in the physical world. But while the Macdonalds envisage rational properties as possessing causal efficacy through co-instanting with physical events, on my account, and

potentially on Jaworski's, these properties have their own causal efficacy, and direct causal and explanatory relevance, in their own distinct causal nexus. In my view, Jaworski sees no need to take this step because of his explanation-based view of causation.

(iii) Constituted Causes and Purely Physical Effects.

In Baker's voting example, which I have been using to illustrate how constituted causation might work, the physical and the ID causes have clearly distinct physical and ID effects – Smith's acquiring a neural state and Smith's getting angry, respectively. But what about the vast number of cases in which an action only has a physical effect – cases in which it seems we act directly on the physical world? The following example is adapted from de Muijnck (2003: 174-175): Jan, walking by the canal, on a whim kicks a stone, which describes a trajectory and falls into the canal, causing the appearance of ripples on the surface of the water. We can describe and explain what happens in two ways; first, Jan's desire (or desiring) caused him intentionally to kick the stone into the water causing ripples, and second, a manifest physical (neurophysiological) causal process, with causal antecedents stretching and branching back indefinitely in time, led to the impact on the stone, its trajectory, and the ripples. On the account I am proposing, the manifest physical causal relation constitutes, in favourable circumstances (such as that Jan is not deceived about his surroundings), the ID causal relation. But here both causes have the same effect – the appearance of ripples on the canal. Are we not, then, confronted once again by familiar Kimian objections – that either the ripples are causally overdetermined, or one of the putative causes is redundant? In the voting example, Smith's anger, clearly an ID effect of an ID cause, is constituted by his neural state (or maybe the event that is its onset), a physical effect of physical causes, and this is

what makes it plausible to argue, first, that the causal relations are distinct, and second, that the ID causal relation subsumes the physical relation and thereby has the status of being the real cause and explanation. But the formation of ripples on a surface is a purely physical phenomenon, which does not seem to constitute anything. It looks as if Jan's action, an ID event, has a physical effect if it has any effect at all. But that ID causes do not have direct, downward physical effects is a central claim of my account, and the physical causal closure principle would seem to dictate that the physical cause is the complete cause of the ripples.

De Muijnck's own suggestion (2003: 174-176), I think, fits well with the picture of ID causation I have been presenting. On his account, first, our commonsense causal explanation of what happens is correct, in that Jan's intentional action of kicking the stone causes its flight and the ripples on the canal's surface. But as none of the philosophers I have been discussing would deny, there is also a complete physical causal story. The exact natures of Jan's neural states when he sees the stone, forms the desire, and acts presumably supervene on a wide spatiotemporal base that includes, for example, the effects of behaviours learned in childhood, as well as his genetic make-up, hence that of his parents, and so on. Thus there is in principle a complete physical causal explanation for the ripples in terms of antecedent causal events, and as we know Kim would insist that this explanation preempts the ID one. De Muijnck argues, correctly I think, that in a case like this the ID event – Jan's acting on his desire – causes ID effects which simply *cease* at a certain point, say when Jan notes the ripples caused by his action. The sequence of ephemeral ID events related to Jan's action that begins when he sees the stone dissipates when Jan turns his attention to something

else. De Muijnck suggests (*ibid*: 175) that perhaps most of our mental states are like this, in that they do not lead to any further effects *via* ID causal processes. Then again, Jan's action *could* have had further ID effects, firstly in Jan himself. These could range from a momentary feeling of satisfaction to, say, his developing an abiding interest in the physics of wave formation. Or, he might have inadvertently disturbed the thoughts of a fisherman along the bank, maybe leading to his deciding to give up fishing for the day. It is possible to claim, then, that the ID and the physical causes do not, after all, have the same effect. While the physical effect is straightforwardly the trajectory of the stone and its change of location, the ID effect is most likely to be something like Jan's brief feeling of satisfaction.

Applying Baker's account to de Muijnck's example, we would say that the onset of a neural state n_1 of Jan at t_1 constituted his perceiving of the stone and forming a desire to kick it, while a further neural state n_2 at t_2 constituted his kicking the stone. On my notion of constituted causation, we could add that the causation of n_2 by n_1 plus additional factors 'ae' constituted the causation of Jan's action by his intention. If we accept Baker's claim that constituted entities subsume their constituters, Jan's action is the real cause, and explanation, of what happens. It might be objected that this does not give us the commonsense explanation of the *ripples*. Do not purely physical causes take over from the instant, t_2 , that the stone is no longer in contact with Jan's foot, so that what Jan *does*, his action, is limited to what happens until that point? A full discussion would involve an engagement with the philosophy of action beyond what is possible here. But if, as seems reasonable, we assume that it was part of Jan's deliberate intention to cause ripples on the surface by his action, it also seems reasonable to include the causing of the ripples in a

description of the ID event that comprises Jan's action. Neural state n_2 , we could then say, constitutes Jan's performing the complete action.

IV. ID Causal Relations.

(i) Marcus on Rational Explanation.

The discussion so far has focused almost entirely on what features constituted ID causation does *not* share with physical causation. It does not work by moving objects or parts of objects by force, whether mechanically, electrochemically or otherwise. It is not nomological, at least not in the way that physical causation is, and it differs from physical causation in not being governed by the other constraints listed in section I:(ii). But if ID causation is not part of the everyday physical causal nexus, and if its guiding principles and constraints are those of the space of reasons, or the constitutive ideal of rationality, should it be classed as 'causation' at all?

Eric Marcus addresses these issues in his *Rational Causation* (2012). This is his term for the distinct kind of causation that, he argues, is operative both in theoretical reasoning and in action, and approximates to what I call ID causation. Marcus uses the term 'efficient causation' for what I have been calling manifest physical, or just physical, causation; in discussing his account, I will use his terms. Like mine, his account can be understood as a rejection of premise (MC) of Harbecke's tetrad, the premise that mental events cause physical events.

In theoretical reasoning, rational causation consists in a subject's self-consciously representing to herself, and perhaps also to others, what Marcus calls the to-be-believedness of one proposition on the basis of the to-be-believedness of another. To represent a proposition as to-be-

believed just is to believe it. In a typical case this takes the form of the subject's articulating a *modus ponens* inference. In Marcus's example, Poirot reasons, "Everyone else has an alibi, so the butler did it" (6)⁶. In general, S knows that p on the basis of her knowledge that q. But S's knowledge that q is not *constitutive* of her knowledge that p. Rather, S's conscious exercise of the rational ability to infer p on the basis of q, Marcus claims, is causal – it is an exercise of rational causation.

In the case of instrumental actions, rational causation consists in the subject's representing (to herself – again, the subject's first-person perspective has an essential role) the to-be-doneness of one action on the basis of the to-be-doneness of another. Thus a subject may express what she is doing by saying "I am ϕ -ing because I am ψ -ing", or "I am ϕ -ing in order to ψ ". The cause, then, is an event in progress. In both practical and theoretical reasoning this is something the subject can say with an authority that is not based on observation or evidence (68). On Marcus's account, intentionally acting, say ϕ -ing, just *is* representing ϕ -ing as to be done – it is practical thought (79). Intending to ϕ and ϕ -ing are not fundamentally different sorts of thing, say, one mental and the other physical. Rather, action is a kind of exteriorized thought (92), whose "nature...is exhausted by its being the representation of to-be-doneness" (87). Explanations of instrumental actions, like "Dara is tasting all the desserts because she is reviewing the restaurant" (8), argues Marcus, are made true by the causal connection contained in the description of the exercise of practical rational ability, between the cause, the reviewing, and the effect, the tasting.

⁶ Page numbers are from Marcus (2012).

Marcus insists that in normal, successful cases, attributions of exercises of rational causation, both theoretical and practical, are world-citing rather than mind-citing (36-41; 107-108), thus agreeing with Stoutland and McDowell (2.III:(iii)) that in responding to reasons we respond to, and are rationally constrained by, the world, in contrast to being caused to believe or act by internal psychological states. Marcus's account thus does not require – indeed it denies – that mental states or events are identical to or realized by physical states or events that are efficient causes of actions. If John believes there is beer in the fridge because he put it there yesterday, he is representing the proposition that there is beer in the fridge as to-be-believed in the light of the to-be-believedness of his putting it there yesterday. We normally explain John's first belief by citing the worldly *fact* that he put the beer there yesterday, not by citing his *belief* that he did so. We would cite his belief in the case in which, say, we doubted or disbelieved that he had put the beer there. Like McDowell and others, Marcus (2-3) favours a disjunctive account (1.III:(ii)), in contrast to the position that he calls "psychologism" and that McDowell calls the highest common factor view (1.III:(i)) – the view that belief acquisitions and actions are (physically) caused by inner psychological states whose natures are independent of whether or not they accurately represent worldly reality. The opposing view is that such acquisitions and actions are rationally caused by external situations themselves in successful cases, while unsuccessful cases in which reasoning is based on error or illusion are parasitic on the successful ones. The latter are the cases in which we withhold world-citing explanations and instead cite agents' psychological states.

I will not discuss Marcus's detailed analysis of the logic and grammar of rational explanation of belief and action (14-115), from which he derives

the structure of rational causation, which I have just summarized. My interest in rational causation is more concerned with its status *vis a vis* efficient causation. There seems no reason to dispute that this analysis gives as accurate a picture as any of the inferential relations that form the foundations of our theoretical and practical reasoning. But as Marcus acknowledges (146), if taken solely as an analysis of these kinds of explanation, his account would be consistent with, say, a Davidsonian account of reason causation and mental causation (see 2:I:(ii), (iii); II:(iii)). The Davidsonian could agree that the analysis accurately captures our explanatory practices and requirements within the conceptually irreducible framework of the constitutive ideal of rationality (Davidson ME: 223). Yet she would maintain that the actual causality involved is (in Marcus's term) efficient, governed by the PNCC.

