A Study on Mental Causation: The Problem of Causal Exclusion

by Mircea Mihail Cucu

Submitted in partial fulfillment of the requirements for

the Degree of Doctor of Philosophy

CEU eTD Collection

Date of submission:

February 16, 2010

Supervisor: Professor Howard Michael Robinson

Abstract

My thesis is an analysis of the too easily claimed conflict between the idea of a causally potent mind of a non-physical kind and the hypothesis of the closure of the physical domain. Such a conflict is known as the Causal Exclusion Problem (CEP).

Interactionism proposes a metaphysical picture of how psychical phenomena intermingle with bodily ones by causally affecting one another, where mental properties are not token-identical with physical properties. CEP gives credit to the idea that, if one is committed to physical closure, the interactionist doctrine practically asks one to admit that in any case of mental causation the physical effect is overdetermined. Thus, interactionism would lead to a widespread and systematic overdetermination, which is a highly implausible situation. Any account of mental causation is more preferable to one implying that our world spends more resources than it needs for its becoming.

The whole thesis is an attempt at showing that CEP is not quite a firm ground for building up an argument that interactionism could not be an answer to the mind-body problem. A supporter of interactionism not willing to defend his doctrine at the cost of contesting physical closure should base his/her comeback on finding some compatibilist scenario. In such a scenario, if it were real, the physical effect would not be overdetermined even though it would have a non-physical cause.

The key point of my thesis is the endeavor to set up such a kind of compatibilist scenario. The strategy I have used rests on finding a *necessary condition on*

CEU eTD Collection

i

overdetermination and some relation tokening among some of the causes of the physical effect such that, if the relation held, the necessary condition would be violated.

Given that 'overdetermination' is a derivative notion, its understanding requires the grasp of other notions as 'causal process sufficient for an effect' and 'excess in bringing about an effect'. Their analysis rests on a nomic account of causation whose main claim is that any causal relation is part of some relation of nomic necessity. Since there is no causation without events, the identity of the causal network involved by CEP comes in part from the identity of the events featuring as causal relata. In my view, any event has a metaphysical structure characterized by a form that is temporal in nature.

The compatibilist scenarios designed under various frameworks for mental causation do cast doubt on CEP. As long as there is no convincing argument that our actual world does not accommodate the compatibilist scenarios, there is no guarantee for the soundness of CEP.

Some of the proposed compatibilist scenarios have in common a causal linkage that can be described as *simultaneous causation*. The idea is not entirely new; it has been mentioned or used by philosophers like David Lewis and E.J. Lowe. Some examples seem to credit this kind of causation with empirical plausibility.

ii

Acknowledgements

Acknowledgements

I thank to my supervisor, Professor Howard M. Robinson, for his constant help and ongoing exchange of ideas during this work. Most of all, he has been a model of subtle philosophical thinking. I also want to thank to my second supervisor, Professor Katalin Farkas, for her sharp critical observations. From her I have learnt what huge a difference clarity and structure make in getting your ideas across. I am grateful to Professor Ferenc Huoranszki for fruitful discussions and insightful suggestions in metaphysical matters. I want to express my gratitude to Professor Joseph Levine (University of Massachusetts Amherst) for his thoughtful guidance during my visit at the Ohio State University. Finally, I thank all from the Philosophy Department that helped me in various ways and, in particular, Krisztina Biber for her kindness and availability.

Introduction1
Chapter 1: No Real Mind-Body Issue without a Proper Mental-Physical
Distinction7
1.1 Monism, dualism, and pluralism8
1.2 Being mental, being physical13
1.3 Dismissing the theory approach14
1.4 The paradigm approach16
1.5 Subjects, minds, and phenomenal experiences
1.6 Resuming the result and stating some possible objections
1.7 Neutral monism 24
Addendum 1.1 Why properties, instead objects?
Addendum 1.2 Supervenience, realizability
Chapter 2: Stating the Problem of Causal Exclusion
2.1 Interactionism
2.2 What would mental causation look like?
2.3 The Causal Exclusion Problem (CEP) 46
2.4 One problem, different reactions and employments 50
2.5 The argument from CEP against interactionism
2.6 The comeback compatibilist strategy 58
2.7 The first version of the CEP argument
Chapter 3: Causation and Laws of Nature67
3.1 Singularist vs. non-singularist theories of causes
3.2 Looking into the singularist camp70
3.3 Visiting the non-singularist camp76
3.4 No laws, no causation
3.5 Leading conditions for an account of causation
3.6 The nomic account of causation95
Chapter 4: A Sketchy Metaphysical Theory of Events

	4.1 Davidson's argument for the existence of events	101	
	4.2 Competing candidates for taking a role	104	
	4.3 With or without internal structure??	108	
	4.4 Some difficulties for Kim's view	113	
	4.5 Troubles brought, this time, by complex events	118	
	4.6 Hunting for a metaphysical account of events	120	
	4.7 Categorial structures and metaphysical structures	121	
	4.8 Events as changes in properties	132	
	4.9 Events as temporal structures of states	137	
	4.10 Criteria of identity	140	
Chap	ter 5: The Constituents of Causation	151	
	5.1 Causal relata	152	
	5.2 Causal histories	156	
	5.3 Causal processes	159	
	5.4 Relations holding between two causal processes	163	
	5.4.1 Connecting two causal processes	163	
	5.4.2 Being part of a causal process	167	
	5.4.3 Sharing a causal process	169	
	5.5 The notion of causal role	173	
	Addendum 5.1 The metaphor of causal power	182	
Chan	ter 6: Overdetermination	185	
Chap		105	
	6.1 Questionable intuitions	186	
	6.2 Going for something more basic	196	
	6.3 Clearing up two notions	200	
	6.4 A necessary condition on overdetermination	205	
	6.5 Unnecessary causes, overdetermining causes	209	
	6.6 Looking around for a case of overdetermination	213	
	6.7 The logic of the threshold function	215	
	6.8 Pseudo-overdetermination, partial overdetermination	218	
	6.9 An argument for rejecting any case of overdetermination	225	
	6.10 Rejecting widespread and systematic overdetermination	227	

Chapter 7: Beating the CEP Argument 235
7.1 General strategy 236
7.2 The notion of physical closure in a weak reading (PC _w)
7.3 A simplified framework for designing compatibilist scenarios 244
7.3.1 A "candid" scenario246
7.4 The notion of physical closure in a mild reading (PC_M)
7.5 A "splitting" scenario
7.5.1 Is it plausible a triangular linkage of laws?
7.6 The notion of physical closure in a strong reading (PC _s)
7.6.1 Lowe's "divine" scenario275
7.7 More working frameworks, more compatibilist scenarios
7.7.1 m and p do not overlap each other
7.7.2 Another "candid" scenario282
7.7.3 Another "splitting" scenario283
7.7.4 m and p overlap each other285
7.7.5 One more "candid" scenario286
7.7.6 One more "splitting" scenario287
Conclusions
Bibliography

Introduction

How should we conceive the relationship between mind and body? This is the mindbody problem. It requires an account of the mechanisms that produce mental phenomena in cognitive subjects like human beings. In other words, to address the mind-body problem is to advance an account for how psychical phenomena intermingle with bodily ones.

One may find an impressive variety of theories aiming at giving a definitive answer to the issue of the interweaving between mental and physiological processes. According to what entities are held to be basic for the actual world, any such theory is committed to a certain metaphysical framework or doctrine. Chapter 1 opens the dissertation with a short presentation of the main contenders to the mind-body problem: monism, dualism, and pluralism. Even though all these views admit the existence of two ontological realms, mental and physical, the mental-physical distinction is still not easily traceable and hence, many times, it has not received a proper cut in the relevant literature. A good deal of Chapter 1 approaches questions like "What is it for an entity to be physical?" and "What is it to be mental?"

The central topic of the dissertation is the Causal Exclusion Problem (CEP, as I shall refer to it). This is the name of an important objection coming from the physicalist camp, the modern version of the monist materialism, against interactionism, the dualist doctrine that argues for the causal interactions among the entities of different natures (mental and physical). As it is made explicit in Chapter 2, the objection is that

CEU eTD Collection

there is a conflict between the idea of a causally potent mind of a non-physical kind, with respect to the behavior and the hypothesis of physical closure. The defenders of CEP claim that the existence of interactionist minds in a world where the physical domain is closed would lead to a highly implausible situation: each and every bit of action we perform (or, more to the point, any neurological event involved by mental processes) would be overdetermined.

The question of whether the threat of such a systematic and widespread overdetermination is genuine has brought a considerable disagreement among philosophers, splitting the theoretic field into two opposite camps: exclusionism and causal compatibilism. Taking the exclusionist stance, many physicalists have used CEP in order to produce an argument against interactionism (the CEP argument). Since, according CEP, interactionist theses lead to an implausible overdetermination in mental causation, the physicalist claims their falsity.

I shall argue that the CEP argument is not conclusive for dismissing interactionism. My attempt at blocking the physicalist attack will focus on showing that there might be cases where the changes in the interactionist mind do not overdetermine their physical effects. In other words, even if the physical domain is closed, there is no guarantee that interactionism does entail overdetermination. Dualism is still a framework of research worth considering by the metaphysicians of mind.

Since CEP is highly dependent on the notion of causation, it is this notion that should come under scrutiny before any discussion on the soundness of CEP and of the arguments it generates. One cannot undertake a substantial consideration of CEP

CEU eTD Collection

and its consequences unless one takes a particular stance on the matter of causation itself. One of the fundamental convictions that shape my analysis of CEP is that the issue of causation should be understood in the light of the metaphysical relations supported by the laws of nature. Chapter 3 presents the nomic account of causation, whose main thesis is that any case of causation involves a case of necessitation holding in virtue of certain laws of nature. By introducing the notion of primitive causal relation, it is suggested a way of distinguishing between cases of genuine causation and various cases of non-causal correlation.

Since my work is built up on the belief that there is no causation without events, Chapter 4 is devoted to a discussion on the nature of events. In the beginning, I present Davidson's argument for positing events. However, his approach leaves unsettled neither whether such entities have any structure, nor what such a structure would be. Inspired by Kim's view on events, it is analysed in some detail what it is to be understood by 'having a structure' when it is about an ontological entity. The whole discussion proves to be crucial in approaching the issue of event identity.

Along the chapter, various objections are addressed to the views proposed by Davidson or by Kim. Finally, I present the theory of events that I favor in a sketchy manner. Starting from the common idea that an event is a change in properties, I suggest a view that takes events to be temporal structures of states. Naturally, if I am right, the identity of an event depends also on the temporal profile of transition through the intermediary states.

Chapter 5 will provide some notional support for the discussion to follow on overdetermination. A series of causal notions referring to some of the constituents of causation will be defined, explained and illustrated by proper examples.

Causes are linked in causal chains that further generate the causal history of any effect. However, in giving causal explanations, causal processes are more workable than causal histories. Therefore, the chapter will define and elaborate at length the notion of causal process. The discussion will focus on matters such as how two causal processes could be connected to each other and what it is for a causal process to be part of another one. Also, it will be analysed the case when two causal processes, each bringing the same ending effect, share one of their parts. The chapter ends up with the notion of causal role, whose main benefit is that it discriminates among the different ways the same effect is or could be directly caused.

Chapter 6 is an analysis of the concept of causal overdetermination. I hold that if there is any way of answering the exlusionist's attack from CEP against interactionism, the first step to be taken is to look for a necessary condition on overdetermination. The chapter undertakes a conceptual analysis that leads to the conclusion that the understanding of 'overdetermination' should be grounded on the grasping of other two notions: 'causal process sufficient for an effect' and 'excess in bringing about an effect'. The study of these notions suggests a certain set of necessary conditions on overdetermination.

Introduction

Next, the chapter reviews a number of causal scenarios that are recurrently cited as cases of overdetermination in today's literature. By appealing to the theory of events sketched in Chapter 4, I shall argue for the conclusion that some of classical examples are highly debatable, many of them illustrating only cases of pseudo-overdetermination or partial overdetermination. Also, I shall propose an argument intended to challenge the very idea of overdetermination. The chapter closes with a study on what would ground our reluctance to admit widespread and systematic overdetermination.

In the last chapter, I take into account three possible working frameworks for mental causation, designing in each of them scenarios in which, even if mental causation is consistent with the interactionist claims, causation does not involve overdetermination. The strategy I shall use is based on finding a necessary condition on overdetermination and some relation tokening among some of the causes of the physical effect such that, if the relation held, the necessary condition would be violated.

Some of the compatibilist scenarios have in common a causal linkage that can be described as simultaneous causation. In such cases, causation is accomplished in such a way that the effect is necessitated independently by two distinct laws whose activation is made by the same cause. I shall try to give empirical plausibility to this idea by crediting it with a number of examples.