(ii) Rational Causation and the Physical.

Marcus's broader, and harder, task, then, is to show that these rational explanations designate, or reflect, real causal connections, which "consist in the exercise of inferential theoretical and practical rational abilities" (167). So, for example, if John is entering the kitchen in order to get beer from the fridge, Marcus's aim is to show that getting beer is not just the explanation but also the *cause* – the rational cause - of his entering the kitchen. He argues (155-156) that Davidson (ARC, ME; 2:I:(iii); II:(iii)(iv) above) unjustifiably draws the conclusions I have just attributed to the Davidsonian from the evident fact that causation belongs to the realm of the represented while causal explanation belongs to that of the representing (*ibid*). Put in that way, Marcus argues, the distinction is innocuous, but, he continues, it is not innocuous to infer, as Davidson does, from the fact that explanantia and explananda are

representational items the idea that the facts that they pick out “are not real in the way that items in other categories – e.g., events – are real” (*ibid*). The illicit move is that from acceptance that propositions are not causes and effects to the view that they do not even *designate* causes and effects (156). To put the point in Kimian terms, Marcus is taking issue with the causal/explanatory exclusion principle, that a given event can have only one cause and explanation. What is objectionable, then, is Davidson’s ontological privileging of the scientific, or physical, to the extent that it is only events – all of which, recall, are physical and bound by strict laws on Davidson’s account, even if some also have mental descriptions – that are the real relata of causation. This, as we saw (2:II:(iv)), is also McDowell’s objection to Davidson’s account.

What, then, is the relation between rational and physical causation on Marcus’s account? Practical rational causation, in particular, seems *prima facie* to be intimately bound up with the physical world. It might seem inescapable that John’s changing his physical location by entering the kitchen is a physical effect of its rational cause, the event in progress of his getting a beer. But Marcus writes:

People often do things as a result of what they want and think. And what they do often involves their moving stuff around... Hence the nearly inescapable conclusion that mental states and events affect stuff... A man may push a sofa down a hallway because he wants to move it into the living room. (But) his moving it is... a rational consequence [of his desire]. And the effects of rational causation are not physical events at all. They are...*thoughts* (230).

In chapter 1:II:(i) I mentioned the parallels between the idea in philosophy of action that “mental states and events affect stuff” and the myth of the given in the philosophy of perception. A non-conceptual sensation, Sellars and McDowell argue, is not the kind of thing that could justify a belief. By analogy, I suggest, an intention or desire or intention is not the kind of thing that could “affect stuff“. Marcus is firm in his rejection of the telekinetic, or “spiritual pushing” view of mental causation – the view that “mental events or states... would be causal powers that fill in the gaps in physical causal chains” (235). Such a scenario would obviously be a breach of physical causal closure, and the view, according to Marcus, “reflects obliviousness to the sui generis character of rational causation”.

And yet the sofa undoubtedly does move, and according to the causal closure principle, its movements are fully accounted for by physical, or efficient, causes. To avoid the fate to which Kim consigns dual explananda, then (III:(i) above), it seems that rational causes and efficient causes must be seen as doing their work quite independently of each other, that is, as not operating in the same causal nexus. As Marcus writes, “the threat [from the causal closure objection] seems to survive so long as we think of mental causation as culminating in physical events” (254). But given that physical events do in fact occur, and accepting, with Marcus, that they are not causally overdetermined, we must ask how, if at all, are the two kinds of causation related to each other? We are discussing practical rational causation, but the same question arises in theoretical examples like Poirot’s deduction, even though in this case the physical events are confined within Poirot’s nervous system.

Marcus says little about the relation between physical and rational causation, but he does accept that rational activity must be under *some* kind of physical constraint – that something physical must accompany an instance of rational causation - but it seems that this must be a *constitutive*, rather than a causal, constraint, such as a relation of global supervenience of the mental on the physical. He writes:

If someone wants [on that basis] to describe my view as a form of physicalism, fine. Physicalism is then consistent with a view of the mind according to which it is a field of non-physical and non-physically realized events, states, and facts and is governed by a kind of non-physical causation (223).

Unlike Baker or Jaworski, then, Marcus recognizes that to be autonomous and not subject to the objections of physical causal closure arguments, rational causation must occur in a nexus – a “field” – that has no causal connection with the world of physical or efficient causation. As with my concept of ID causation, rational causes operate according to their own rational constraints.

As I said above, what I have called manifest physical causation is what Marcus calls efficient causation, and I have argued in 4:II:(v) and elsewhere that the idea that mental content could partake in this kind of causation is fundamentally flawed. Many of the examples of constituted ID causal relations that I have given and discussed have the same grammatical form (*i.e.*, “*c* caused *e*”) as the expressions of their constituting physical causal relations, and hence might look like instances of efficient causation at a higher level, which is exactly what Marcus denies that rational causes are (233). But the principles that

determine the relations are, or can be traced back to, the same principles that govern the self-conscious manifestations of rational ability that define Marcus's concept of rational causation, that is, inferential, normative relations. Just as the causal relation in 'the window shattered because the baseball hit it' is an elliptical expression of an instance of physical law, so the one in 'Smith became angry because Jones voted against him' can be interpreted as picking out the exercise of an inferential ability, such as Smith's inferring, from the to-be-believedness of Jones's having voted against him, the to-be-believedness of Jones having betrayed him in a manner to which angry resentment is an appropriate response.

The same kind of transformations can be applied to the mixed physical/ID and the pure ID but non-mental causal descriptions I used as examples in 4:I:(v) and above. For example, in

He purposely threw the ball that smashed the window,

'he purposely threw the ball' is analyzable in accord with Marcus's formulation of practical rational causation; from his first person perspective, he threw the ball in order to (because he was) ψ (-ing) (exacting revenge, playing cricket...). But, as with de Muijnck's example, we can say that his purposely throwing the ball was constituted, in favourable circumstances, by lower-level events, which are what caused the smashing of the window, another lower-level, manifest physical event.

We can, however, identify a difference that emerges between mental and non-mental ID causes. Like McDowell (1:II:(i)) and Baker (3:I), Marcus

((i) above) emphasizes the essential role of the first person perspective in agency, whether theoretical or practical – in inference or action, the subject *represents* to-be believedness or to-be-doneness to herself. In the purely ID causal description

Excessive sub-prime mortgage lending caused the recession

there is no explicit appeal to rational causation in this first-person sense. But again, the proposition that it expresses implies that, on both the cause and the effect sides of this efficient causal relation, many individuals self-consciously exercised both theoretical and practical rational causation on many occasions, and I showed in II:(iii) above how the notion of constituted causation is applicable to these kinds of causal statements. ID causal statements like this, then, are secondary, or derivative, from what we might call primary statements of rational causation – those involving an individual’s exercise of the capacity.

It looks, then, as if the notion of constituted causation and Marcus’s concept of rational causation are quite compatible. As I noted, Marcus himself says very little about how rational and efficient causes are related, merely noting that his account is compatible with a kind of constitutive (*i.e.*, non-causal) dependency of the rational on the physical. I suggest that my development of Baker’s constitution account supplies this constitutive dependency through constituted causation. Instances of rational causation like Poirot’s believing that the butler did it on the basis that everyone else has an alibi (6), or Dara’s tasting all the desserts in order to review the restaurant (8), consist of the self-conscious representation of causal connections from the to-be believedness or to-be-doneness of one thing to the to-be-believedness or to-be-doneness of

another, according to Marcus. If we accept Marcus's claim that these are indeed causal, they are representative of a rational, normative form of causality that has "no echo in physical theory" (Davidson PP: 231). It is a causality that is bound by the constraints of the space of reasons rather than the realm of law. Yet this is compatible, I have been arguing, with these causal relations' being constituted, in favourable circumstances, by causal relations that are constrained only by physical laws.

We can, then, equate those cases of constituted ID causation that are mental – those that I have suggested are the primary form of ID causation - with instances of theoretical or practical rational inference. That is to say, mental ID causes are a person's reasons; reasons for belief or for action that are consciously entertained by an individual. This primary form of ID causation is then revealed as corresponding precisely to the rational constraint by the world that is at the core of McDowell's account of mind and nature. On the McDowellian account (1:II:(i)), we can include in the to-be-believed category the passive actualization of conceptual capacities that occurs in experience, which, through integration into our conceptual content, become the causes of other beliefs. Smith is passively presented with the content *that* Jones has voted against him, and this fact's to-be-believedness is the basis of his active exercise of the judgment with the same content, from which he infers further judgments. And, as in the example of the man moving the sofa, the intention and the action are both part of the same "representation of to-be-doneness" (87). The network of ID causal relations that, if this is right, permeates seamlessly through the continuum of our mental content and the world of facts, structures the unbounded conceptual of McDowell's account.

V. Constituted Causation and the Space of Reasons.

(i) *Direct Causation by Content.*

On the constitution view, ID causes are directly causally efficacious *qua* the rational causal relations that they are. So on this account, unlike, say, Dretske's (III:(i) above), there is a convergence of cause and causal explanation, in that we do not account for causal efficacy by reference to causal-nomological mechanisms and then have to resort to a separate explanation, such as that a causal mechanism has become hooked up in a certain way, to account for the explanatory role of content. Since semantic content, rationality, and normativity are integral to ID causation, it seems that to articulate an ID causal relation is *thereby* to provide a rational explanation.

Also, on this model of ID causation and explanation, both mental and 'secondary' ID, but non-mental, instances of causation, such as, say, an economic recession's causing a fall in house prices, are seamlessly integrated within the ID causal nexus. As I argued above, non-mental instances of ID causation are ultimately analyzable into mental instances, and it seems to be the case that both mental and non-mental ID causation belong to the same causal nexus, in the sense that mental events can cause non-mental ID events, and *vice versa*. There is no problem, for example, with the proposition that a fall in house prices causes Mary to decide not to sell her house. McDowell's claim (MW:27) that there is no ontological gap between content and worldly facts, so that our minds are open to the world, fits this situation exactly, and the idea blends naturally with Stoutland's (1998) views on the externality of content. According to Stoutland and Marcus, Mary's decision here is

typically a response to an external situation, not to her belief that prices are falling (her response may be to her belief if, say, (a) the belief is false or (b) prices are falling, but her belief with that content has been acquired in some deviant way, such as from a fortune teller). Since the causality under discussion is ID, with ID causal relata, we can call Mary's response, of deciding not to sell, a *causal* effect. ID causation, then, is causation of content *by* content, and I follow Marcus (IV:(i), (ii) above) in designating this a species of causation because it is a non-constitutive way of explaining how things come to be a certain way. Content is not confined to our heads, but pervades the space of reasons. It is through the sharing of content that our minds and the world interpenetrate, in McDowell's phrase (STEIS: 241; see chapter 1).

In Stoutland's example (2:III:(iv)), a driver's encountering a stop-sign, seeing it *as* a stop-sign and not as just a piece of metal, and stopping accordingly, comprises an ID causal relation. In Marcus's terms, this is the exercise, by the driver, of a mixture of theoretical and practical reason; presented with the to-be-believedness of the sign, he infers the to-be-doneness of stopping. This, I suggest, is the sense in which McDowell and Stoutland concur that reasons are causes, but as I have presented constituted ID causation, reasons do not exert their effects in the guise, so to speak, of physical, causes in the way that Davidson claims – *via* token identity – but directly, as ID causes, with ID effects. Physical causes, although omnipresent, do not figure in the ID description or explanation.

(ii) *Constitution, Causation, and the Space of Reasons.*

I began my discussion of Baker with the intention of seeing whether the constitution account could furnish the resources to clarify and resolve the inconsistency in McDowell's treatment of causality I identified in chapter 2:II:(v). I have found it necessary to modify and develop Baker's account, but having done so I think it is clear that it can provide what is needed. There turns out to be a close correspondence between the patterns of relations in ID causation and the structure of the space of reasons as McDowell envisages it; "The space of reasons is the space within which thought moves, and its topography is that of the rational interconnections between conceptual contents; we might equally speak of the space of concepts" (KI: 408). There are several other points of correspondence. The space of reasons is irreducible to the realm of law in just the same way that, as we have seen, constituted entities, and ID entities in particular, are irreducible to their constituters. Also, the notion of ID causes, comprising both mental and other non-mental causes which freely interact with each other, seems to provide just the right basis for the externality of content and the interpenetration of mind and world that is central to McDowell's account (chapter 1).

Causal relations within the space of reasons are constrained by normative notions such as justification, purpose, and value, as these are exercised by a self-conscious subject, a possessor of or participant in second nature (1:I:(ii)), and are bound by principles of holistic consistency and coherence. The relata of constituted ID causation are, as McDowell would put it, thinkables. The *exercise* of ID causation consists in a subject's drawing of inferential connections among thoughts (thinkables – not necessarily true – that are being entertained), but on McDowell's account those thinkables that are true, and the relations amongst them, are "there anyway" (MW: 91), whether or not

they feature, or have ever featured, as part of any subject's content. Together, the relata and the ID causal relations linking them form a rational, meaningful pattern.

The picture of ID causation as distinct from physical causation and the main elements of McDowell's anti-Cartesian account (1:III:(iv)), such as content externalism, unmediated interpenetration of mind and world, and the disjunctive theory of perception are, I think, mutually supportive. The Cartesian idea that only the contents of an inner realm are available to consciousness seems to follow directly from the fact, obvious at least since the scientific revolution, that our only *physical* connection with the external world is *via* our various sensory organs. Our understanding of the general input-integration-output structure of nervous systems, and of the tendency of the complexity of the central, integrative component to be correlated with that of organisms themselves, encourages us to see our intentionality as merely the upper limit of this continuous line of development. If content is thus thought of as in some way a property of the physical structures or properties of the nervous system, and thus subject to the same constraints, the Cartesian picture can easily seem inescapable – how could we be in touch with a reality beyond the filter of our sensory apparatus? But if content, or intentionality, is not bound by physical constraints, then the claim that when we are not deceived we are in direct contact with reality appears much more plausible. Further, one might suggest that veridical perceptual content is *constituted* by a neural state in the favourable circumstances of its being a true reflection of the worldly situation, while the favourable circumstances for, say, a hallucination, involve the same neural state's being constituted in some deviant way.

It seems to me that the parallels between McDowell's view of reality and the picture of what we might call a constituted ID world that we get from our development of Baker's account are too striking to be ignored. McDowell, as we have seen, denies that there are philosophical questions to be addressed about the relation between the realm of law and the space of reasons, insisting only that the latter is both irreducible to and invisible from the standpoint of the former, and that, despite the spectacular results achieved in the last 400 years by the deliberate adoption of the realm of law standpoint, the picture of reality it provides is only partial. How "events that manifest freedom are related to events that are intelligible by the methods of natural science" is "a good question", but one for "scientific investigation of the machinery of mindedness" rather than philosophy (RGM: 239). The constitution account, however, suggests a basis for a relation between the two. The notion of ID causation appears to correspond precisely to the rational constraint by the world that is at the core of McDowell's account of mind and nature. The network of ID causal relations that, if this is right, permeates seamlessly through the continuum of our mental content and the world of facts, structures the unbounded conceptual of McDowell's account. In a McDowellian spirit, then, I suggest that the difference between explaining a causal event in the terms of the ID causal nexus and explaining it in the terms of the physical causal nexus is the difference between an explanation from the subjective viewpoint that is only attainable from within the space of reasons, in which notions such as normativity, meaning, and purpose figure, and a realm of law kind of explanation, framed entirely in terms of "how things generally tend to happen" (FAM: 328).

According to Michael Williams, on McDowell's account of the two logical spaces, "Rational justification and subsumption under law constitute distinct modes of intelligibility, neither being reducible to the other" (Williams 2006: 305). In view of what the constitution account reveals, I think this is not entirely correct. Certainly, the space of reasons mode is irreducible to that of the realm of law, but the relation in the other direction is more subtle than irreducibility. We should not fall into the trap of thinking that while, say, Jan's kicking of the stone as an instance of ID causation clearly takes place in the space of reasons, the lower-level causal process does not. Recall (1:II:(ii), (iii)) that on McDowell's account nothing is outside the conceptual. The space of reasons/ realm of laws distinction is an epistemological one that distinguishes two different *ways* of thinking about the world – and, of course, all thinking belongs in the space of reasons. When we explain an event in terms of manifest physical causal process, we make use of common sense knowledge of how inanimate objects behave, supplemented to a variable extent by knowledge from special sciences – mechanics, neurophysiology, and so on. But of course "*depictions* of nature are linked by relations of justification" (MW: 70n1, emphasis added). "Subsumption under law", as a form of explanation, belongs in the space of reasons just as does rational justification. It is just that "there are no such linkages in what is depicted" (*ibid*) – the *content* of realm of law explanations is non-normative. It might be better to say of realm of law explanations, in cases in which we can also give an ID explanation, that they are incomplete. If so, the physical explanation of Jan's action does not exclude the ID one. Rather, the physical explanation is what we are left with when the ID explanation is not applicable or is discounted - in cases, that is, when the favourable circumstances that bring the explanation into the space of reasons are

absent or deliberately abstracted from. If this is right, the presence of favourable circumstances is just what makes the difference between Jan's deliberate kicking of the stone, Jane's raising her arms to be searched, and all other such cases, on the one hand, and causal processes that are entirely explicable in realm of law terms on the other.

(iii) Subsumption, Priority, and Physical Causal Closure.

Perhaps the most vulnerable part of Baker's account, which I have helped myself to in developing the idea of constituted causation (**II**:(iii) above), is her claim that the primary kind property of a constituted entity subsumes or encompasses that of its constituter, and that the latter has no independent existence (PB:33, 46; MEL:166). As I discussed (3:V:(i)), one way of interpreting this is as the idea that what is constituted is prior to what does the constituting, in such a way that facts about the former rather than the latter count as the real explanation of an event or state of affairs. The motivation for the subsumption claim appears to be the prior claim that constituted objects or properties are, through their constitution, endowed with irreducible, novel causal powers. For Baker the priority is ontological, but as I discuss in (iv) below I am uncommitted as to whether the causal nexus of the space of reasons should be thought of as an ontological 'realm', rather than as an epistemic space, or an 'aspect' of the world (Hornsby ACE: 150).

In any case, my constituted causation account stands or falls on the viability of this notion of priority. If ID causes do not subsume lower-level ones, the account is pointless. Proposing that there are ID causes would be "an empty verbal ploy" in Kim's sense (PSNE: 62; 3:V:(ii) above) - they would suffer the same fate as his 'supervenient causes',

being no more than causal descriptions that do not designate real causes, while the actual causation occurs at the lower level. It is comparatively easy, I think, to accept that once Piece, in favourable circumstances, constitutes *David* (3:II:(ii)), what we have is *really* a statue rather than a piece of marble, perhaps because *David*'s most salient property is its place in the Renaissance artistic canon. But in a case involving causation, such as Jan's kicking the stone into the canal, even if we accept that the ID causal process does not belong in the same causal nexus as the manifest physical one, it seems somehow counterintuitive to think of the former as the real cause and explanation of the stone's flight in its physical detail. This may be because whatever happens at the intentional level, we have a strong intuition that the physical causal process, occurring in accordance with well-established and immutable laws, was pre-determined and inevitable – an intuition that, as I argued in 2:II:(v), arises and belongs *within* the manifest image. It would seem natural to explain how Jan was able to kick the stone by referring to the physical mechanisms, but if asked to explain why a particular sequence of neurophysiological events was instantiated in Jan between t1, when he formed the intention, and t2, when the stone left his foot (III:(iii) above), I think we would tend to point to the relevant physical causal history up until t1, rather than to the fact that the events constituted his action of kicking the stone.