CEU eTD Collection

As long as there is no convincing argument that our actual world does not accommodate the compatibilist scenarios, there is no guarantee for the soundness of CEP and it cannot be used in a conclusive argument against interactionism.

Chapter 1: No Real Mind-Body Issue without a Proper Mental-Physical Distinction

The content of this chapter is rather preliminary to the central topic of the dissertation, namely, the Causal Exclusion Problem (CEP). Since it has become known as the main attack coming from the physicalist camp against the interactionism, § 1.1 prepares the battlefield by introducing in general terms the protagonists of the dispute: monism, dualism, and pluralism. Also, of crucial importance for the understanding the true meaning and the real force of CEP is the mental-physical distinction. In § 1.3 and § 1.4 investigates two alternative routes on which people usually proceed, when making the distinction: the theory approach and the paradigm approach. My own stand is sympathetic with the latter one: the objects devoid of sentience are the best guide in searching for physical properties. Since the concept of subjectivity appears to be crucial for the metal-physical distinction, § 1.5 houses a discussion about what it is for a subject to have phenomenal experiences: is there a relation of constitution? § 1.6 concludes what it means to be a mental property and what it means to be a physical one, and signals the possible objection that the obtained result leaves no room for a doctrine like neutral monism. Therefore, the chapter closes with a survey on the main arguments for neutral monism, ending with a discussion on the some difficulties of the notion of 'neutral'.

1.1 Monism, dualism, and pluralism

In the context of a discussion in the metaphysics of mind, there are three general ontological doctrines: monism, dualism, and pluralism. Since they are intended as general views about the nature of the *actual* world¹, each of them could provide a certain framework for various attempts to tackle the mind-body problem.

Among monism, dualism, and pluralism, what essentially differentiate each doctrine are the kinds of entities it holds that *ultimately* there are. Within the framework given by each doctrine, the relationship between the mental entities and the physical entities *might* be traced via the constitutional relationships holding between any of these entities and the *basic* entities posited by the respective doctrine. (See Figure 1.1.) Here, 'basic' is used to characterize entities that are not constituted by others.²

It should be noted that there is one more sense of 'basic' which, this time, is applicable for properties. A property is taken to be basic if its instances do not supervene on others, that is, if its instantiation in some entity is not necessitated by the instantiation of other properties in that entity or in its constituents. On contrary, the instances of any property that is not basic are bound to supervene on the instances of some basic properties.³

¹ They are not always only about the actual world. Some versions of the mentioned doctrines have even significant consequences for how the world *would/might* have been under certain conditions, or even for the nature of any *possible* world.

² Almost all metaphysics take for granted the existence of such basic entities, ignoring the possibility of an ontologically infinite detailing without reaching the *bottom of being*.

³ For more on supervenience and realization (as a particular case of supervenience), see Addendum 1.2.



Fig. 1.1

Another point to be noticed here is that the mind-body problem could be seen also as a problem of mixed entities, because it requires an answer to the following question: how should a mixed constitution be accounted for? Assuming that both mental and physical entities feature in the same mixed entity, how is that possible? As I have mentioned before, I believe that the answer strongly depends on what view one decides to take on the nature of the basic elements that constitute or necessitate a mental or a physical entity.⁴

⁴ I do not claim that an account of the mental-physical relationships should be nothing else but an analysis of the constitution of mental entities and of physical entities involved in the mind-body relation. Most likely, a complete account should also make use of the laws governing all these entities.

To begin with, monism is the view that all that exists in the universe is ultimately made up of one and only one kind of entities. This opens three doctrinaire options:

- Physicalism (or materialism) holds that, from a bottom-up perspective, all primitive constituents of our world are physical entities; or, changing the viewpoint, anything is ontologically reducible to some basic physical entities; in this sense, everything that exists is physical,⁵ mental entities being only a specie of this universal kind (see Figure 1.2a);
- 2) Phenomenalism (or idealism⁶) holds the converse thesis, namely, that physical entities are only members of a particular species of the comprehensive kind of mental entities (see Figure 1.2b); the idealist contends that all entities objects, properties, events, etc. are constituted of mental entities, where a mental entity is a mind-dependent entity, that is, an entity whose reality is only in virtue of being in, or of being for, a mind; thus, the idealistic thesis is that any reality is mind-dependent: *to exist is to be for a mind*;
- 3) Neutral monism is the commitment to the view that there are ultimately entities that are neither physical, nor mental; the neutral basis accounts for any physical or mental reality.

⁵ Going bottom-up, by the way of constitution, an entity whose basic constituents belong to a certain ontological kind will be of the same kind. For instance, if physicalism is true, any entity is constituted of physical entities and hence would be itself physical. The rule of thumb is that if the basic constituents of an entity, *x*, belong to the *N*-kind (e.g., physical or mental), *x* will belong to the same kind. Yet, at an intermediary level of constitution, if *x*'s basic constituents make up some higher order constituents which belong to the *P*-kind – a species of the *N*-kind –, *x* will be also of the *P*-kind.

Of course, even though the principle of the kind preservation is true on the bottom-up route of constitution or, more generally, of necessitation, it might not be true on the top-down route.

⁶ Paradigmatic for this doctrine is Berkeley's idealism, holding that that the world is made of ideas.



THE BASIC LEVEL

a) Physicalism:

Everything that exists is physical; the mental is only a specie of the physical. b) Idealism: Everything that exists is mental; the physical is only a specie of the mental.



Whereas on the monist view everything that exists is ontologically reducible to one single, either mental or physical, kind of basic entities, the dualist doctrine claims the existence of two kinds of basic entities - mental and physical.⁷ Accordingly, any nonbasic entity has a constitution made up either of only pure physical entities, or of only pure mental entities, or of mixed entities. (See Figure 1.3.)

⁷ Actually, the matter is not so clearly cut. What I have presented above is a radical form of dualism, namely, substance dualism. In fairness, there is also the property dualist doctrine. In some sense, it seems to be rather somewhere in-between substance dualism and physicalism, borrowing some features from each of the two disputing parts. I shall return to property dualism on the next chapter, where more will be said about each of its versions.



Figure 1.3 Dualism: There are two kinds of basic entities: mental and physical.

Finally, pluralism is the philosophical doctrine according to which the bricks of existence are of many kinds; hence physical and mental entities could be reduced ontologically to a large variety of entities. For the sake of simplicity, in what follows I shall ignore the pluralist doctrine. Given that it is only a sophistication of dualism, in the face of the causal exclusion problem, pluralism and dualism either survive together, both of them being immune to the alleged threat, or succumb together. For one thing, they share exactly the criticized idea, namely, that there are causal interactions holding between some physical entities and some other entities of a very different nature.

Further, given that the idealist alternative is symmetrical to the physicalists one, the considerations made with respect to the latter option mirrors the considerations to be made with regard to the other one. Therefore, for the sake of simplicity, I shall

disregard the idealist doctrine in the development of my argumentation about mental causation.

1.2 Being mental, being physical

The previous section inevitably involved us in a good deal of talk about two kinds of entities – mental and physical –, but left unaddressed the mental-physical distinction as such.⁸ So, we have to face questions like 'What is it for an entity to be physical?' and 'What is it to be mental?' My preliminary claim is that, whatever answers the question may receive, they should not exclude by themselves the "candidature" of any psycho-physical theory. Or, at least, one should do one's best to come with such non-discriminatory answers.

Let us begin with the notion of 'physical'. In defining the realm of physical entities, there are two alternative routes on which people usually proceed. One, which we may call 'the theory approach', is to rely on the theory/discourse in which the concept 'physical' is typically and coherently used. The other, which we shall call 'the paradigm approach', is to look for entities that, under a large agreement, fall under the respective concept.⁹

⁸ We should ignore here any form of eliminativism (either mental or physical). For example, within the borders of eliminative materialism, whose defenders do not acknowledge the existence of any mental entity (they aim to get ride of the whole folk mentalistic vocabulary, considering that the only legitimate discourse is one in physicalistic terms), there is no answer to the mind-body problem simply because there is no such thing as the mind-body relationship.

⁹ To a certain extent, here I am following Stoljar, who distinguishes two conceptions which usually are taken to account for the term 'physical property': the "theory-based conception of a physical property" and the "object-based conception of a physical property". See Stoljar (2001a, 2001b).

1.3 Dismissing the theory approach

According to the theory approach, the notion of 'physical' could be understood in relation to the content of the physical sciences. Thus, an entity would be physical if and only if there is of the same sort as those entities considered or posited in the discourses of the completed physical theories.

It should be said that by 'physical theories' we usually mean those theories advanced in such domains as microphysics, physical chemistry, biology, medicine and the like.¹⁰ Of course, psychological or sociological theories cannot be included by default among physical theories. For this is partly what is at stake in the mind-body problem: whether or not entities assumed by certain psychological theories – i.e. mental entities, if you like – are reducible to physical entities. On the one hand, if it is shown that psychological concepts are reducible to physical ones, then psychology should be eventually included among physical sciences. On the other hand, if it is shown that mental entities are constituted of physical entities, we should admit that anything mental is also physical, following from this that mental entities should be explained by a physical science and, again, psychology would be rendered as a physical science. Finally, if it were proved that a complete psychological theory is explanatorily reducible to a complete physical theory,¹¹ then, granting that an entity accounted by

¹⁰ Note that not all authors consider the above range relevant for the notion of 'physical'. For instance, Chalmers (1996) appears to restrict it to micro-physics. He defines the physical properties as "the fundamental properties that are invoked by a completed theory of physics" (p. 33), excluding from this class "high-level properties" like the biological property of being a giraffe. Instead, Braddon-Mitchell and Jackson (1996) appear to be uncommitted to Chalmers' restriction, since they use term 'physical' for ingredients "posited in the physical and biological sciences" (p. 13).

¹¹ The notion of 'explanatory reduction' is a key notion for some defenders of emergentism, the view that mental properties are nomically necessitated by physical ones, but not in virtue of some basic physical laws. An example of such a philosopher is Tim Crane, who takes explanatory reduction to be

physical entities and physical laws is also physical, we should welcome psychology among physical sciences.

Nevertheless, the question remains: What does it mean for a theory or a science to be *complete*? It simply means that the laws of the theory, its posited entities, and the generated models fully explain all observable phenomena occurring in a certain domain of reality, which is in the scope of that theory. If so, given that it seems that we are far from being in the possession of any complete physical theories, one might worry that 'physical' is defined via reference to some sciences whose evolutions and final contents are unpredictable. Along its history, a science – given the fact that new theories come into stage – gets rid of some kinds of entities from its ontology and refreshes it with newly posited ones.

But this opens two possibilities: one is that the future physical theories, especially micro-physics, might posit entities completely different from those they assumed now; the other is that, eventually, the ontologies of the complete physical sciences might include mental items. Such possibilities open the door to the treat of getting a concept of 'physical' devoid of content or, anyway, a mental-physical distinction completely vague.¹²

Such a threat comes only from understanding 'physical' through the content of physics. It seems to be a reason enough for approaching the mental-physical distinction from another perspective.

[&]quot;a relation between theories: for instance, when one theory explains why another theory is true, or gives us an insight into the underlying mechanisms which explain how the entities of the reduced theory work." (2001a: 54-5). See also Crane (2001b).

¹² A similar version of the objection is known as 'Hempel's dilemma', originating in Hempel (1969).

1.4 The paradigm approach

As long as the theoretic approach lays the burden of tracing the mental-physical distinction on what happens in sciences, it should be of no wonder that neither the concepts of 'physical' and 'mental', nor their distinction is on a firm ground. But this is so only because both physical and mental entities are mainly conceived of as theoretical entities, i.e. items that feature into the explanations of observable phenomena. Then, the question that pops up naturally in one's mind is: Why would the meaning of 'physical' not be relied on certain observable entities?

I believe that the paradigm approach makes use to a *certain* extent of the abovementioned intuition. The approach picks up the sound idea that our commonsensical knowledge of some paradigmatic exemplars of physical entities may provide us with an appropriate conception of what sort of entities are physical. An example of a philosopher who takes that route is Jackson (1998). According to him:

[P]hysicalists can give an ostensive definition of what they meant by physical properties and relations by pointing to some exemplars of non-sentient objects – tables, chairs, mountains, and the like – and then say that by physical properties and relations, they mean the kinds of properties and relations needed to give a complete account of things like them. (p. 7)

Note that what Jackson takes to be paradigmatic physical objects are "non-sentient objects". The idea is that any object devoid of mentality should be an exemplary physical object. The whole reasoning appears to run as follows.¹³ A paradigmatic physical object must be one very likely inhospitable for any mental property

¹³ Actually, I do not claim that this is Jackson's argument, but I believe that it is a good defence for his defining of physical properties.

instantiation. Since sentience is a necessary mark of this kind of instantiation,¹⁴ it is plausible to maintain that non-sentient objects do not house any form of mentality and, apparently, they are physical objects of a paradigmatic sort.

Now, if one follows Jackson's suggestion and one has a pretty good idea of how to identify a non-sentient being or object,¹⁵ one would be committed to the idea that a physical property is nothing else but a property of the same sort like any property featuring among those to be mentioned in an exhaustive account of a paradigmatic physical object.¹⁶

The nice thing about such an account is that it provides a sufficient condition for being a physical property without the need for a science (in particular, for a micro-physical theory) able to ultimately and completely explain the physical realm. We know *a priori* where we can find physical properties, namely, in any complete explanation of an object devoid of sentience, that is, of an object devoid of phenomenal (subjective) experiences or, in Nagel's terms, of an object without a what-is-like-to-be.

¹⁴ That is, usually, it cannot be the case that some being is equipped with mentality, but it is devoid of sentience.

¹⁵ Supposedly, inorganic things and primitive organisms are the prime candidates of non-sentient objects.

¹⁶ It may be noted that there is a shift to a discourse in terms of properties. For more details about the possibility of discussing only in terms of property, see *Addendum 1.1*.

1.5 Subjects, minds, and phenomenal experiences

The assumption of objects devoid of subjectivity appears to be crucial for the mentalphysical distinction. Levine (2001) makes an explicit point of it by saying that psychophysical doctrines are coherent as far as the assumption stands. Taking rationality, representation and consciousness to be distinct features of mentality, he says:

What is it about tables and chairs that make them paradigmatic examples of the physical? ... it's their non-mentality. As far as I know, tables and chairs, as well as rocks and avocado trees, do not support mental life. In particular, their states do not possess either representational properties or phenomenal, qualitative properties. There is nothing it is like to be a chair, nor does the chair represent; none of its states are about anything. If this is false, then tables and chairs would not be good paradigmatic examples of the purely physical. If it turns out to be false of everything, then, as Boyd says, all bets are off. Materialism and dualism would both be false.

It appears that the ground conceptual level in tracing the mental-physical distinction is given by the concept of subjectivity or phenomenal experience. When we are talking about being subjective or having a first-person perspective we know what we are talking about. For one thing, it is just our way of being as cognitive subjects. We, human beings, are paradigmatic sentient objects. And properties of our phenomenal experiences are paradigmatic mental properties. It follows that by 'mental properties' it should be meant *the kind(s) of properties instantiated by (or in) a phenomenal experience*. Coherently, phenomenal experiences are nothing but purely mental entities.

Any phenomenal experience requires a subject, that is, a mind or a being having a mind or a being capable of mentality.¹⁷ One may conceive of experiences as being

¹⁷ It does not make too much sense to say that x is a phenomenal experience because, even if x is not actually an experience for any subject, there can be a subject for which x is an experience. To say that

modifications of a substance-subject, not necessarily taken to be in the Cartesian sense a mental substance. Or, in the Humean or Russellian style, one may think of a subject being nothing but a bundle of experiences.

The above discussion raises the question of whether there are things as a mind or a consciousness understood as a unity principle for various phenomenal experiences. Any of your phenomenal experiences requires a subject – what you refer by 'l' –, but is there one and the same subject for "your" phenomenal experiences? Do you refer to the same thing each time you utter 'l'? Are your phenomenal experiences related with the same self?

Even though we take a subject to be something capable of having phenomenal experiences, it does not follow that a subject *should* have more than one experience. Nor that a subject *could* have more than one experience. True, it would be mad not to admit that a regular human being is *one* subject having *more than one* experience. (The idea of more than one subject – each having its experiences – in the same human being is not only somewhat bizarre, perhaps, but also too speculative to my mind.) However, we should also admit that nothing precludes the possibility of a subject having only one experience, that is, the possibility of one-experience-minds.¹⁸

x is a phenomenal experience just because there *can* be a subject able to have that experience is the same type of wrong claim as y is a mother just because y can become pregnant and have a child. ¹⁸ Actually, for the adherent to the idea that a mind (including a human mind) is a construction of phenomenal experiences, the only way to avoid the unwanted consequence that phenomenal experiences ontologically precede minds seems to be the positing of one-experience-minds. That would be to say that some phenomenal experiences are minds as such, even though the most primitive minds. The direct consequence would be that some primitive minds feature into the constitution of any human mind.

There is no doubt that most of us are reticent in face of such a speculation. There are at least two reasons. First, it is not at all easy to work out a theory that makes sense of a mind having as constituents other minds. For instance, Nagel (2000), even though he is committed to the idea that our minds are composed, points out that it is not very clear how to tackle such an idea of composition.

He takes split-brain cases to favor the thesis of composed mind:

In an intact brain, the two halves do not lead distinct conscious lives: they support a single consciousness. But the fact that each of them can support a distinct consciousness when separated seems to show that *the normal unified consciousness is composed of mental parts* embodied in the physical parts. These parts are 'mental' in a derivative but nonetheless real sense. (p. 463, my italics)

In Nagel's view, we should look after a "new theory of composition - mental parts and

wholes". However, he appears to be aware of the lack of both empirical evidence and

conceptual resources (or, if you want, intuitions) for supporting and, respectively,

working out such a theory. He says:

Further experiments to investigate the results of combining parts of different conscious nervous systems would be criminal if carried out on human subjects—the only kind who would be able to tell us about the experiential results. (p. 467)

And:

The analogue of a presocratic speculation of psychic atoms that are just like animals, only smaller, is not even a starter in this case, because we don't have ready a coherent idea of larger conscious subjects being composed of smaller ones—as we did have the perfectly clear geometrical idea of larger physical objects or processes being composed of smaller ones. (p. 462)

What seems to be for Nagel the main problem, even though only a conceptual one, in

carrying out a theory of composed minds, is that of dealing with the subjective aspect

of mentality:

The real conceptual problems would come in trying to describe elements or factors of subjectivity too basic to be found as identifiable parts of conscious experience. (p. 467)

In the same vein, Levine (2001) expresses a similar worry about the difficulty of tackling the constitutional relation between a subject like us and more primitive ones:

Are we to imagine that the objects in which [phenomenal] properties inhere, whether they be particles or points in a field, are themselves subjects of experience? It's hard to see how our phenomenal properties are supposed to be related to these very basic ones. What relation is there between me, a subject who is clearly constituted by trillions and trillions of these things, and the little subjects who serve as the ultimate bases? How is my conscious experience explained by reference to theirs? (p. 25)

The second reason against the idea of a mind composed of very primitive minds is that the whole story seems to envisage a panpsyhistic picture of the world where the primitive minds exist not only in the constitution of complex minds like ours, but also in other constitutions, which we usually suppose to be unminded (i.e. devoid of any phenomenal experience). Of course, we cannot reject the possibility that stones, for example, house primitive phenomenal experiences. Instead, we cannot be asked to take it seriously as far as the working hypothesis of this dissertation is that mentality is causally effective. One's phenomenal experiences reflect in one's behavior. But there is no apparent similarity between our paradigmatic intelligent (i.e. rational and intentional) behavior and the manifestations of any regular stone.¹⁹

¹⁹ Levine (2001) points to a similar reason in dismissing the possibility that things like tables, chairs and the like are constituted by mind.

1.6 Resuming the result and stating some possible objections

In my search for a mental-physical distinction available for all mind-body theories, I have reviewed two influential approaches aiming at defining the notion of 'physical'. I have seen that the theory approach is apparently vulnerable to certain difficulties. The main reason for its failing is that, in order to "locate" the basic physical entities, the approach relies on the ontology of the current micro-physics. Since we do not have (if ever) a complete micro-physics, we do not know whether or not the ontology of such a fundamental science will contain what we definitely take to be mental entities. Instead, what we know for sure is that, if the future explanations given in micro-physics appeal to some mental entities, the mental-physical distinction traced by theory approach will collapse.

The analysis of the paradigmatic approach has suggested us a better way to start in tracing the mental-physical distinction, namely, to settle first the meaning of 'mental properties' and only then to derive from it the meaning of 'physical properties'. Thus, we have come to the conclusion that:

(MP) By 'mental properties' it should be meant the kind(s) of properties instantiated by (or in) phenomenal experiences.

One may reasonably suppose that it is this meaning of 'mental' that guides Jackson in proposing:

(PP) A physical property is nothing else but a property of the same sort as any property featuring among those to be mentioned in an exhaustive account of a

paradigmatic physical object; such an object would be one devoid of phenomenal experiences.

Remarkably enough, Jackson's definition avoids the objection that we do not have yet a theory able to account ultimately and completely for any non-sentient object.

However, one might note that in defining 'physical' Jackson seems to be assuming, even though in a weak sense, the physicalist thesis that ultimately any non-sentient object would be completely accountable in physical terms. There is at least one objection here: there are plenty of non-sentient objects refractive to any physical account. For instance, it is debatable whether aesthetic properties are accountable in such a way. Would it not be too strong a claim to state that *Gioconda*, say, is completely accountable in physical terms? Maybe it is, but there are still strong doubts about it.

A way out could be a restraint to the intrinsic features of the non-sentient object to be accounted for. Yet, the defender of (PP) may seem to be in a more difficult situation in answering another related objection: (PP) leaves no room for neutral monism, to remember, the view that there is a basic reality, neither physical, nor mental, which ultimately accounts for any physical or mental reality. Prima facie, if there is such a neutral basis, the way depicted by (PP) does not lead to physical properties, but rather to neutral ones.

Note that the same sort of objection may also regard Levine's strategy in drawing the mental-physical distinction. Levine argues that neutral monism – at least, that of Russellian inspiration – is much like an implausible version of panpsychism, one

which assumes that chairs and tables, for example, are supportive for mental life.²⁰ However, even if Russell's neutral monism is hardly acceptable, it is not a sufficient reason for rejecting all neutral monist versions. Therefore, a brief review of neutral monism and its rationales would shed some light on this issue.

1.7 Neutral monism

Bertrand Russell was the first who adopted the name 'neutral monism' in order to define his own view. However, in what follows, 'neutral monism' is used to refer to any view committed to the minimal thesis of the existence of a neutral stuff underlying the whole reality. In terms of properties, the neutral monist thesis is:

(N) Any property, either physical or mental, would be dependent on and, therefore, accountable by basic neutral properties, which are neither mental, nor physical.

As I have already stated, by 'basic properties' it may be meant either properties which made up the subvenient base of all other properties or, as a particular case, properties featuring in the reductive base (or the realization base) of some high-order properties.²¹ Even though all versions of neutral monism hold that the mental and the physical are reducible to the neutral, there is not always a clear commitment to ontological reducibility. Sometimes, 'reducible' must be read in an explanatory sense. Therefore, in what follows, I shall take neutral entities presumed in (N) to provide a basis for accounting for all other entities, whether mental or physical, under the

²⁰ See Levine (2001: 25).
²¹ I take realization to be a particular case of supervenience. See Addendum 1.2.

assumption that a consistent account of some property can be rendered only by identifying what property base necessitates its instantiation and in virtue of what laws.

There are two main rationales one usually uses in defending neutral monism. The first (we may call it 'the argument from intrinsic nature') relies on the thesis that any account given by any scientific theory built up in the traditional way will make use of relational, structural properties. However, as the argument proceeds, it cannot be that the properties inhering in the basic entities are all relational properties; admittedly, one would expect that the features of our basic ontology would be at least in part intrinsic. Therefore, the theory able to provide ultimate and complete accounts should be a theory of a new kind, its concepts referring to intrinsic and basic properties. If the mental and the physical properties are defined relative to the current kind of scientific theories, two consequences proceed from it. First, they are all relational properties and, second, they all must be realized by the properties posited by the new theory. It is in this sense that the basic properties should be understood as neutral, i.e. neither mental, nor physical, properties. The neutral properties should and would be mentioned in any description of the basic reality and constitute (realize) any highlevel property. Hence any mental or physical property will count as a high-level, relational or dispositional property being token-identical with a pattern of basic neutral properties.

Note that the key point in the above argument is the thesis of relational properties. However, why would anyone hold that the current scientific theories are able to render accounts only in terms of relational concepts? And, moreover, what could be

the reason for making the same claim about any future theory conceived in the same manner as the current ones?