As I suggested above (II:(iii)), if we extend Baker's general theory of constitution as unity without identity to my constituted causation account, we can claim that Jan's action *just is* (again, the 'is' of constitution) the neurophysiological and mechanical causal chain of events between t1 and t3, when the ripples appear on the water, given, of course, that the stone-kicking-favourable circumstances are present –

that Jan is a possessor of a first-person perspective, that he is not deluded in some way about his situation, and so on. In taking Jan's intentional action to be the real cause of what happened, then, we are not presenting this as an *alternative* to the physical cause. We are, rather, enriching the account by adding the extrinsic features that transform it. When the event is described from Jan's point of view or that of an observer – from the standpoint, that is, of the space of reasons – it becomes an ID event, an instance of ID causation.

This appeal to the relation of unity without identity may still, however, appear too weak to dispel the intuition that the real causation happens at the lower level, and that this is thus the locus of the real explanation. In our experience or in recorded history (setting aside religious texts), we cannot point to a single universally authenticated instance in which a manifest physical event that was independent of human agency had anything other than a manifest physical cause. This, together with our understanding of the physical and physiological basis of agency, leads, or so I have been arguing, to the inescapable conclusion that our own interaction with the physical world is part of this same all-encompassing and predetermined manifest physical causal nexus. So if the relation of constituted causation ensures that, say, Jane's voluntary action of raising her arms *just is*, given favourable circumstances, a part of this nexus, how can we escape the conclusion that it, too, is predetermined, and that Jane's conviction that she is voluntarily complying with airport security regulations is illusory?

I have been arguing throughout that ID causation operates under quite different constraints from manifest physical causation. It is subject to the constraints of the space of reasons and not those of manifest physical

causal closure, and if it can be said to be under nomological constraint, the laws governing it are not physical laws but those of a different, perhaps psychological or semantic, jurisdiction. The problem is to explain how two sets of causal relations that are united by the ‘is’ of constitution could nevertheless form patterns that are quite distinct and do not map on to each other at all. The answer I suggested in **II:(iii)** above was that properties like being bound by the physical causal closure principle or physical nomological sufficiency are modal and hence that they cannot be shared derivatively. But this may seem inadequate as a support for an account of mental and ID causation as subject only to rational constraints, given the relentless playing out of physical necessity at the lower causal level, which does not seem compatible with Jane’s, and our, conviction that in willing to raise her arms she made a conscious choice and could have chosen otherwise.

Here I think we must reject, or at least modify, one component of Baker’s account. Baker plays down the mental/physical, or ID/physical distinction, thinking of it as just one of the vast number of constituted hierarchical ontological distinctions in nature; “(T)here is not just one big divide in nature between two disparate realms – mental and physical” (MEL: 177). This is reflected in the final clause of her definition (C) of material constitution (**3:II:(i)**); where x constitutes y ,

- (f) If x is of one basic kind of stuff, then y is of the same basic kind of stuff.

The stipulation is independent of the remainder of the definition, and underpins Baker’s view that all properties are ultimately physical (**3:IV:(ii)**). However, if I am right that physical and ID causes operate in

distinct nexuses, it seems that their relata cannot be “of the same basic kind of stuff”. As I suggested above, I am not sure if ID causal relata should be regarded as *any* kind of stuff. *David* and *Piece*, for example, share properties; *David* derivatively shares *Piece*’s molecular structure, and *Piece* similarly shares *David*’s value. But, barring perhaps spatiotemporal location, it is difficult to think of any properties that are common to Jane’s willing causing her arm-raising and the physical causal relation that constitutes it. And this, it appears, is intimately connected to the fact that (as I argue) ID causal relations “make no difference in the [manifest physical] causal order of the world” (de Caro and Voltolini 2010: 78).

(iv) *Ontology*.

If ID causes and other ID phenomena are not part of the physical causal order, what are they? This is the ontological question that has been looming throughout – that Kim (EEPNC: 133; III:(i)), for example, raises when he suggests that if higher causes are not physical we have an overt dualism. McDowell, on the other hand (RGM: 238), claims that there is a *distinction*, not a dualism, between “two kinds of happenings in nature: those that are subsumable under natural law, and those that are not”. But this distinction between kinds of event “does not imply that the composition of human beings includes something just like a kind of stuff except that it is not material” (*ibid*).

I think the account I have given of mental and ID causation is in line with McDowell’s view as expressed here. What I hope to have done is just to show that holding this view consistently requires that we regard these forms of causation as occupying a distinct causal nexus. So the

relata of ID causation are not a “kind of stuff” at all. We seem drawn, once again, to Davidson’s view (PEA: 114; 2:III:(i)) that the mental is not an ontological but a conceptual capacity, and Hornsby’s (ACE: 150; 2:III:(iv)) that the mental (or intentional, including our first-person perspective) is an aspect of reality rather than a portion of it. Jane’s self-conscious raising of her arms, in the richness of its context, is an aspect of the world visible only from within the space of reasons.

If we take Baker’s claim that constituted entities subsume their constituters seriously, then the claim that to understand something – say Jane’s performance of her deliberate, rational action – as intentional is to have in view an aspect of reality that is only visible from the space of reasons perspective need not, I think, imply that reality viewed in this way is somehow less real than reality viewed from the physical standpoint. To think that this view *is* implied, on the interpretation I am following, would be to remain in thrall to a picture that McDowell argues is incoherent – a ‘sideways-on’ picture of a “reality outside a boundary enclosing the conceptual” (MW: 82; 1:I:(ii)). As we saw (1:II:(iv)), on McDowell’s modification of the Kantian picture, what is real is the conceptually structured world that we experience, whose existence is a precondition for our having experiential, and hence any, content at all.

All we can be sure of, it seems, is that in our corner of the cosmos creatures have somehow developed the capacity (or, at least, have developed in such a way that they interpret themselves as having the capacity) to be attuned to “the world’s own language” (1:II:(ii)) - not only to the physical nomological patterns governing inanimate nature, nor to the biofunctional patterns of natural selection, but to the rational,

meaningful, and ethical patterns of their own distinctive form of life. That these various kinds of patterns, according to the standards of our rationality, are generally remarkably consistent, and that truths about them appear to be objective in the recognition-independent sense that McDowell, Baker, and others share, suggest that they are indeed *real* aspects of the world (see Rescher 2009). I am inclined towards McDowell's view (recently reaffirmed; CCP: 144n18) that the world is conceptually structured, made up of facts, and so apt for our understanding. And self-evidently, the world has contained the possibility of there coming to exist creatures who would 'resonate to meaning', as McDowell would say, from the very beginning. However it came about, with the appearance of such creatures on the scene, one might say, the favourable circumstances were in place for the realm of law to constitute the space of reasons.

BIBLIOGRAPHY

Titles of works by Baker, Davidson, Hornsby, Kim, McDowell and Sellars are abbreviated and listed separately, below.

Works by Lynne Rudder Baker, with date of first publication.

AA. ‘Attitudes in Action. A Causal Account’, in Leclerc, A, de Querioz, G, and Wrigley, M, eds, *Proceedings of the Third International Colloquium in Philosophy of Mind*. Campinas, Brazil: Center for Logic, Epistemology and History of Science, State University of Campinas, 2002: 47-78

ABBS. ‘Are Beliefs Brain States?’ in A Meijers, ed, *Explaining Beliefs: Lynne Rudder Baker and her Critics*. Stanford: CSLI Publications, 2001: 17 – 38

ARP. ‘Against Reductive Physicalism: Causation Without Reduction’, workshop presentation 2006.

www-unix.oit.umass.edu/~lrb/files/bak06agaM.pdf

Accessed November 2010.

EA. *Explaining Attitudes*. Cambridge: Cambridge University Press, 1995

MEL. *The Metaphysics of Everyday Life*. Cambridge: Cambridge University Press, 2007.

MMC. 'Metaphysics and Mental Causation', in J. Heil and A. Mele, eds, *Mental Causation*. Oxford: Oxford University Press, 1993: 75 – 95.

MSVK. 'Must Science Validate All Knowledge?', in A. J. Sanford, ed, *The Nature and Limits of Human Understanding*. London: T&T Clark, 2003: 163-208.

NFPP. *Naturalism and the First-Person Perspective*. Oxford: Oxford University Press, 2013.

NRM. 'Nonreductive Materialism', in B. McLaughlin, A. Beckermann, and S. Walter, eds, *The Oxford Handbook of Philosophy of Mind*. Oxford: Clarendon, 2009: 109 – 127.

OMTU. 'On Making Things Up: Constitution and Its Critics', *Philosophical Topics* 30: 1: 31 – 51, 2002

PB. *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press, 2000.

SB. *Saving Belief*. Princeton, N.J.: Princeton University Press, 1987.

Works by Donald Davidson with date of first publication.

ARC. 'Actions, Reasons, and Causes', in EAA: 3-19. 1963

CR. 'Causal Relations', in EAA: 149-162: 1967

CTTK. 'A Coherence Theory of Truth and Knowledge', in SIO: 137-158. 1983

EAA. *Essays on Actions and Events*. Oxford: Clarendon, 1980.

ITI. *Inquiries into Truth and Interpretation*. Oxford, Clarendon, 1984.

LC. 'Laws and Cause', in D. Davidson, *Truth, Language, and History*. Oxford: Clarendon, 2005: 201 – 219. 1995

ME. 'Mental Events', in EAA: 207-224. 1970

PEA. 'Problems in the Explanation of Action', in D. Davidson, *Problems of Rationality*. Oxford: Clarendon, 2004: 101-116. 1987

PP. 'Psychology as Philosophy', in EAA: 229-239. 1974

SIO. *Subjective, Intersubjective, Objective*. Oxford: Clarendon, 2001.