A possible line of reasoning would basically run as follows. Arguably, what any scientific theory aims to achieve amounts to accounting for the data (or the empirical evidence) regarding a certain domain by appealing to a set of laws that correlate instances of some properties counting as primitive for the ontology of the respective theory.²² The interesting thing about the available data is that they are not only what for the theories aim at accounting, but they are also guidance in choosing what kind of primitive properties the theories posit. If the data are acquired in a mediated way, that is, if they derive from our knowledge of some interactions which affect our "external" senses, the properties depicted by those data are only relational in kind. This way of acquiring data is also known as the third-person perspective. Almost everybody agrees that this kind of data makes up the empirical base of any current science. A direct consequence would be that, as long as the empirical base of the theory comes exclusively from the third-person perspective, its primitive properties will be always relational in kind.

But, if not by the third-person perspective which provides us with public data and hence with a ground for objectivity, how else are we supposed to acquire data in

²² Here, being primitive should not be taken for *being basic*. The former marks what is elementary in the ontology of a scientific theory, the latter signals what there ultimately is. An identity between the two kind kinds of properties is acquired only when the respective scientific theory is the fundamental one, i.e. it provides the complete and ultimate accounts of any phenomenon.

Note that, unlike a high-level theory which works in its explanations with observable entities, a fundamental theory assumes unobservable (usually called: 'theoretical') entities. Paradigmatic examples of theoretical entities are the micro-physical ones.

order to elaborate a scientific theory? Russell²³, one of the champions of the intrinsic nature argument, points at the first-person perspective. His 'knowledge by acquaintance' refers at our direct way of accessing the intrinsic features of the entities. The data we acquire from our phenomenal experiences are data about the intrinsic nature of our intimate reality.

Actually, Russell claims about phenomenal properties more than their intrinsic character. He maintains also that they are basic properties, realizing psychological, neurological and micro-physical properties.²⁴ Russellian monism, in its full-fledged form, is the commitment to the thesis that all entities have the same intrinsic nature in the sense that they are realized by the instantiation of some phenomenal properties. He says:

The data of psychology do not differ in their intrinsic character from the data of physics. I have maintained that sensations are data for psychology and physics equally, while images, which may be in some sense exclusively psychological data, can only be distinguished from sensations by their correlations, not by what they are in themselves. (1921: 297)

It is for us to note that if one grants, as we have done, that phenomenal properties are *par excellence* mental properties, the charge of panpsyhism bears on Russellian monism. However, one may not agree with the Russelian thesis that phenomenal properties are the basic properties of our world, and yet sympathize the intrinsic

²³ See Russell (1921, 1927).

²⁴ Russell appears to defend the *possibility*, to put it mildly, that all physical entities – more exactly, events – have an intrinsic nature of the same kind as mental entities. According to Russell (1927), what pleads for such a hypothesis is that it helps us a lot to explain our cognitive relationship with the external world: "There is no theoretical reason why a light-wave should not consist of groups of occurrences, each containing a member more or less analogous to a minute part of a visual percept. We cannot perceive a light-wave, since the interposition of an eye and brain stops it. We know, therefore, only its abstract mathematical properties. Such properties may belong to groups composed of any kind of material. To assert that the material must be very different from percepts is to assume that we know a great deal more than we do in fact know of the intrinsic character of physical events. If there is any advantage in supposing that the light-wave, the process in the eye, and the process in the optic nerve, contain events qualitatively continuous with the final visual percept, nothing that we know of the physical world can be used to disprove the supposition." (pp. 263-4)

nature argument, and being committed to its conclusion that there is a basic reality, neither mental, nor physical. In other worlds, one may be convinced by the intrinsic nature argument that there is a basic reality, whose intrinsic properties are not accessible from the third-person perspective, and one still holds coherently that the basic reality realizes both phenomenal properties and physical properties. Maybe we have a direct and accurate access to phenomenal properties, which are intrinsic properties, but there is a strong possibility that we do not have any kind of access to the basic reality, i.e. the reality responsible for both phenomenal properties and physical properties and physical properties.

On the other hand, it is true that the conclusion of the intrinsic nature argument is compatible with the neutral monist thesis, (N). The idea that the primitive properties of our present fundamental theories are not basic opens the possibility that there is a basic reality of a neutral nature that realizes what we take presently to be both physical properties and mental properties. However, the compatibility does not amount to an implication: the argument is not conclusive for the thesis that there is a shared base of realization for both mental properties and physical properties

of a basic reality of a neutral nature. Maybe the accounts given by our present theories make use of relational properties and, moreover, basic properties cannot be all relational properties, but these considerations do not lead to the conclusion that. One may be persuaded that none of our traditional theories²⁶ can provide us with the

²⁵ For example, Chalmers (1996) is in agreement with the idea that micro-physical properties are not basic since they are defined in a relational (or dispositional) way, and suggests that both phenomenal and physical properties might be realized by a third kind of properties, "protophenomenal properties."

²⁶ By 'traditional theories' we should understand here theories built up on the empirical evidences available for the third-person perspective.

so-much-needed basic ontology. However, this could only be an innuendo regarding the unbeatable limitation of our scientific theories and not a reason to believe that we may find a new science positing intrinsic properties. Moreover, even if such an advance in scientific knowledge were possible, why could not be two separated and at the same time equally fundamental theories, each of them positing a distinct ontology populated by basic entities (hence having intrinsic properties)? There is no guarantee that reality meets our desideratum of unification; the basic reality could still be split in two or more kinds of intrinsic entities (or properties), in conformity with the dualist or pluralist doctrine. Finally, the whole notion of neutrality assumes a mentalphysical distinction issuing from the theory approach. But, as we have seen, the approach is objectionable. Then, to what degree does 'neutral' make sense if 'physical' is not defined relative to a certain (traditional) theory? I shall return to the point in the next subsection.

If correct, the above analysis shows that the intrinsic nature argument is not conclusive for inferring the need for a neutral reality of a single kind. Yet, there are some people wanting to take another line of reasoning to the effect of concluding the existence of such a reality. The starting assumption is that there is a relation of mutual necessitation among the mental properties and the physical properties involved in the mind-body relationship. Then, granting the multiple realizability argument against mental-physical type identity, the suggestion is that there is a basic reality whose properties necessitate (possibly, in virtue of constituting a common realization base) both mental and physical properties. The claim is that if a realization/subvenient base supports the instantiations of two different properties, then, indirectly or relative to this base, there is a mutual necessitation between the
two high-level kinds of properties. The formula is intended to offer an alternative to the thesis that mental (phenomenal) properties are realized in the physical ones – which is the ontological ground for the program of conceptual reduction of the mental to the physical that proceeds via a functionalist analysis of mental concepts.

For instance, Nagel (2000) suggests that there is a deeper reality, neither mental, nor physical, but entailing both of them. Thus, claims Nagel, in virtue of having the same sufficient base that realizes both kind of properties, mental and physical necessitates each other. If an object sustains mental life (phenomenal experiences) in our world, this is because there is a basic reality, whose properties realize phenomenal properties. However, the same basic properties realize also physical/physiological properties inhering in the respective sentient body. Nagel says:

"[T]he hypothesis would be not that the physiological state causes the phenomenological, but that there is a third term that entails both of them, but that is not defined as the mere conjunction of the other two. It would have to be a third type of variable, whose relation to the other two was not causal but *constitutive*. This third term should not leave anything out. It would have to be an X such that X's being a sensation and X's being a brain state both follow from the nature of X itself, independent of its relation to anything else." (p. 458)

However, if there would be something that constitutes both X's being a sensation and X's being a brain state, why were X a third type of variable, i.e. neutral, neither mental, nor physical? Why would X not be physical, for example? I believe that there is no reason for saying that what realizes physical is not still physical, but something of a different type.

Besides the difficulties one encounters in defending neutral monism, the hardest is probably to fix the content (the meaning) of 'neutral'. What makes problematic the notion of 'neutral' is our difficulty in conceiving something capable of constituting both phenomenal experiences (or minds or subjects) and non-sentient objects. More general, there is a difficulty in conceiving how a sentient object could be a constituted object. On the one hand, if neutral stuff is devoid of subjectivity, how can unminded entities made up one having a mind, i.e. a subject? On the other hand, even if the panpsychist view is true, it is not less difficult to see how little subjects come to arrange themselves in a coherent sentient entity.

However, I do not take the constitution difficulty to be one that should be addressed only by neutral monism. Instead, what appears to be a specific question for this view is where the domain of physical entities ends and the domain of non-physical neutral constituents begins. What makes hard to accept any border between the physical entities and the non-physical constituents of them (again: basic neutral entities) is that the neutral monism does not provide us with any criterion: we do not know what makes some entity to be physical, and, at the same time, what is missing from its constituents such that they are only neutral.

In the same vein, Nagel acknowledges the radicalism of the idea that:

there is something more fundamental than the physical—something that explains both the physical and the mental. How can the physical be explained by anything but the physical? And don't we have ample evidence that all that needs to be postulated to get ever-deeper explanations of physical phenomena is just more physics? (p. 469)

Granting that the entities posited by a scientific theory have the same nature, what seems to alarm Nagel is that, once a physical entity is explained by other entities in virtue of their constituting the physical entity, there is nothing to prevent us from considering the constituting entities as being physical as well. We may conclude that if the neutral monist doctrine does not collapse in panpsychism – such an example is Russellian monism –, by defining neutral properties as the shareable base of realization both for micro-physical and phenomenal properties, neutral monism and its thesis, (N), is nothing but disguised physicalism.

Addendum 1.1 Why properties, instead objects?

One may wonder why some scholars, like Jackson, Stoljar, and many others, are concerned only with physical properties, states, or events, not bothering with physical objects (or substances). One likely answer would be that, on their view, all concrete objects are physical. It seems to me that this is an unfortunate prejudice of our times; it is not reasonable to address the distinction mental-physical, but still ignoring the alternative proposed by substance dualism – the initial engine of such a distinction.

However, one may adopt a more charitable attitude in interpreting Jackson's view. It might be held that, in order to understand the notion of 'physical object' it is enough to master the notion of 'physical property': a physical object is nothing else than an object having physical properties. However, this proposal probably would work fine provided that all objects having physical properties had only physical properties.²⁷ But we cannot grant such a thing because there are people – namely, the property dualists – holding that some objects do not have all their properties physical. Why should we call such objects 'physical'?

²⁷ Arguably, it is the kind of objects Jackson suggests us to look at in order to define 'physical property'. I shall say more about this strategy in a moment.

The main reason – whether or not made explicit – would be that the physical properties are taken to be primitive (or basic) relative to the mental ones. That is, the mental properties are dependent on physical ones, the latter either realizing or simply necessitating the latter. But these are theses of a physicalistic breed. The direct consequence is that the distinction mental-physical is usually drawn on the physicalistic assumptions. Could we say that it looks like a fair physicalist-antiphysicalist debate on the issue of the mental-physical relationship?

Maybe some people would say: 'Well, there is no debate between the defenders and contenders of the physicalist doctrine anymore, but rather a debate among physicalists – namely, those scholars involved in a coherent program and working under some common ground physicalistic assumptions.' We should admit that as a matter of fact those people would be right. However, let us not mix up matters of fact with matters of principle and try to draw the mental-physical distinction in such a manner that is open to any view present on the battlefield of the dispute.

Thus, let us save the phrases 'physical object' and 'mental object' only for those objects (if any) whose properties are all (but not necessarily only) physical and, respectively, mental.²⁸ What about the rest of objects? Actually, I have already introduced the term 'mixed entity' for referring to any entity having in its constitution entities belonging to different kinds or species.²⁹ (Presumably, a human being is a good example of a mixed entity.) Given that an entity could be constituted even of

²⁸ As it has been made plain in the previous section, on some views, an object could be both mental and physical in the sense that any member of specie is also a member of the respective genus.

²⁹ Such a heterogeneity may be rendered by entities of different kinds (e.g., mental and physical – see dualism) or different species (e.g., mental and non-mental – see physicalism).

entities belonging to another ontological kind, an object not having all its properties of the same kind or specie will be called 'mixed object'.

Now, if physicalism proves to be true, then, from one perspective, any object is physical – as a feature inherited from its basic constituents. However, from another perspective, some objects are mixed because both mental entities and non-mental entities feature in their constitution.

But what if mental properties are not realized by physical properties? What if the physical properties just nomically necessitate the instantiation of mental properties? In that case, some objects are mixed because both physical entities and non-physical entities make up that the respective objects.

Addendum 1.2 Supervenience, realizability

Let us consider a property, *M*. We may call *M* a high-level property if it can be instantiated either by a non-basic entity (that is, an entity having a reducible constitution) or by a number of entities which do not aggregate one. For purposes of simplicity, let us consider only the first case, namely, that when the high-level property inheres in only one entity.

In the most general sense, we say that *M* supervenes on properties P_1 , P_2 ... from a lower level – that is, properties instantiated by the constituents of the entity in which *M* inheres – iff the instantiation of *M* is necessitated by the instantiations of P_1 , P_2

That is, for any entity, p, by matter of necessity, if p's constituents instantiate P_1 , P_2 ..., then p instantiates M.

Now, if *M* supervenes on properties P_1 , P_2 ..., one alternative is that *M*'s instantiation is something *over and above* the instantiations of the properties from the subvenient base. The other alternative is that *M*'s instantiation is *absorbed by*, or *realized in*, the instantiations of P_1 , P_2

We say that *M* is type identical with properties P_1 , P_2 ... iff all instantiations of *M* are realized by the instantiations of P_1 , P_2 If *M* is multiple realizable, that is, if *M* is realizable not only by one set of properties, there is only a relation of token identity that holds between *M* and P_1 , P_2 Given that realization is a particular case of supervenience, multiple realizability is a particular case of supervenience where a property can have more than one subvenient base.

It is also possible to have one subvenient base for two different high-level properties. In such a case, the base necessitates them both. Some authors take such an alternative to be a way of showing indirectly the necessity between the two properties. The idea is the following. If *N* necessitates *M* and *N* necessitates *P* then, given *N*, *M* is necessarily correlated with *P*.

Chapter 2: Stating the Problem of Causal Exclusion

According to what kinds of entities there are held to ultimately be. Chapter 1 has briefly introduced the main contenders to the mind-body problem: monism, dualism, and pluralism. In this chapter, § 2.1 takes a closer look at the dualist doctrine, and in particular at its interactionist version, by exposing its core theses. In § 2.2, it is presented a common scenario where mental events causally interact with bodily events. In the next subsection, the Causal Exclusion Problem (CEP) is stated as an alleged inconsistent set of assertions regarding the protagonists of mental causation. However, in § 2.4, it is emphasized a cleavage of the theoretic field into two opposite camps, the one that grants the legitimacy of CEP (exclusionism) and the other that contests it (causal compatibilism). Taking the exclusionist stance, that is, claiming that the truth of the CEP assertions leads to an implausible systematic and widespread overdetermination, the physicalist uses CEP in order to produce an argument against interactionism (the CEP argument). In § 2.5, it is presented the general structure of such an argument, and, in § 2.6, it is exposed the strategy I shall follow in trying to block the physicalist attack. In order to make the physicalist objection more vivid, the chapter closes with a particular version of the CEP argument.

2.1 Interactionism

To put it bluntly, the dualist doctrine holds that there are two basic ontological realms: mental and physical. By 'basic ontological realm', it is meant a domain whose entities – whether particulars or universals – are basic or primitive. Recall that a particular is basic if no other particulars (of the same ontological kind) feature into its constitution. In other words, a basic particular has no parts; it is simple. For instance, a sub-atomic particle is basic if it is not made up of smaller particles. Instead, we think of a property as being basic if its instantiation does not depend on the instantiations of other properties. Thus, properties realized by other properties are not to be included among basic properties. All in all, it can be said that a basic ontological realm contains a "slice" of the ultimate realities of our world.

Descartes (1642) held that there are two basic realms – mental and physical. His believing in the existence of two kinds of substances led to his doctrine being called *substance dualism*.³⁰ This has been traditionally distinguished from *bundle dualism*, a view attributed to Hume (1739). If Descartes thought of mind as a substance in the scholastic sense, namely something that possesses properties, but it is over and above its properties, Hume, instead, suggested that the mind is nothing but a collection or succession of mental episodes or occurrences (probably, instantiations of mental properties). Thus, the difference between the two kinds of dualism stems from the existence of two distinct conceptions about substances: the substratum-attribute view and the bundle theory of the substance. However, on both views, a basic substantial particular (individual) belongs to the realm it belongs to according to

³⁰ Examples of contemporary philosophers who argue for substance dualism, but not necessarily of a Cartesian sort, are Foster (1989, 1991), Lowe (1996, 2000), Robinson (2003a, 2003b).

what kind of basic properties – either mental or physical – it possesses.³¹ For instance, on the Cartesian view, my mind is a basic mental substance, and it possesses only mental properties. The mental property of *being in pain* may be one of them, and its instantiation is nothing but the mental state of *my mind's being in pain*.

The suggestion is that substance dualism requires the separateness of basic mental properties from basic physical ones. There is also *property dualism* that holds the existence of two basic and hence separated kinds of properties, even though under the proviso that, in the actual world, mental properties are always co-instantiated with physical properties.³² All in all, the thesis that urges the separateness or distinctness of basic mental properties from basic physical ones is a core dualist thesis.³³

One should be aware that the dualist thesis of distinctness makes a claim of nonidentity in the strongest sense; it rejects not only any type-identity between mental

³¹ It is an open question for the dualist *whether* a basic substance instantiates mental properties in virtue of its being mental (or its having a mental nature) *or*, the other way round, a basic substance is mental because it instantiates only mental properties. If the former is true, then, for a substance, to be mental is to have such a nature that requires only mental property instantiations. If the latter is the right alternative, then to be mental is nothing but to have only mental properties. The difference amounts to the distinction between being necessarily and being contingently mental.

³² That is, any individual exemplifying a mental property should also exemplify a physical property. Of course, the requirement that any mental property should be co-instantiated with some physical property is a concession to physicalist ideology. This is why it is not property dualism, but only substance dualism, that is also called *proper dualism*. The co-instantiation requirement may seem *prima facie* to be in conflict with the original claim of dualism, namely, that there are basic mental properties. For, by definition, the instantiation of any basic property does *not depend* on the instantiation of other properties. But, it should be noted, by 'no instantiation of property X does not depend on any instantiation of property Y', it is not mean that no instantiation of X is attended by any instantiation of Y, but rather that no instantiation of X is realized by any instantiation of Y. Moreover, co-instantiation is not a relation of necessity, but rather a contingent relation holding between instances of properties inherent in the same entity.

³³ In the crusade against the idea of reducibility of mental properties (especially, the phenomenal properties of conscious experience) the knowledge-type arguments have a major contribution. See Robinson (1982), and Jackson (1982, 1986). More recently, Chalmers (1996) brings into debate the conceivability-type arguments initiated by Descartes and, at same point, by Kripke (1980).

and physical properties, but also any realization relation and hence any token-identity between the instances of these properties.

Now, if one embraces the dualist thesis of distinctness, one is left with the question of the relationship between the two posited basic realms. If there is no type-identity between the two kinds of properties, no bidirectional metaphysical necessitation relation holds among their instances. What about supervenience relations?

Supervenience relations involve a one-way necessitation from subvenient base properties to all of their supervenient properties.³⁴ We say that property *M* supervenes on properties P_1 , P_2 , ... when, for any entity *p*, by matter of necessity, if *p*'s constituents instantiate P_1 , P_2 , ..., then *p* instantiates *M*.

It follows automatically that substance dualism is incompatible with any relation of supervenience between the mental and the physical. The reason for incompatibility is simply that substance dualist doctrine implies the existence of two kinds of substances and hence none of the substances of the one kind is constituted of substances of the other kind; this goes in conflict with the idea of supevenience, which presupposes that the subvenient properties are instantiated by the constituents of the entity in which the supervenient property inheres. Thus, the idea of supervenience is interesting only for the property dualist.

³⁴ Since there could be more than one supervenient property, a subvenient base could necessitate the instantiation of more than one property. See Addendum 1.2.

By far, the most popular case of supervenience is the one where the supervenient property instances are realized by the instances of the respective subvenient properties. But, a realization relation is nothing else than a token-identity relation between properties, and, as I have shown, the content of the dualist thesis of distinctness is incompatible with the assertion of any identity between the mental properties and the physical properties.

It is not worthless to say here that, regarding the relationship between the mental properties and the physical properties, the doctrine that takes it to be a relationship of realizability, nay, of multiple realizability is non-reductive physicalism. It is important to note that non-reductive physicalism is *non-reductive* only in the sense that, on this view, even though mental properties are seen as realized in physical systems, given its presumed multiple realizability, any mental property is taken to be type non-identical with the physical property-pattern that realizes it. Therefore, non-reductive physicalism, in fairness, is as reductive as any monist doctrine: the instances of mental properties are *made up of* instances of physical particulars. This is why, even though its having a reasonable stance against mental-physical type identity, non-reductive physicalism seems to be at most a *dualism of predicates*. Therefore, I shall stop short of considering it a (property) dualist view.

Next, let us consider a case of supervenience, where property M supervenes on properties P_1 , P_2 , ..., such that their instances do not stand in any relation of realization, that is, M's instantiation is not absorbed by, or realized in the instantiations of P_1 , P_2 , In other words, the considered case is one where the

CEU eTD Collection

instantiation of the supervenient property is something *over and above* the instantiations of the properties from the subvenient base. Now, if mental instances are over and above physical instances, the dualist thesis of distinctness is met and there is no reason intrinsic to dualist doctrine for rejecting the idea that mentality is supervenient on physicality.³⁵

However, the relation of necessitation requires further explanations. Philosophers usually distinguish between metaphysical necessitation (i.e. the relation holds in any possible world) and nomic or empiric necessitation (i.e. the relation holds just in those possible worlds where the laws actually underlying the relation hold).

I believe that, if one is sympathetic with the dualistic thesis of distinctness, one should discard the idea of metaphysical necessitation between mental and physical. The reason comes from the difficulty in conceiving such a relation between two distinct entities.³⁶

³⁵ By defining a basic property as one which does *not depend* for its instantiation on the instantiations of other properties (excepting those which belong to the same realm), it does not mean that its instantiation cannot be *necessitated* by the instantiations of other properties. It is the difference between necessary conditions and sufficient conditions for (coming into) existence. It is one thing to exist *only if* condition *C* is met (*C* would be a *necessary* condition), and another thing to exist *if C* is met (*C* would be a *sufficient* condition). The existence of a mental basic realm defended by dualism requires the instantiation of the basic mental properties to have no necessary conditions in the physical realm.

³⁶ Probably, the stance the dualist takes on the issue of the mental being metaphysically necessitated by the physical (more precisely, the issue of whether certain physical configurations may be sufficient conditions for the instantiation of certain mental configurations in any possible world) seems to depend on the motives one has to commit oneself to dualism. For instance, if one is attracted by dualism due to one's sympathy for the idea of the possibility of zombies, then one will also take a stand against the thesis that physical metaphysically necessitates mental (see Chalmers, 1996). Instead, if one's main reason for adopting a dualist view is just the appealing for the idea of the possibility of mind's existing without any body, then one has no reason related with one's dualist convictions to be against the idea of a mentality metaphysically necessitated by physical configurations. Instead, what would be mainly rejected in this latter case is the thesis that a necessary condition for the existence of a mind is a body or a world populated with material objects in order to join or attend (not realize!) mental phenomena.

However, given that the notion of nomic necessitation enjoys enough popularity, nothing prevents us from thinking of supervenience in terms of nomicity. The idea of nomic relations holding between the mental and the physical seems to be supported further by our common beliefs that there are causal interactions between the two realms. If one believes that mind has a causal role in bringing about bodily events and, vice versa, that some physiological phenomena have psychical effects, then, plausibly, one has to appeal to psycho-physical laws (nomic necessities) in accounting for such mutual causal actions.

Traditionally, any substance dualist version acknowledging the reciprocal causal relationships between mental and physical has been referred to by the term 'interactionism'. Even though the Cartesian view is the paradigmatic example of interactionism doctrine, in what follows, regardless what concepts of *substance* and *causation* one has in mind, I shall take the interactionism to be a doctrine that entails a conjunction of two theses:

(D) Dualist thesis: Mental properties are not token-identical with physical properties;

(C) Causal thesis: Mental phenomena and physical phenomena can causally affect one another.

2.2 What would mental causation look like?

Let us consider the following imaginary, but still not very uncommon, scenario. I am in a bus station and hear a very familiar voice coming from behind. Believing that it is my good old friend, Dan, the person whose voice I hear, and wanting to check out whether he is really there, I form the intention to turn my head in order to see who is talking over there. Consequently, I turn my head, expecting to recognize Dan's face. It seems entirely reasonable to say that the situation I have just described is a succession of events/states³⁷ related in a causal manner. Moreover, some of these events are mental, others are bodily or behavioral (terms I use interchangeably); the causal chains seem to travel from the bodily level to the mental one and the other way round. Briefly, it is a case when the mental is "intermingled" with the body in a causal fashion: there is a causal interaction between the mind and the body in the described scenario. On the bodily-mental route, for instance, the sound that strikes me causes me to recognize a certain vocal pattern.³⁸ On the mental level, my hearing the sound, my recognizing Dan's voice together with other relevant beliefs³⁹ bring about my believing that Dan is just behind me and, plausibly, further mental events like the excitement that I feel about the unexpected opportunity of meeting Dan, and my intention of finding out whether I am right in supposing Dan's presence. Further, still on the mental level, my desire of meeting Dan and my belief that he is right behind me are at least two of the mental events causally responsible for the

³⁷ I shall work under the hypothesis that both events and states could have causal powers or, in more neutral terms, could matter in bringing about other events or states. For simplicity, I shall use the generic notion 'event' for referring either a proper event or a state.