TC. 'Thinking Causes', in J. Heil and A. Mele, eds, *Mental Causation*. Oxford: Oxford University Press, 1993: 3-17.

VICS. 'On the Very Idea of a Conceptual Scheme', in ITI: 183-198. 1974

Works by Jennifer Hornsby with date of first publication.

ACE. 'Agency and Causal Explanation', in SMH: 129-153. 1993

CPCP. 'Causation in Intuitive Physics and in Commonsense Psychology', in SMH: 185-194. 1992

SMH. *Simple Mindedness: In Defense of Naïve Naturalism in the Philosophy of Mind*. Cambridge, Mass.: Harvard University Press, 1997.

WPEME. 'Which Physical Events are Mental Events?', in SMH: 63-77. 1981.

Works by Jaegwon Kim with date of first publication.

BCD. 'Blocking Causal Drainage and Other Maintenance Chores with Mental Causation', *Philosophy and Phenomenological Research*, LXVII 1 (2003): 151-176.

CMC. 'Causation and Mental Causation', in EMM: 243-262. 2007

CS. 'Concepts of Supervenience', in SMK: 53-98. 1984

DREB. 'Dretske on How Reasons Explain Behavior', in SMK: 285-309. 1991

EEPMC. 'Explanatory Exclusion and the Problem of Mental Causation', in C. Macdonald and G. Macdonald, eds, *Philosophy of Psychology*. Cambridge, Mass.: Blackwell, 1995: 121-141.

EKMD. 'Explanatory Knowledge and Metaphysical Dependence', in EMM: 167-186. 1994

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

EPE. 'Events as Property Exemplifications', in SMK: 33-52. 1976

ESC. 'Epiphenomenal and Supervenient Causation', in SMK: 92-108. 1984.

LWMC. 'The Layered World: Metaphysical Considerations', in EMM: 41-65. 2002

MC. 'Mental Causation', in B. McLaughlin, A. Beckermann, and S. Walter, eds, *The Oxford Handbook of Philosophy of Mind*. Oxford: Clarendon, 2009: 29 – 42.

MNRM. 'The Myth of Nonreductive Materialism', in R. Warner and T. Szubka, eds, *The Mind-Body Problem*. Oxford: Blackwell, 1994: 242 – 260

MPEE. 'Mechanism, Purpose, and Explanatory Exclusion', in SMK: 237-264. 1989

MPW. *Mind in a Physical World*. Cambridge, Mass.: MIT Press, 1998.

MRMR. 'Multiple Realization and the Metaphysics of Reduction', in SMK: 309-335. 1992

MSDC. 'Making Sense of Downward Causation', in Andersen et al, eds, 2000: 305-321.

NRTMC. ‘The Non-Reductivist’s Troubles with Mental Causation’, in SMK: 336-358. 1993.

PM. *Philosophy of Mind*. Boulder, Colorado: Westview, 1996.

PPL. ‘Psychophysical Laws’, in SMK: 194-215. 1985

PSNE. *Physicalism, or Something Near Enough*. Princeton: Princeton University Press, 2005.

SG. ‘Supervenience’, in S. Guttenplan, ed, *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1994: 575–583

SMK. *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993.

Works by John McDowell, with date of first publication.

AOFB. ‘Aesthetic Value, Objectivity, and the Fabric of the World’, in MVR: 112 – 130. 1983

CCP. ‘Conceptual Capacities in Perception’, in HWV: 127-144. 2009

CDK. ‘Criteria, Defeasibility, and Knowledge’, in MKR: 369-394. 1982

CPE. ‘The Content of Perceptual Experience’, in MVR: 341-358. 1994.

DRS. ‘*De Re Senses*’, in MKR: 214-227. 1984

EF. 'Evans's Frege', in EI: 163-185. 2005

EI. *The Engaged Intellect: Philosophical Essays*. Cambridge, Mass.: Harvard University Press, 2009.

ENN. 'Avoiding the Myth of the Given' and 'Responses', in J. Lindgaard, ed, *Experience, Norm, and Nature*. Oxford, Blackwell, 2008.

FAM: 'Functionalism and Anomalous Monism', in MVR: 325-340. 1985

HWV. *Having the World in View: Essays on Kant, Hegel, and Sellars*. Cambridge, Mass.: Harvard University Press, 2009.

IIW. 'Intentionality and Interiority in Wittgenstein', in MVR: 297-321. 1991

KI. 'Knowledge and the Internal', in MKR: 395-413. 1995

MKR. *Meaning, Knowledge and Reality*. Cambridge, Mass.: Harvard University Press, 1998.

MVR. *Mind, Value and Reality*. Cambridge, Mass.: Harvard University Press, 1998.

MW. *Mind and World*. Cambridge, Mass.: Harvard University Press, 1994.

- NPM. 'Naturalism in the Philosophy of Mind', in EI: 257-275. 1999
- OSPLA. 'One Strand in the Private Language Argument', in MVR: 279-296. 1989
- PMM. 'Putnam on Mind and Meaning', in MKR: 275-291. 1992
- PMW. 'Précis of Mind and World' and 'Reply to Commentators', *Philosophy and Phenomenological Research*, LVIII, 365 – 368: 403 – 431. 1998
- PPD. 'Physicalism and Primitive Denotation: Field on Tarski', in MKR: 132-154. 1978
- RAB. 'Response to Akeel Bilgrami', in C. Macdonald and G. Macdonald, eds, *McDowell and his Critics*. Oxford: Blackwell, 2006: 67 – 72
- RGM. 'Response to Graham Macdonald', in C. Macdonald and G. Macdonald (eds), *McDowell and his Critics*. Oxford: Blackwell, 2006: 235 – 239
- RN. 'Experiencing the World', and 'Responses' in M. Willaschek, ed, *John McDowell: Reason and Nature*. Munster: Lit Verlag: 1999.
- RS. 'Responses', in N. Smith, ed: *Reading McDowell: On Mind and World*. London: Routledge, 2002: 269-305.

SCDE. 'Scheme-Content Dualism and Empiricism', in Kotatko, Pagin, and Segal, eds, *Interpreting Davidson*. Stanford: CSLI, 2001: 143-154

SSPR. 'On the Sense and Reference of a Proper Name', in MKR: 171-198. 1977

STE. 'Sellars's Transcendental Empiricism' in J. Nida-Rumelin, ed, *Rationality, Realism, Revision*. Berlin: W. de Gruyter, 2000: 42-51.

STEIS. 'Singular Thought and the Extent of Inner Space', in MKR: 228-259. 1986

TRO. 'Towards Rehabilitating Objectivity', in EI: 204-224. 2000

TSN. 'Two Sorts of Naturalism', in MVR: 167-197. 1996

WFR. 'Wittgenstein on Following a Rule', in MVR: 221-262. 1984

WL. 'Having the World in View: Sellars, Kant, and Intentionality' (The Woodbridge Lectures), in HWV: 3-65. 1998

Works by Wilfrid Sellars with date of first publication.

EPM. 'Empiricism and the Philosophy of Mind', in SPR: 127-196. 1956.

PSIM. 'Philosophy and the Scientific Image of Man', in SPR: 1-41. 1962

SMS. *Science and Metaphysics*. Atascadero, Cal.: Ridgeview, 1992.
1967

SPR. *Science, Perception and Reality*. Atascadero, Cal.: Ridgeview, 1991. 1963

Alexander, S., *Space, Time, and Deity*. London: Macmillan, 1920.

Allison, H. E., 'We Can Only Act Under the Idea of Freedom', *Proceedings and Addresses of the American Philosophical Association* 71:2 (1997); 39 –50.

Andersen, P., Emmeche, C., Finnemann, N., and Christiansen, P., eds, *Downward Causation: Minds, Bodies and Matter*. Aarhus: Aarhus University Press, 2000.

Anscombe, G.E.M., 'Causality and Determination', in Sosa and Toley, eds, 1993: 88 – 104 (first published 1971).

Armstrong, D., 'Going Through the Open Door Again: Counterfactual versus Singularist Theories of Causation', in Collins, Hall, & Paul, eds, 2004: 445-457.

Bakhurst, D., *The Formation of Reason*. Chichester: Wiley-Blackwell, 2011.

Barber, M.D., 'Phenomenological *Wissenschaftslehre* and John McDowell's Quietism', in T. Nenon and P. Blosser, eds, *Advancing Phenomenology*. London: Springer, 2010: 433-454.

Beaney, M., Penco, C., and Vignolo, M., *Explaining the Mental: Naturalist and Non-Naturalist Approaches to Mental Acts and Processes*. Newcastle-upon-Tyne: Cambridge Scholars Publishing, 2007.

Bechtel, W., *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. New York: Routledge, 2008.

Beckermann, A., 'The Real Reason for the Standard View', in Meijers, ed, 2001: 51 – 67.

Beebe, H., 'Causing and Nothingness', in J. Collins, N. Hall, and L. Paul, eds, *Causation and Counterfactuals*. Cambridge, Mass.: MIT Press, 2004: 291 – 308.

Beebe, H., Hitchcock, C., and Menzies, P., eds, *The Oxford Handbook of Causation*. Oxford: Oxford University Press, 2009.

Bermudez, J.L., *Philosophy of Psychology*. New York: Routledge, 2005.

Bernstein, R., 'McDowell's Domesticated Hegelianism', in Smith, ed, 2002: 9 – 24.

Bilgrami, A., *Self-Knowledge and Resentment*. Cambridge, Mass.: Harvard University Press, 2006 (a).

Bilgrami, A., 'Some Philosophical Integrations', in Macdonald and Macdonald, eds, 2006: 50 – 66 (2006b).