³⁸ Of course, I do not want to say that the mentioned cause is the *only* cause of the mentioned effect. Nor is it in my intention to say that the mentioned effect is the *only* effect the mentioned cause has. All that I try to do here is to propose a plausible example of some incomplete causal chains.

³⁹ An example could be my "tacit" belief that, in normal conditions, hearing one's voice is an evidence that there is one around.

occurrence in me of the intention of turning my head. Finally, on the mental-bodily causal route, following my intention, I turn the head. (See Figure 2.1.)

MENTAL LEVEL



If we use '*m*' to denote the mental event of my intending to turn the head and '*b*' the behavioral event of my turning the head, it is very tempting to assume that *m* is a mental cause of b – that is, it is a case of mental causation. Let us see what metaphysical scenario would require it such an assumption.

Taking seriously neurobiology, we would not expect that a mental event is an unmediated cause of a behavioral event, but rather that such a causation is accomplished with the help of at least one neural event. Thus, let us consider a neural event – call it e – such that m's influence upon b takes effect through e. (See Figure 2.2.) Maybe it is not implausible to assume that e is directly caused by m. For

purposes of simplicity, let us also assume that there is no other mental event that causes *e* and is causally or ontologically (mereologically) unrelated to m.⁴⁰ For instance, besides my intention of turning the head to see who is speaking behind me, there is no such a thing as my having a pain in my neck that *reflexly* makes me to turn my head in order to relief the respective pain. It is not that something like that is not possible, but rather that, for stressing more strongly what is at stake here, I prefer to consider that *m* is the most comprehensive mental event causally responsible in an immediate way for the occurrence of *e*. Henceforth, *m*, *e*, and *p* will feature recurrently in the discussions to come about any case of mental causation.



Fig. 2.2

⁴⁰ That is to say that any other mental cause of *e* is either a cause of *m*, or a constituent of the more complex event *m*, or simply an effect of *m*.

2.3 The Causal Exclusion Problem (CEP)

It is a familiar point that there are two main kinds of phases required by any comprehensive advocacy of a certain doctrine (or theory) which addresses a given issue. Granting that the doctrine is already defined, its promoter should proceed either to defend it against the main objections, it encounters or to bring forward arguments that straighten its plausibility. Relative to the arguments, it became common to distinguish between positive and negative ones. A positive argument for a philosophical doctrine is supposed to provide it with a rationale independent of other contending doctrines. Instead, negative arguments play the card of comparative analysis. Sometimes, it is a presumed desideratum that, unlike the promoted doctrine, some of its contenders do not meet; other times, it is claimed a difficulty that stands as an objection for most competitors and leads to the elimination of some opponents, which leaves the respective doctrine in a select company, if not even solitary.

The modern materialist monism (expressed by the physicalist doctrine) is a main stream in the philosophy of mind that has occurred from dissatisfaction with substance dualism and nowadays it is the main combatant against any newcomer in the camp of property dualism. Many arguments brought by physicalists committed to the idea of the causal efficacy of the mental on the physical are just negative arguments. They aim to show that the dualist, unlike the physicalist, cannot make room into his/her doctrine for the phenomena of mental causation. In other words, what is common for this kind of physicalistic argument is their pointing to the failure of

any theory claiming for mental events both a different nature from physical ones and the capability of causally interacting with the physical domain.

My dissertation focuses on a contemporary version of these arguments, which is based upon what is usually known as 'the Causal Exclusion Problem' ('CEP' for short): the claimed conflict between the idea of a causally potent mind with respect to the occurrence of some physical phenomena and the physicalist hypothesis of the causal completeness of the physical domain.⁴¹ The conflict would be due to an alleged inconsistent set of assertions:

(NI) Mind-body causal non-identity: Mental causes are not identical with physical causes;

(MC) Mental causation: The mental is able to cause physical occurrences;

(PC) Physical closure,

where (PC) involves at least that, for any physical occurrence, there was a physical cause or some set of physical causes sufficient for bringing about the respective occurrence.

The inconsistency of the CEP assertions is claimed to stem from their allegedly entailing of a *highly implausible situation*. The truth of the CEP assertions implies the notion that *each* physical effect involved in mental causation is overdetermined. However, it is held, rational constraints prevent us from admitting such a widespread overdetermination. In particular, it is implausible that each and every bit of action we

⁴¹ Malcolm (1968), Schiffer (1987), and Kim (1989, 1998) are among those who have initially raised the problem.

perform (or, more to the point, any neurological event involved by mental processes) is overdetermined.

Before closing this section, there is one more point that I want to bring into discussion. Given that there is no consensus about the nature of causal relata, we face the issue of what particular expression we are going to take for each of the CEP assertions. In terms of what is CEP to be construed? Or, is there any way of working with the CEP assertions without committing ourselves in particular to any analysis of causation?

I believe that one cannot undertake a substantial consideration of CEP and its consequences unless one takes a particular stance on the matter of causation itself. However, as a regulative methodological principle, it is undeniable that, in approaching CEP, one should try to keep one's discourse close to generally accepted claims (that is, truisms) about causation as much as possible. In this vein, let us note that, beyond any conception about the nature of causal relata, it seems to be widely acknowledged that properties make a causal difference. On the one hand, if one grants that causal relata are particulars (like objects, events and states), it would be of no wonder that one finds that the causal powers they have are in virtue of certain properties their existences involve. On the other hand, if one is committed to the idea that causal relations hold between universals, it is natural for one to see properties as regulating any case of causation.

Prima facie the idea suggested by the above considerations is that a (type) nonidentity in causal properties entails a non-identity of causes. Yet, given that two (type)

CEU eTD Collection

distinct properties can be token identical, it is too strong a claim to be made. Instead, it is plausible to assume that *the non-identity of two causes is entailed by the token non-identity of causal properties in virtue of which they are causes.* If this is correct, it follows that the dualist thesis of interactionism, (D), implies the first CEP assertion, (NI), where the causal non-identity between mind and body is propounded.

As will become more explicit in the following chapters, I consider that the protagonists of causal relations are events (including states). Yet, given that I take events to be instantiations of properties by objects in the Kimian style, there should not be considerable disagreement between the view I adopt and any other one committed to objects as causal relata as long as the objects are seen to be causally empowered by some of their properties.

For instance, let us consider that in a case of causation are involved three objects, x, y, and z. According to Horgan (1989), to say that x and y are some causes of z is to say that there are certain properties, P, Q, and R, such that x qua P and y qua Q cause z qua R. That is, x in virtue of having P and y in virtue of having Q cause z to have R. Then, bottom line, in identifying a cause in the considered case, the difference between pointing to *the instantiation of P by x* (or, better, x's *being P*), say, like I do in the Kimian style, and pointing to x qua P, like one as Horgan does, is just a matter of different emphases made in the same metaphysical structure, the one which is both the truth-maker of 'x is P' and causal responsible for the occurrence of the truth-maker of 'z is R'.

Further and more importantly, on both philosophical views, the non-identity of causes should amount either to the non-identity of the objects $(x \neq y)$ or, if the objects are identical, to the non-identity of the instances of the causal properties. But, then, the non-identity of causes is assured by the non-identity of property tokens. That the tokening of *P* in *x* is different from the tokening of *Q* in *y* implies the difference of causes either in virtue of there being *different objects* causally empowered or in virtue of there being a single object endowed with distinct causal powers.

2.4 One problem, different reactions and employments

According to Bennett (2003), there is considerable disagreement among philosophers who are concerned with mental causation whether the CEP assertions make up an inconsistent set in the sense of implying the unacceptable idea of a systematic and widespread overdetermination. To be precise, regarding the legitimacy of CEP, there is a cleavage of the theoretic field into two opposite camps: "exclusionism" and "causal compatibilism". Those from the latter camp – the compatibilists – want to keep all of the CEP assertions – "completeness, the distinctness of the mental and physical, and the causal efficacy of the mental." (p. 473) The exclusionists, in contrast, believe that the CEP assertions lead to a false metaphysical proposition which sets a logical tension among the three assertions, a tension to be relieved only by rendering one of them false:

[CEP] can obviously be dissolved by denying one of the primary claims that generate it. We could deny the distinctness of the mental and physical ... Or we could resign ourselves to denying the efficacy of the mental... And we could in principle deny the completeness of physics ... (p. 472)

Thus, the exclusionist may use CEP in order to make a case for epiphenomenalism: the physical does all the causal work, and there is nothing left for the mental to do – that is, a denial of the second CEP assertion.⁴² Note that (MC), the second CEP assertion, is not exactly (C), i.e. the causal thesis entailed by the interactionism, only the latter exhibiting the commitment to the occurrence of causal processes in both directions, including from physical to mental. But, since (C) implies (MC), by denying (MC), interactionism is also rejected.

However, the exclusionist may use the same problem for physicalistic purposes as well. Being under the belief that the CEP assertion set is inconsistent, and being keen both on the causal efficacy of some mental phenomena and on physical causal completeness, the physicalist may argue against the first CEP assertion which claim the distinctness between mental causes and physical causes involved in the cases of mental causation. Given that, as I have mentioned, the first assertion of CEP is implied by the dualist thesis, (D), it would follow, by *modus tolens*, that (D) is false and hence that any dualist doctrine is to be rejected.

Reaching this point, I would like to call attention to a style of spelling out the above physicalistic argument, which, even though often used, is still objectionable for logical reasons. Here is how Levine (2001) argues:

For my part, the materialist case essentially rests on the phenomenon of mentalphysical causal interaction. ... It seems overwhelmingly obvious that mental phenomena are both causes and effects of non-mental, physical phenomena. What's more, within the realm of non-mental physical phenomena, the hypothesis that what determines the distribution of matter and energy is *exclusively* determined (to the extent there is determination) by non-mental, physical forces, seems very well confirmed. ...Thus only if mental phenomena are somehow constructible from, or

⁴² In Jackson (1982) can be found a defence of the idea that mental events are caused by physical events, but have no physical effects. See also Chalmers (1996: 150-160, 192-203).

constituted by, the physical phenomena that serve as the ultimate causal basis for all changes in the distribution of matter and energy does it seem possible to make sense out of mental-physical causal interaction. (p. 5, my italics)

In the same vein, consider also:

[W]e know that changes in the distribution of mass-energy in spacetime are *only* caused by other such changes, and since mental states cause (and are caused by) such changes it has to be that they do so by way of physical mechanisms. Discovering the mechanisms involves discovering *how physical states realize mental states*, for otherwise it wouldn't be clear why these physical events constitute the mechanisms we're looking for. (p. 36, my italics)

It appears that a crucial assumption of the above argument is that:

(α) "what determines the distribution of matter and energy is *exclusively* determined by physical forces".

It expresses the claim that the physical domain is *causally* closed. Note that the idea of causal closure of the physical domain is the strongest and the most aggressive understanding (interpretation) of (PC).

Nevertheless, it is not the aggressiveness (and unjustified boldness) of the interpretation that should bother us so much. Rather, the worry is that, taking that particular interpretation, the assumption of physical causal closure is used in an argument whose bet is to prove that interactionist doctrine, unlike physicalism, cannot account for mental causation on pain of inconsistency. Given that a domain is causally closed if and only if there is no element from the outside able to causally interact with any element from the inside, to assume such a closure of the physical domain is nothing else than to assume the falsity of interactionism. Thus, to make use of such an assumption within an argument whose conclusion is against interactionism amounts to begging the question against interactionism. In other words, you cannot reasonably assume that the distribution of matter and energy is

exclusively determined by physical occurrences as long as what is at stake (to be proved) is exactly whether or not the distribution of matter and energy is *also* influenced (determined) by non-physical occurrences.

It is important to note that one can mistakenly take the above criticized argument for the one advanced by Kim against non-reductive physicalism.⁴³ As I have said, even though the non-reductive physicalist doctrine promotes a dualism of predicates, its defenders should not be considered dualists. Why is that? On the one hand, any genuine dualist doctrine (regarding either substances, or properties) rejects the thesis that mental properties are realized by physical properties, that is, that the instances of mental properties are *made up of* instances of physical properties. On the other hand, the non-reductive physicalists go only half way in departing from the full-fledged physicalism: they are against the property type-identity thesis, but, unlike the dualists, stick to the property token-identity thesis.