Bird, G., 'McDowell's Kant: *Mind and World*', *Philosophy*, 71 (1996): 219-243

Block, N., 'Do Causal Powers Drain Away?' *Philosophy and Phenomenological Research*, 67 (2003): 133–150.

Brandom, R., *Making It Explicit*. Cambridge, Mass.: Harvard University Press, 1994.

Brandom, R., *Articulating Reasons*. Cambridge, Mass.: Harvard University Press, 2000a.

Brandom, R., ed, *Rorty and his Critics*. Oxford: Blackwell, 2000b.

Brandom, R., 'Non-inferential knowledge, perceptual experience, and secondary qualities: placing McDowell's empiricism', in N.H.Smith (ed), 2002: 92 – 105.

Bransen, J., and Cuyper, S., eds, *Human Action, Deliberation and Causation*. Dordrecht: Kluwer Academic Publishers, 1998.

Brewer, B., 'Mental Causation: Compulsion by Reason', *Proceedings of the Aristotelian Society, Supplement* 69 (1995): 237-254.

Brewer, B., 'Realism and Explanation in Perception', in J. Roessler, H. Lerman & N.Eilan, eds, *Perception, Causation, and Objectivity*. Oxford: Oxford University Press, 2011: 68-81.

Bridges, J., 'Review of *Experience and the World's Own Language: A Critique of John McDowell's Empiricism* by R.Gaskin'. Online at Notre Dame Reviews, 2007. <http://ndpr.nd.edu/review.cfm?id=8743>

Burge, T., 'Mind-Body Causation and Explanatory Practice', in Heil and Mele, eds, 1993: 97 – 120.

Burge, T., *Foundations of Mind: Philosophical Essays, Volume 2*. Oxford: Clarendon, 2007.

Charles, D., and Lennon, K. eds, *Reduction, Explanation, and Realism*. Oxford: Clarendon, 1992.

Chemero, A. and Silberstein, M., 'After the Philosophy of Mind: Replacing Scholasticism with Science', 2007. <http://edisk.fandom.edu/tony.chemero/papers/philmindisover4.pdf>

Child, W., 'Anomalism, Uncodifiability, and Psychophysical Relations', *The Philosophical Review* 102: 2 (1993): 215 – 245.

Child, W., *Causality, Interpretation and the Mind*. Oxford: Clarendon, 1994.

Churchland, P.M., *Matter and Consciousness*. Cambridge, Mass.: MIT Press, 1984.

Churchland, P.S., *Neurophilosophy*. Cambridge, Mass.: M.I.T. Press, 1986.

Crane, T., 'The Mental Causation Debate', *Proceedings of the Aristotelian Society, supplement*, 69 (1995): 211 – 236. URL=
<http://emile.uni-graz.at/pub/05s/2005-04-0115.pdf>

Crane, T., 'Dualism, Monism, Physicalism', 2000.
http://web.mac.com/cranstim/Tims_website/Online_papers_files/Dualism,%20monism,%20physicalism.pdf

Crane, T., *Elements of Mind*. Oxford: Oxford University Press, 2001a.

Crane, T., 'The Significance of Emergence', in Gillett and Loewer, eds, 2001: 207 – 224 (2001b).

Craver, C., *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon, 2007.

Craver, C., and Bechtel, W., 'Top-down causation without top-down causes', *Biology and Philosophy*, 22 (2007): 547-563.

Craver, C., and Darden, L., *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago: University of Chicago Press, 2013.

De Caro, M., 'The Claims of Naturalism', in Beaney, Penco, & Vignollo, eds, 2007: 222 – 237.

De Caro, M., and Macarthur, D., eds, *Naturalism in Question*. Cambridge, Mass.: Harvard University Press, 2004.

De Caro, M., and Macarthur, D., eds, *Naturalism and Normativity*. New York: Columbia University Press, 2010.

De Caro, M., and Voltolini, A., 'Is Liberal Naturalism Possible?', in de Caro and Macarthur, eds, 2010: 69-86.

De Gaynesford, M., *John McDowell*. Cambridge: Polity, 2004.

De Muijnck, W., *Dependencies, Connections, and Other Relations: A Theory of Mental Causation*. Dordrecht: Kluwer Academic Publishers, 2003.

De Vries, W., *Wilfrid Sellars*. Chesham: Acumen, 2005.

De Vries, W., 'McDowell, Sellars, and Sense Impressions', *European Journal of Philosophy* 14:2 (2006): 182 – 201.

De Vries, W., 'Ontology and the Completeness of Sellars's Two Images', *Humana.Mente Journal of Philosophical Studies* 21 (2012): 1-18.

Dewey, J., *Logic: The Theory of Inquiry*. New York: Henry Holt, 1938.

Dodd, J., 'McDowell and identity theories of truth.' *Analysis* 55:3 (1995): 160 – 165.

Dodd, J., 'McDowell's identity conception of truth: a reply to Fish and Macdonald.' *Analysis* 68: 1 (2008a): 76 – 85.

Dodd, J., *An Identity Theory of Truth*. Basingstoke: Palgrave Macmillan, (originally published 2000), second (paperback) edition, 2008 (b).

Dowe, P., *Physical Causation*. Cambridge: Cambridge University Press, 2000.

Dowe, P., 'Causal Process Theories', in Beebe, Hitchcock, and Menzies, eds, 2009: 213 – 233.

Dretske, F., *Explaining Behavior: Reasons in a World of Causes*. Cambridge, Mass.: MIT Press, 1988.

Dretske, F., 'Reasons and Causes', *Philosophical Perspectives*, 3 (1989): 1 – 15.

Dretske, F., 'Mental Events as Structuring Causes of Behaviour', in Heil and Mele, eds, 1993: 121 – 136.

Dretske, F., 'Does Meaning Matter' and 'Reply: Causal Relevance', in Macdonald and Macdonald, eds, 1995: 107 – 120 and 142 – 151 (1995a).

Dretske, F., *Naturalizing the Mind*. Cambridge, Mass.: MIT Press, 1995(b).

Ehring, D., 'Causal Relata', in Beebe, Hitchcock, and Menzies, eds, 2009: 387-413.

Evans, G., *The Varieties of Reference*. Oxford: Clarendon, 1982.

Fair, D., 'Causation and the Flow of Energy', *Erkenntnis*, 14, (1979): 219-250.

Fish, W. and Macdonald, C., 'On McDowell's identity conception of truth', *Analysis* 67: 1, (2007): 36 – 41.

Fish, W. and Macdonald, C., 'The Identity Theory of Truth and the Realm of Reference': where Dodd goes wrong', *Analysis* 69: 2, (2009): 297 – 304.

Fish, W. and Macdonald, C., 'McDowell's Alternative Conceptions of the World', *International Journal of Philosophical Studies*, 19: 1 (2011): 87-94.

Fodor, J.A. 'Making Mind Matter More', *Philosophical Topics*, 17 (1989): 59-80.

Fodor, J.A. 'A Modal Argument for Narrow Content', *Journal of Philosophy*, 88 (1991): 5-26.

Fodor, J.A., *In Critical Condition*. Cambridge, Mass.: MIT Press, 1998.

Follesdal, D., 'Causation and Explanation: a Problem in Davidson's View on Action and Mind', in LePore and Mclaughlin, eds, 1985: 311 – 323.

Foucault, M., *The Order of Things*. London: Routledge, 2002. Originally published 1966.

Frege, G., 'On Sense and Reference', in A. W. Moore, ed, *Meaning and Reference*. Oxford: Oxford University Press, 1993: 23 – 43. Originally published 1892.

Friedman, M., 'Philosophical Naturalism', *Proceedings and Addresses of the American Philosophical Association*, 71:2 (1997): 5 – 21.

Friedman, M., 'Exorcising the philosophical tradition', in N.H.Smith, ed, *Reading McDowell*. London: Routledge, 2002: 25 – 57.

Gaskin, R. *Experience and the World's Own Language*. Oxford: Clarendon, 2006.

Gasser, G., ed, *How Successful is Naturalism?* Lancaster: Gazelle, 2007.

Gibb, S.C., Lowe, E.J., and Ingthorsson, R.D., eds, *Mental Causation and Ontology*. Oxford: Oxford University Press, 2013.

Gibbons, J., 'Mental Causation without Downward Causation', *Philosophical Review*, 115:1 (2006): 79-103.

Gillett, C., 'The Metaphysics of Mechanisms and the Challenge of the New Reductionism', in Schouten and de Jong, eds, 2012: 76-100.

Gillett, C. and Loewer, B., eds, *Physicalism and Its Discontents*. Cambridge: Cambridge University Press, 2001.

Glennan, S., 'Mechanisms and the Nature of Causation', *Erkenntnis*, 44:1 (1996): 49-71.

Glennan, S., 'Rethinking Mechanistic Explanation', *Philosophy of Science*, 69: 3 (2002): s342-s353.

Glennan, S., 'Mechanisms', in Beebe, Hitchcock, and Menzies, eds, 2009: 315-325.

Glennan, S., 'Mechanisms, Causes, and the Layered Model of the World', *Philosophy and Phenomenological Research*, 81:2 (2010): 362-381.

Groff, R., ed, *Revitalizing Causality*. New York: Routledge, 2008.

Groff, R., and Greco, J., eds, *Powers and Capacities in Philosophy: The New Aristotelianism*. London: Routledge, 2013.

Guttenplan, S., ed, *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1994.

Haddock, A., 'McDowell and Idealism', *Inquiry*, 51:1 (2008): 79 – 96.

Haddock, A., and Macpherson, F., eds, *Disjunctivism: Perception, Action, Knowledge*. Oxford: Oxford University Press, 2008.

Hall, N., 'Two Concepts of Causation', in J. Collins, N. Hall, and L. Paul, eds, *Causation and Counterfactuals*. Cambridge, Mass.: MIT Press, 2004. http://fitelson.org/269/Hall_TCOOC.pdf

Harbecke, J., *Mental Causation: Investigating the Mind's Powers in a Natural World*. Heusenstamm: ontos verlag, 2008.

Harre, R., and Madden, E., *Causal Powers. A Theory of Natural Necessity*. Oxford: Blackwell, 1975.

Heil, J., and Mele, A., eds, *Mental Causation*. Oxford: Oxford University Press, 1993.

Heßbruggen-Walter, S., 'Spontaneity and Causality: McDowell on the Passivity of Perception', in M. Willaschek, ed, 1999: 21-25.

Honderich, T., 'The Argument for Anomalous Monism', *Analysis*, 42 (1982): 59 – 64.

Horgan, T., 'Actions, Reasons, and the Explanatory Role of Content', in McLaughlin, ed, 1991: 73 – 101.

Hume, D., *A Treatise of Human Nature*. London: J.M.Dent & sons, 1911.

Illari, P., and Williamson, J., 'What is a Mechanism? Thinking about mechanisms *across* the sciences', *European Journal for the Philosophy of Science*, 2:1 (2012): 119-135.

Jackson, F., *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon, 1998.

Jackson, F., Pettit, P., and Smith, M., *Mind, Morality, and Explanation*. Oxford: Clarendon, 2004.

Jaworski, W., 'Mental Causation from the Top-Down', *Erkenntnis*, 65: 2 (2006): 277-299.

Jaworski, W., *Philosophy of Mind*. Chichester: John Wiley, 2011.

Kant, I., *Critique of Pure Reason*. Translated by Norman Kemp Smith. Basingstoke: Palgrave Macmillan, 1929.

Kistler, M., *Causality and Laws of Nature*. London: Routledge, 2006 (published in French 1999).

Kistler, M., 'Mechanisms and Downward Causation', *Philosophical Psychology*, 22 (2009): 1-17.

Kitcher, P., 'Explanatory Unification', *Philosophy of Science*, 48 (1981): 507-531.

Kripke, S., *Wittgenstein on Rules and Private Language*. Oxford: Blackwell, 1982.

Kutach, D., *Causation and Its Basis in Fundamental Physics*. Oxford: Oxford University Press, 2013.

Ladyman, J. and Ross, D., *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press, 2010.

Laurence, S. and Macdonald, C., eds, *Contemporary Readings in the Foundations of Metaphysics*. Oxford: Blackwell, 1998.

LePore, E., and McLaughlin, B., eds, *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell, 1985.

Lewis, D., 'Causal Explanation', in Ruben, ed, 1993: 182-206. (Originally published 1986).

Lindgaard, J., ed, *John McDowell: Experience, Norm, and Nature*. Oxford: Blackwell, 2008.

Lipton, P., 'Causation and Explanation', in Beebe, Hitchcock, and Menzies, eds, 2009: 619-631.

Loewer, B., 'Comments on Jaegwon Kim's *Mind in a Physical World*', *Philosophy and Phenomenological Research*, 65 (2002): 655-662.

Loewer, B., 'Mental Causation, or Something Near Enough', in McLaughlin and Cohen, eds, 2007: 243-264.

Lombard, L., *Events: A Metaphysical Study*. London: Routledge & Kegan Paul, 1986.

Luntley, M., *Contemporary Philosophy of Thought*. Oxford: Blackwell, 1999.

Macarthur, D., 'Naturalizing the Human or Humanizing Nature: Science, Nature and the Supernatural', *Erkenntnis*, 61:1 (2004): 29 – 51.

Macbeth, D., 'Pragmatism and Objective Truth', in Misak, ed, 2007:169-192.

Macdonald, C., *Mind-Body Identity Theories*. London: Routledge, 1989.

Macdonald, C. and Macdonald, G., 'Mental Causes and Explanation of Action', *Philosophical Quarterly*, 36: 143 (1986): 145 – 158.

Macdonald, C. and Macdonald, G., 'Mental Causation and Nonreductive Monism', *Analysis*, 51:1 (1991): 23 – 32.

Macdonald, C. and Macdonald, G., eds, *Philosophy of Psychology*. Cambridge, Mass.: Blackwell, 1995a.

Macdonald, C. and Macdonald, G., 'How to be Psychologically Relevant', in Macdonald and Macdonald, eds, 1995a: 60 – 77 (1995b).

Macdonald, C. and Macdonald, G., 'Introduction: Causal Relevance and Explanatory Exclusion', in Macdonald and Macdonald, eds, 1995a: 86 – 106 (1995c).

Macdonald, C. and Macdonald, G., eds, *McDowell and his Critics*. Oxford: Blackwell, 2006a.

Macdonald, C. and Macdonald, G., 'The Metaphysics of Mental Causation', *Journal of Philosophy*, 103 (2006): 539 – 576. (2006b).

URL=

http://www.phil.canterbury.ac.nz/personal_pages/cynthia_macdonald/publications/MMC.pdf. Accessed June 2007.

Macdonald, C., and Macdonald, G., 'Emergence and Downward Causation', in Macdonald and Macdonald, eds, *Emergence in Mind*. Oxford: Oxford University Press, 2010: 139 – 178.

Macdonald, G., 'The Nature of Naturalism', *Proceedings of the Aristotelian Society Supplement*, 66 (1992), 225 – 244 (1992b).

Macdonald, G., 'The Two Natures: Another Dogma?', in Macdonald and Macdonald, eds, 2006: 222 – 235.

Macdonald, G., 'Emergence and Causal Powers', *Erkenntnis*, 67:2 (2007): 239 – 253.

Machamer, P., Darden, L., and Craver, C., 'Thinking about Mechanisms', *Philosophy of Science*, 67:2 (2000): 1-25.

Maguire, E.A, Gadian, D.G., Johnsrule, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S., and Frith, C.D., 'Navigation-related Structural Change

in the Hippocampus of Taxi Drivers'. *Proceedings of the National Academy of Sciences*, 97(8): 2000: 4398-4403.

Malcolm, N., 'The Conceivability of Mechanism', *The Philosophical Review*, 77 (1968): 45-72.

Marcus, E., 'Mental Causation in a Physical World', *Philosophical Studies*, 122 (2005): 27-50.

Marcus, E., *Rational Causation*. Cambridge, Mass.: Harvard University Press, 2012.

McCulloch, G., *The Mind and its World*. London: Routledge, 1995.

McGinn, C., 'The Structure of Content', in C. McGinn, *Knowledge and Reality: Selected Essays*. Oxford, Clarendon, 1999: 111 – 151 (1999).

McLaughlin, B., 'Anomalous Monism and the Irreducibility of the Mental', in LePore and McLaughlin, eds, 1985; 331 – 368.

McLaughlin, B., ed, *Dretske and His Critics*. Oxford: Blackwell, 1991.

McLaughlin, B., 'On Davidson's Response to the Charge of Epiphenomenalism', in Heil and Mele, eds, 1993: 27 – 40.

McLaughlin, B., and Cohen, J., eds, *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell, 2007.

McLaughlin, B., Beckermann, A., and Walter, S., eds, *the Oxford Handbook of Philosophy of Mind*. Oxford: Clarendon, 2009.

Meijers, A., ed, *Explaining Beliefs: Lynne Rudder Baker and her Critics*. Stanford: CSLI Publications, 2001.

Melnyk, A., *A Physicalist Manifesto*. Cambridge: Cambridge University Press, 2003.

Menzies, P., ‘The Causal Efficacy of Mental States’, in Walter and Heckmann, eds, 2003: 195 – 223.

Menzies, P., ‘Mental Causation in the Physical World’, in Gibb, Lowe, and Ingthorsson, eds, 2013: 58-87.

Merricks, T., *Objects and Persons*. Oxford: Clarendon Press, 2001.

Meyering, T., ‘The Causal Powers of Belief’, in Meijers, ed, 2001: 119-137.

Montmarquet, J., ‘Review of *Rational Causation* by Eric Marcus’, online at Notre Dame Philosophical Reviews, 2012. <http://ndpr.nd.edu/news/30933-rational-causation/>. Accessed March 2013.

Mumford, S., ‘Powers, dispositions, properties or a causal realist manifesto’, in Groff, ed, 2008: 139-151.

Mumford, S. and Anjum, R., *Getting Causes from Powers*. Oxford: Oxford University Press, 2011.

Nagel, T., *The View from Nowhere*. Oxford: Oxford University Press, 1986.

Nash, O., *Candy is Dandy*. Chatham: Mackays, 1983: 133.

Olson, E., 'Review of *Persons and Bodies: A Constitution View*', *Mind*, 110 (2001): 427-430.

O'Shea, J.R., 'On the Structure of Sellars' Naturalism with a Normative Turn', paper presented at the 'Empiricism and the Philosophy of Mind after 50 Years' Conference, London, June, 2006.

www.philosophy.sas.ac.uk/Empiricism-Mind-Sellars.htm.

Accessed March 2007.

O'Shea, J.R., 'Sellars and McDowell on the myth of the given'. Unpublished, 2005.

O'Shea, J.R., *Wilfrid Sellars*. Cambridge: Polity, 2007.

Papazoglou, A., 'Nature and the Space of reasons'. <http://cas.uchicago.edu/workshops/conteurophil/files/2010/10/Papazoglou---Nature-and-the-Space-of-Reasons.pdf>. Accessed April 2012.

Paul, L., and Hall, N., *Causation: A User's Guide*. Oxford: Oxford University Press, 2013.

Pereboom, D., 'Constitution and Nonreductive Physicalism', paper given at a conference on Lynne Rudder Baker's work, University at Buffalo (SUNY), 2009.

<http://philosophy.buffalo.edu/events/conferences/LynneRudderBaker>

Accessed June 2011.

Perry, J., 'The Same F', *The Philosophical Review*, 79 (1970): 181-200.

Pettit, P., 'The Nature of Naturalism', *Proceedings of the Aristotelian Society Supplement*, 66 (1992): 245 – 266.