Thus, Kim, assuming the truth of the physicalism, in arguing against the nonreductive physicalism, is free to assert, against interactionism, the physical causal closure, a specific thesis of minimal physicalism:

There is a further assumption that I believe any physicalist would grant, namely <<the causal closure of the physical domain>>. Roughly, it says this: *any physical event that has a cause at time t has a physical cause at t*. This is the assumption that if we trace the causal ancestry of a physical event, we need never go outside the physical domain. To deny this assumption is to accept the Cartesian idea that some physical events need nonphysical causes, and if this is true there can in principle be no complete and self-sufficient physical theory of the physical domain. If the causal closure failed, our physics would need to refer in an essential way to nonphysical causal agents, perhaps Cartesian souls and their psychic properties, if it is to give a complete account of the physical world. I think most physicalists would find that picture unacceptable. (1993c: 280)

⁴³ See, e.g., Kim (1993b, 1993c, 1998).

Let us consider one assertion made in the course of the above argumentation:

(β) "if we trace the causal ancestry of a physical event, we need never go outside the physical domain".

As it appears, Kim takes it to express the idea that the physical domain is causally closed. However, one may find (β) to be a weaker assertion then (α), by which Levine conveys the true meaning of the notion of causal closure of the physical domain. For one thing, (β) may be understood as speaking only about the *sufficiency* of the physical in affecting any physical event, whereas (α) clearly speaks about the causal *exclusivity* or *exclusive sufficiency* of the physical domain. By (α) it is held not only that, for any physical event, there are physical events causally sufficient for it, but also that the physical domain is the only "place" where one may find a cause for a physical event.⁴⁴

Further, I shall use, at some length, for the assertion of the physical closure the working version introduced in section § 2.3 – namely, (PC): *for any physical occurrence, there was a physical cause or a set of physical causes sufficient for bringing about the respective occurrence*. Its main merit (advantage) is that it does not imply by itself the falsity of interactionism.⁴⁵ However, it should be stressed that, even if the physicalist chooses the above interpretation for the assertion of physical closure, he is called to back it up with some plausible reasons.⁴⁶ In Chapter 7, I shall

⁴⁴ Actually, (α) expresses only half of the idea of causal closure, the other half being that no physical event causes a non-physical event.

⁴⁵ Instead, as Kim rightly notices, its negation implies the truth of interactionism.

⁴⁶ After all, if the exclusionist is not very found of physicalism (or, at least, he is not willing to buy the full option physicalist offer), in order to relief the annoying (alleged!) inconsistency pointed out by CEP, he may be tempted to discard the third CEP assertion.

analyze some reasons one may try to bring as supports for various understandings of (PC).

2.5 The argument from CEP against interactionism

As I have mentioned, what gives a certain credit to the idea that CEP is a genuine problem is the belief that the CEP assertions make up an inconsistent set. Its inconsistence would be given by the conflict between their truth and the implausibility of the view that any case of mental causation is one of overdetermination. In other words, the truth of the CEP assertions seems to contradict the largely shared belief that the commitment to a widespread and systematic overdetermination is not reasonable. The idea is that any account of mental causation is more preferable to one requiring a commitment to a widespread and systematic overdetermination.

By 'systematic overdetermination' it is meant that all causal relations of a certain kind lead to overdetermined effects. For example, in the case of mental causation, there is systematic overdetermination if, whenever a causal relation of the psychophysical sort holds, the effect is overdetermined. Next, by 'widespread and systematic overdetermination' it is meant that the causal relations characterized by systematic overdetermination occur with a considerable frequency and in a large number of cases. Thus, if our behavioral events were overdetermined, i.e. if each and every time we move intentionally our bodies a case of overdetermination were exemplified, then mental causation should be described as being a widespread and systematic overdetermination.

It is almost unanimous that an acceptable view on world can not make room for systematic and widespread overdetermination. For instance, it is said that postulating systematic and widespread overdetermination would be "absurd" or "extremely odd" (Kim, 1993a: 281) or, at least, would "look suspiciously *ad hoc*" (Lowe, 2000: 572)

In other words, the acceptance of systematic overdetermination on a large scale would be in conflict with a principle of metaphysical parsimony or economy, which is a principle of rationality of the same kind as Ockham's razor. If the parsimony expressed by Ockham's razor regards the entities one posits in order to explain anything, the parsimony I have in mind here regards any metaphysical account of the world: one should not admit a metaphysics implying a world characterized by an oversized structure. That is to say that one should avoid a picture of the world like a divice whose components are functionally oversized.

Thus, in a nutshell, CEP amounts to the following:

- 1. The truth of the CEP assertions implies that any physical effect obtained by mental causation is overdetermined.
- 2. It is irrational (unreasonable) to hold the existence of any systematic overdetermination.
- 3. Therefore, at least one of the CEP assertions is false.

Going further, we may spell out in general terms the argument from CEP against interactionism (the CEP argument, for short) as follows:

4. There are reasons for holding the physical closure;

- 5. Therefore, we have to give up either the non-physicality of the mental causes or their causal efficacy on physical phenomena or both.
- But, interactionism requires both mind-body causal non-identity and mental causation on physical;
- 7. Therefore, interactionism is false.

My dissertation focuses on the possibility of rejecting the above argument, and the reply I forward mainly consist in an endeavor to show that the implication claimed by line (1) is unwarranted. Usually, the implication comes as the conclusion of a line of reasoning that assumes a certain sufficient condition, (SC), on overdetermination. The reasoning line would be like this: considering a case of mental causation, the truth of CEP assertion assures the satisfaction of (SC) and hence the overdetermination of the effect; but, giving that the case of mental causation is arbitrarily chosen, it follows that each and all effects of mental causation are overdetermined. In a schematic way, the whole thing would look like this:

CEP assertions imply SC

SC implies overdetermination

:. CEP assertions imply overdetermination

The above reasoning is true of any case of mental causation.

:. CEP assertions imply a systematic overdetermination.

Let us note that the soundness of the above line of argumentation depends very much both on how one construes each of the CEP assertions (i.e. in what terms one chooses to express each assertion) and on what one understands by 'overdetermined effect'. It is also line (4) that puts a burden on the proponent of the argument. Depending on what reading is used for the notion of the physical closure, the burden is more or less manageable. However, the physicalist who uses the argument should come up with such a reason for asserting physical closure that stands independently of the truth of interactionism.

The physicalist is not allowed to say that the physical closure (whatever it would mean) should be granted because, otherwise, we have to face the unacceptable alternative that there are some non-physical phenomena affecting in the causal manner the physical domain. For the alternative does not have the status of being unacceptable, but rather that of being under discussion, the purpose of the argument being that of rejecting it. This, however, can be done only in the conclusion, not in a premise.

2.6 The comeback compatibilist strategy

In the face of the CEP argument, it appears that a supporter of interactionism, who happens to be also fond of causal compatibilism or, at least, not willing to defend his/her favored doctrine at the cost of contesting the physical closure, should confine his/her comeback to a rejection of the first line of the argument. As far as I can see, this can be done by two different strategies. The first is to provide a counterexample to what the exclusionist proposed to be a sufficient condition for the overdetermination of an effect.

I am not very fond of this kind of strategy because it, if successful, stops the exclusionist offensive only temporarily, until the exclusionist comes with a new

CEU eTD Collection

proposed sufficient condition for overdetermination, good to be used in backing up the first line of the CEP argument. The compatibilist interactionist can escape from the exclusionist's objection once for all if and only if he shows that, as a *matter of principle* (that is, independent of any proof the opponent would like to bring into discussion), the implication (from the truth of the CEP assertions to overdetermination) is false or, at least, doubtful.

Surveying the literature, one may notice that any compatibilist show of the required kind (magnitude) has two guest stars intended to play the parts of an evil character and, naturally, of a hero. Of course, in a compatibilist spirit, the evil character is a necessary condition, (NC), for any effect to be overdetermined. Further, the hero is a relation, *R*, tokening among some of the causes featuring into the causal history of the effect such that (NC) is not met and hence killing any exclusionist hope for a case of overdetermination. The plot is simply as follows:

overdetermination implies NC

... Not-NC implies not-overdetermination

CEP assertions and R imply not-NC

:. CEP assertions and R imply not-overdetermination

That is, if *R* then *CEP* assertions imply not-overdetermination.

In other words, under the circumstances of *R*'s instantiation, the truth of CEP assertions do not imply overdetermination as the exclusionist holds, but rather its negation.

Thus, if the compatibilist is able to find the proper actors for the two parts, (NC) and R, s/he will manage to direct a play of a smashing success, bearing the name not

very different from 'The Death of the Beast'.⁴⁷ In the rest of my dissertation, given that I am enthusiastic about the above plot, but not very happy with the productions I have met until now,⁴⁸ I shall undertake the task of finding a potent cast that fits into the logic (semantic) of the play/plot.

It appears that the task of the compatibilist is more or less similar to that of the mathematician trying to solve a system of mathematical equations containing two variables or unknowns. The system of equation may be true only for certain pairs of the values of its variables. Such a pair is called a solution of the system.⁴⁹ In the simple cases, the natural way of finding a solution (if any) consists in following the so-

⁴⁷ In all fairness, this is not enough for killing, but rather for *not being* killed by the "beast". By showing that there could be a circumstance in which the truth of the CEP assertions does not amount to a case of overdetermination, the compatibilist manages only to eschew a mortal attack addressed to the possibility of having all the CEP assertions true. However, the constructive and *offensive* phase of pleading the truth of these assertions should bring on scene reasons that show more than the possibility of avoiding overdetermination (or, better, show more than that overdetermination is not unavoidable). The compatibilist should convince us not only that there is such a possibility, but rather that this is the way our actuality looks like sometimes. Since the ambitions of my dissertation are more modest, I confine here the scope of my argumentation to the primary phase, namely that of assuring the *defense*, which has its difficulties enough.

⁴⁸ One might expect here to be delivered a long list of names of people each associated with his/her contribution and seasoned with some plausible reasons for dismissing the respective enterprise. Actually, the list is not long at all, and I prefer to introduce the relevant authors in pages to come a bit later, where the specification of each of their solution has much more relevance for my plea. Moreover, the reticence of introducing most of them at this point is fueled by the fact that I know no compatibilist who explicitly puts forward the comeback strategy quite in the above form. True, it transpires somehow through many compatibilists' works. From my point of view, Bennett (2003) seems to be the closest to my approach. The following series of quotations expresses at least two ideas of those made explicit above:

a) The need of a necessary condition on overdetermination:

[&]quot;If compatibilism is to stand a chance, then, its proponents need to give us some genuine reason to think that the effects of mental causes do not count as overdetermined. They need to provide us with some sort of test, some way of deciding whether or not an effect is overdetermined." (p. 474) And "the test is not intended to provide a sufficient condition on overdetermination. ... I am only claiming that the test provides a necessary condition on overdetermination. That is all the compatibilist needs." (p. 477)

b) The solution of a relation holding among the elements involved in the causal history of an effect such that its instantiation precludes the satisfaction of a necessary condition on overdetermination:

[&]quot;In short, the physical causes that the compatibilist wants to say do not causally compete with mental causes are precisely those that are somehow tightly related to the mental causes. Thus it looks as though the compatibilist should indeed make some sort of appeal to a tight relation between the mental and the physical." (p. 476)

⁴⁹ Note that the set of solutions can be empty (there are no solutions), singleton (there is exactly one solution), finite, or infinite (there are infinitely many solutions).

called substitution method: try to solve one of the equations for one of the variables and substitute the result into the other equation, thereby reaching a single equation with a single variable, which (hopefully) can be solved.

If you like, the two *unknowns* of the compatibilist argument appear to be R and (NC). And, hunting for a *solution* here amounts to looking for a pair of *values* for R and (NC) that renders the compatibilist defence sound. Playing further with the metaphor of equations, a tempting way of *solving* the argument is the *substitution method*: once it is found a value of (NC), one searches for R in an informational context free of other unknowns.

2.7 The first version of the CEP argument

It is time now to make the CEP argument more vivid by particularizing some of its protagonists by making use of a possible reading. As I have already announced, I am going to work under the assumption that the causal relata are events and states (henceforth, I call them all events for a generic use). Thus, the mind-body causal non-identity assertion reveals as follows:

(NI) The mental events involved in any causal process are distinct from any physical events featuring into the same process.