Pettit, P., 'A Definition of Physicalism', *Analysis*, 53: 4 (1993): 213 – 223.

Pettit, P., 'Causality at Higher Levels', in D. Sperber, D. Premack, A. Premack, eds, *Causal Cognition: A Multidisciplinary Debate*. Oxford: Clarendon, 1996: 399-421.

Platts, M., *Ways of Meaning: An Introduction to a Philosophy of Language*. London: Routledge and Kegan Paul, 1979.

Price, H., 'Metaphysical Pluralism'. *Journal of Philosophy*, 89:8 (1992): 387-409.

Price, H., 'Naturalism without Representationalism', in De Caro and Macarthur, eds, 2004: 71 – 88.

Price, H. and Corry, R., eds, *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Clarendon, 2007.

Putnam, H., *The Threefold Cord*. New York: Columbia, 1999.

Ramberg, B., 'Post-ontological Philosophy of Mind: Rorty versus Davidson', in Brandom, ed, 2000b: 351-370 (2000).

Rescher, N., 'The Future of Naturalism: Nature and Culture in Perceptual Duality', in Shook & Kurtz, eds, 2009: 15-23.

Rorty, R., *Philosophy and the Mirror of Nature*. Princeton, New Jersey: Princeton University Press, 1979.

Rorty, R., *Objectivity, Relativism, and Truth*. Cambridge: Cambridge University Press, 1991.

Rorty, R., 'McDowell, Davidson, and Spontaneity'. *Philosophy and Phenomenological Research*, 63, (1998a): 389–394.

Rorty, R., 'The Very Idea of Human Answerability to the World: John McDowell's Version of Empiricism'. In his *Truth and Progress: Philosophical Papers, Volume 3*. Cambridge, Cambridge University Press, 1998(b): 138 – 152.

Ross, D., 'Rainforest Realism: A Dennettian Theory of Existence', in D. Ross, A. Brook, and D. Thompson, eds, *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, Mass.: MIT Press, 2000: 147-168.

Ross, D., Ladyman, J., and Spurrett, D., 'In Defence of Scientism', in Ladyman and Ross, eds, 2010: 1-65.

Russell, B., 'On Denoting', *Mind*, 14: 56 (1905): 479-493.

Russell, B., 'Knowledge by Acquaintance and Knowledge by Description', *Proceedings of the Aristotelian Society*, 11 (1910-1911): 108-128.

Russell, B., 'On the Notion of Cause', *Proceedings of the Aristotelian Society*, 13 (1913): 1-26.

Sachs, C., 'Resisting the Disenchantment of Nature: McDowell and Naturalized Transcendental Philosophy..
<http://umassed.academia.edu/documents/0042/1206/resistance.doc>
 Accessed April 2010.

Sachs, C., 'The Shape of a Good Question: McDowell, Evolution, and Transcendental Philosophy'.
http://umassed.academia.edu/documents/0042/1208/good_question.doc
 Accessed April 2010.

Sacks, M., *The World We Found*. London: Duckworth 1989.

Sacks, M., *Objectivity and Insight*. Oxford: Clarendon, 2000.

Sainsbury, R.M., *Reference without Referents*. Oxford: Oxford University Press, 2005.

Salmon, W., *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press, 1984.

Savellos, E., and Yalcin, U., *Supervenience: New Essays*. Cambridge: Cambridge University Press, 1995.

Schaffer, J., 'Causation by Disconnection' *Philosophy of Science*, 67 (2000): 285-300.

Schaffer, J., 'Is There a Fundamental Level?', *Nous*, 37 (2003): 498-517.

Schaffer, J., 'The Metaphysics of Causation', *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2008/entries/causation-metaphysics/>>. Accessed March 2009.

Schueler, G.F., *Reasons and Purposes*. Oxford: Clarendon, 2003.

Shook, J.R., and Kurtz, P., *The Future of Naturalism*. Amherst, N.Y.: Humanity Books, 2009.

Sider, T., 'Review of *Persons and Bodies; A Constitution View*', *Journal of Philosophy* 99 (2002): 45 – 48.

Smith, J., and Sullivan, P., eds, *Transcendental Philosophy and Naturalism*. Oxford: Oxford University Press, 2011.

Smith, N., ed, *Reading McDowell: On Mind and World*. London: Routledge, 2002.

Sosa, E., and Villanueva, E., eds, *Philosophy of Mind: Philosophical Issues Volume 13*. Oxford: Blackwell, 2003.

Stanford, P.K., 'The Eyes Don't Have It: Fracturing the Scientific and Manifest Images' *Humana. Mente Journal of Philosophical Studies*, 21 (2012): 19 – 44

Steward, H., *The Ontology of Mind*. Oxford: Oxford University Press, 1997.

Stoljar, D., *Physicalism*. London: Routledge, 2010.

Stoutland, F., 'Davidson on Intentional Behavior', in LePore and McLaughlin, eds, 1985: 44 – 59.

Stoutland, F., 'Reasons, Causes, and Intentional Behaviour', *Analyse und Kritik*, 8: (1986): S. 28 – 55.

Stoutland, F., 'The Real Reasons', in J. Bransen and S. Cuypers, eds, *Human Action, Deliberation, and Causation*. Dordrecht: Kluwer Academic, 1998: 43 – 66.

Stoutland, F., 'Reasons for Action and Psychological States', in Leist, ed, 2007: 75 – 94. (2007a)

Stoutland, F., 'Commonsense Psychology and Scientific Explanation', 2007, online at www.fil.lu.se/hommageawlodek (accessed May 2010) (2007b).

Stoutland, F., 'Determinism, Intentional Action, and Bodily Movements', in Sandis, ed, 2009: 313 – 337.

Stoutland, F., 'Interpreting Davidson on Intentional Action', in Malpas, ed, 2011: 297-324 (2011b).

Strawson, P.F., *The Bounds of Sense*. London: Routledge, 1966.

Strawson, P.F., 'Causation and Explanation', in Vermazen and Hintikka, eds, 1985: 115 – 135.

Strawson, P.F., *Scepticism and Naturalism: Some Varieties*. Abingdon: Routledge, 2008 (first published 1985).

Stroud, B., 'Sense-experience and the grounding of thought', in N.H.Smith, ed, *Reading McDowell: On Mind and World*. London: Routledge, 2002: 79 – 91.

Stueber, K., 'Mental Causation and the Paradoxes of Explanation', *Philosophical Studies*, 122 (2005): 243 – 277.

Suhm, C., Wagemann, P., and Wessels, F., 'Ontological Troubles with Facts and Objects in McDowell's *Mind and World*', in Wilaschek, ed, 1999: 27 – 33.

Thomasson, A., *Ordinary Objects*. Oxford: Oxford University Press, 2007.

Thornton, T., *Wittgenstein on Language and Thought: The Philosophy of Content*. Edinburgh: Edinburgh University Press, 1998.

Thornton, T., *John McDowell*. Chesham: Acumen, 2004.

Tomberlin, J.E., *Philosophical Perspectives, 11, Mind, Causation, and World*. Oxford: Blackwell, 1997.

Travis, C., 'Reason's Reach', in Lindgaard, ed, 2008: 176 – 199

Tuomela, R., *Science, Action, and Reality*. Dordrecht: Reidel, 1985.

Van Brakel, J., 'Empiricism and the Manifest Image', in I. Douven & L. Horsten, eds, *Realism in the Sciences*. Leuven: Leuven University Press, 1996: 147-164.

Van Brakel, J., *Philosophy of Chemistry*. Leuven, Leuven University Press, 2000.

Van Fraassen, B., 'The Manifest Image and the Scientific Image', in D. Aerts, J. Broekaert, & E. Mathijs, eds., *Einstein Meets Magritte: The White Book – An Interdisciplinary Reflection*. Dordrecht: Kluwer, 1999: 29-52.

Van Gulick, R., 'Who's in Charge Here? And Who's Doing all the Work?', in J. Heil & A. Mele, eds, 1993: 233-256.

Vicente, A., 'The Dual *Explanandum* Strategy', *Critica*, 34: 101 (2002): 73-96.

Von Wright, G. H., *Explanation and Understanding*. Ithaca, N.Y.: Cornell University Press, 1971.

Wagner, S. and Warner, R., *Naturalism: A Critical Appraisal*. Notre Dame, Indiana: University of Notre Dame press, 1993.

Wahlberg, M., *Reshaping Natural Theology*. Basingstoke: Palgrave Macmillan, 2012.

Walter, S. and Heckman, H-D., eds, *Physicalism and Mental Causation*. Exeter: Imprint Academic, 2003.

Warner, R. and Szubka, T., eds, *The Mind-Body Problem*. Oxford: Blackwell, 1994.

Willaschek, M., ed, *John McDowell: Reason and Nature*. New Brunswick: Transaction Publishers, 1999.

Williams, M., 'Exorcism and Enchantment', *The Philosophical Quarterly*, 46:182 (1996): 99 – 109.

Williams, M., 'Science and Sensibility: McDowell and Sellars on Perceptual Experience', *European Journal of Philosophy* 14:2 (2006): 302 – 325.

Wilson, J., 'Supervenience-based Formulations of Physicalism', *Nous*, 39:3 (2005): 426-459.

Wilson, J., 'On Characterizing the Physical', *Philosophical Studies*, 131 (2006): 61-99.

Wittgenstein, L., *Tractatus Logico-Philosophicus*. London: Routledge Classics, 2001.

Wittgenstein, L., *Philosophical Investigations*. Oxford: Blackwell, 1958.

Woodward, J., *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press, 2003.

Wright, C., 'McDowell's Oscillation'. *Philosophy and Phenomenological Research* 58, (1998): 395 – 402.

Wright, C., 'Human Nature?', in Smith, ed, 2002: 140 – 173.