According to the second CEP assertion, there are mental events endowed with causal powers (or causal efficacy) on the physical realm. That is to say that mental causation is not an illusion, and, at least, some of our mental episodes make a

difference for our behavior.⁵⁰ At this point, one aspect is in need of clarification: Is it claimed by the enthusiasts of mental causation that the mental is able *by itself* to causally affect the physical realm? Or just that, along with certain physical causes, some mental phenomena act as *partial causes* in bringing about some physical occurrences? It seems pretty clear to me that the interactionist doctrine does not require for the mental to be causally efficacious in the sense of being causally sufficient for some physical effects. Anyway, we do not have to clear waters in this matter, since, if there is any compatibilist solution, it will be good for both kind of mental causes, either sufficient or partial. Thus, for the time being, I am going to work with the following CEP assertion:

(MC) There are some mental events causally relevant for some other events of the physical world.⁵¹

If each of the first two of the CEP assertions derives from the interactionist doctrine, the third one, namely (PC), which claims the physical closure, is specific to physicalism. To my mind, (PC) may receive a variety of versions. Each of them provides a particular meaning for *being a closed domain*. In order to set up the first

⁵⁰ Note that, in our days, we are *psychologically* (or *culturally* as some might prefer) prepared to accept much easier a scenario where the mentality is epiphenomenal than one where the physical world does not affect causally our minds. However, it seems to me hard to understand how one would eventually give up of any form of causal "communication" between mental and physical.

⁵¹ Another expression of (MC) is: some mental events are causes of physical events. As I have said, some causes are sufficient for their effects, other causes are just partial causes of their effect. However, note that c_1 is a partial cause of e if and only if there is another cause, c_2 , of e such that c_1 and c_2 make up a condition causally sufficient of e. Thus, if (MC) is used in its weak form – i.e., there are some mental events playing the role of partial causes of physical events –, then (MC) entails the claim that there are some sets made both of mental causes and of physical causes such that the sets are causally sufficient conditions for some physical effects.

version of the CEP argument, I shall consider the most liberal reading, the one already expressed in § 2.3:⁵²

(**PC**_w) For any physical event, *x*, there were some physical events, y_1 , y_2 , ..., y_i , ..., no causal relations holding among them, such that their occurrences were causally sufficient for the occurrence of *x*.

If we call such a set as $\{y_1, y_2, ..., y_i, ...\}$ 'causally sufficient condition for *x*', it can be given the following formulation:

(PC_w) For any physical event, there is a physical causally sufficient condition.⁵³

Finally, many times, by 'overdetermined effect', it is meant an effect brought about by more than one causally sufficient condition. Inspired by the popularity of this usage, Bennett comes to make the claim that the exclusionists *in corpore* take the above definition to deliver the true meaning of overdetermination, that is, they use it as (SC) in delivering the CEP.⁵⁴ The generalization is for sure an exaggeration; the truth is

⁵² On the way of the analysis, the expression will gradually become more and more radical. Accordingly, the CEP argument as such will gradually become more and more aggressive (corrosive). ⁵³ Note that if a causally sufficient condition for *e* is made of only one physical event, then the respective event is a sufficient cause of *e*.

⁵⁴ Bennett puts it in the following way: "[T]he exclusionist at least has a definition of overdetermination up his sleeve—an effect is overdetermined just in case it has more than one sufficient cause—and, in accepting the various other components of the exclusion problem, the compatibilist has acknowledged that this is precisely what is going on whenever the mental causes anything. (p. 474)" We should note here that a couple of inadvertences (errors) crept into Bennett's phrasing. First, even though the author speaks initially as if an effect's having more than one sufficient cause is treated by the exclusionist as a *sufficient* condition of overdetermination – "no effect can have more than one sufficient cause *unless* it is overdetermined (p. 473, my italics)" –, by saying that "an effect is overdetermined *just in case* it has more than one sufficient cause" she ends up speaking about a *necessary* of overdetermination. Incoherence is also brought by term 'definition', which regularly would convey the idea of *necessary and sufficient* condition. Anyway, the inadvertence aside, it seems to me pretty clear that Bennett's intention is to show that the exclusionist defence of the CEP as a true problem is based on the idea that an effect's having more than one sufficient cause entitles us consider it as being overdetermined.

that not all exclusionists are so naïve as Bennett portraits in her article.⁵⁵ Moreover, I doubt that there is anybody who, without having any implicit qualification in mind or, rather, without supposing that there should be a certain additional qualification to be discovered, uses the existence of more than one sufficient cause of an effect as a sign enough for the overdetermination of the respective effect. However, given that what matters in an argument are not the qualifications one would like to find, but rather those one is able to write down, in stating the first version of the CEP argument, by using the claim that an effect that has more than one sufficient cause is overdetermined, we are not far from what many other people use.

Thus, we are ready now to deploy, in its first version (rather, a pilot production), the argument against interactionism particularized to the sort of cases exemplified in section § 1.2:

- (i) If (PC_W) is true, then plausibly, by construction p is a causally sufficient condition for *e*.
- (ii) If (NI)&(MC) is true, then by construction m is an event distinct from p such that m makes a causal contribution to e; for the sake of simplicity, let us consider that m is causally sufficient for e.⁵⁶
- (iii) Therefore, the truth of $(NI)\&(MC)\&(PC_W)$ entails that *e* has more than one causally sufficient condition, namely *p* and *m*.
- (iv) It is enough that *e* has more than one causally sufficient condition to be said that *e* is overdetermined by these causal conditions.

⁵⁵ In fairness, in order to entail the overdetermination of an effect, some of them speak about the satisfaction of a more restrictive condition than the simple existence of two sufficient causes of the respective effect. For instance, Kim (1989) says that the physical effect is overdetermined in cases where its mental cause and its physical cause are each an "independent" sufficient cause.

⁵⁶ Later, in Chapter 7, it will be considered both cases, when m is just a partial cause of e and, respectively, when m is enough for causing e.

- (v) Therefore, the truth of (NI)&(MC)&(PC_W) entails that e is overdetermined by p and m.
- (vi) But, *m*, *p*, and *e* may be the protagonists of any case of mental causation.
- (vii) Therefore, the truth of conjunction (NI)&(MC)&(PC_W) entails that in any case of mental causation the physical effect is overdetermined. That is to say that the truth of the CEP assertions implies that there is a systematic overdetermination in the cases of mental causation.
- (viii) It is irrational/unreasonable to hold the existence of a systematic and widespread overdetermination.
- (ix) Therefore, we should deem conjunction $(NI)\&(MC)\&(PC_W)$ as being false.
- (x) There are reasons for holding the truth of (PC_W) and hence, relative to *e*, the causal sufficiency of *p*.
- (xi) Therefore, conjunction (NI)&(MC) is false, there not being room for any mental cause, *m*, to be both distinct from *p* and causally sufficient for *e*.
- (xii) But, according to the interactionist doctrine, conjunction (NI)&(MC) is true, the mental having a distinct and sufficient causal role in bringing about some physical changes.
- (xiii) Therefore, interactionism cannot be but false.

The argument appears to be valid, and if it is also sound it shows that there is no room for mental causation as long as it is meant to be a process originated outside of physical domain. The simple idea employed here is that since it is unreasonable to accept that the same job is systematically accomplished by more than one processes and there are reasons to believe that the job is successfully accomplished within the physical domain, any other intervention coming from the exterior (if any) of the
physical domain should not be admitted as being responsible for the job. In other words, interactionism cannot receive what it asks for the mental in relation with the physical, i.e. both distinctness and causal efficacy.

Moreover, if the argument is sound and one is committed to mental causation, one should acknowledge that mental causation boils down to physical causation. As I have mentioned, such a line of reasoning is taken by many physicalists to show that mental is ontologically reducible to physical. In other words, the above argument is also seen by some as providing a good reason for the general thesis that mental causes (if any) are constituted of physical stuff. Such a reduction would be the only way for the mental to cause the physical.

Could the above conclusion – that, unless we are ready to give up to mental causation, we cannot but accept psycho-physical identity (either type or token) – be blocked? I believe that there is a way by which we could try to prevent it.

Chapter 3: Causation and Laws of Nature

Since CEP or, at least, its working version I have presented in the previous chapter is highly dependent on the notion of causation, it is this notion that should come under scrutiny before any discussion on the soundness of CEP and of the arguments it generates.

One of the fundamental convictions that shape my analysis of CEP is that the issue of causation should be understood in the light of the metaphysical relations supported by the laws of nature. In the beginning, the chapter presents the main positions regarding the relationship between causality and nomicity. As is shown in § 3.2, some people, assuming that there is no conceptual connection between causes and laws, argue for the idea that causality has nothing to do with nomicity. Whereas, others insist that the ground for stating a causal-law relationship is empirical in nature. Next, § 3.3 is focused on the difficulties to be overcome by any approach that goes along with the hypothesis that any causal relation falls under certain laws. The discussion is mainly concerned with the difficulty of distinguishing between cases of genuine causation and various cases of non-causal correlation.

The view that I favor is that any case of causation involves a case of necessitation holding in virtue of certain laws of nature. I call it 'the nomic account of causation'.

Employing our commonsensical notions, § 3.4 argues for there being an a priori correlation between causes and laws. After having introduced the notion of primitive causal relation in § 3.5, I am using it in the last subsection to outline the main theses of the nomic account of causation.

3.1 Singularist vs. non-singularist theories of causes

To explicitly articulate, or make more vivid, the nomic account of causation I shall draw the landscape of the main positions regarding the relationship between causation and laws of nature. Thus, by way of explanation, let us take the first step and consider two views that are each committed to a different kind of causal relation. Following Bigelow and Pargetter (1990), one way of classifying theories of causation is into "global" and "local" accounts. Local or singularist theories try to capture the intuition that what connects two events is causal as far as it depends only on the relata and what goes between them, not on conditions that hold elsewhere in the world. Instead, global or non-singularist theories hold that whether a connection between two events is causal should depend on what else goes on in the world, e.g., laws or regularities that hold in the world (Donald Davidson, John Clendinnen), or similarity relations to other possible worlds (David Lewis).

These two views are competing conceptions of the truthmakers for singular causal judgments, as Menzies (1999) puts it. On the one hand, there is the singularist view, according to which:

The truthmaker for a singular causal claim is supposed to be a local relation holding in single instances - a relation that does not depend on the existence of widewidespread patterns of occurrences. ... The causal relation does not depend on any other events occurring in the neighborhood: the causal relation is intrinsic, in some sense, to the relata and the process connecting them. (p. 314)

On the other hand, in agreement with the Humean intuitions about causation, there is non-singularist view, claiming that:

Causal claims are made true, not by intrinsic, local ties between events, but widewidespread patterns of occurrences. (p. 316) Thus, Menzies defends the idea that the crux of the issue between singularist and non-singularist views of causation is whether causation is an intrinsic or an extrinsic relation between events.

Following Lewis (1986b), a causal relation is taken in Menzies' paper to be intrinsic in the sense of being determined by, or supervening on, the perfectly natural properties of the relata and the perfectly natural relations holding between them. The notions of perfectly natural property/relation debut in Lewis (1983). There are two different theories that Lewis considers as equally possible candidates in accounting for the natural properties/relations. More or less, the central idea behind both theories is that these properties and relations carve nature at its joints; sharing them does count for true similarity. It needs to be added further that Lewis claims that the naturalness of properties/relations is a matter of degree, the most basic ones being also the most natural, i.e. perfectly natural. In his view, they form the basis on which all other less natural properties/relations supervene.

Instead, a causal relation would be characterized as being extrinsic if and only if it was fixed by "features of reality extraneous to the causally related events, in particular by widewidespread regularities holding throughout the world." (p. 320)

3.2 Looking into the singularist camp

Now, let us come closer to the main issue of this chapter, namely the relationship between causation and laws of nature. Even if singularists agree about the thesis that a causal relation is an intrinsic relation – a relation of power, energy, or necessary connection – holding as a local matter of fact, they divide among themselves over the proper stance to take on the question whether or not the causal relations are necessarily law-governed. Usually, the dispute amounts to the debate over the existence of an *a priori* necessary connection between causes and laws in virtue of a *conceptual* connection. As Menzies points out in his paper, some singularists (Anscombe, Armstrong, and Tooley) claim that, once granted that the causal relations are intrinsic, there is no conceptual restraint preventing one from thinking of the causal relations as being free of nomological support. Menzies expresses the reasoning as follows:

If the causal relation is an intrinsic relation that can hold independently of whatever events occur in its neighborhood, it should also be thought of as a relation that can hold independently of whatever laws obtain. (p. 315)

To my mind, the alleged way from the assumption of intrinsicness of the causal relations to the *possibility* of their holding independent of the obtaining of any laws should be mediated by two steps. The first argumentative step should show that the singularist conception of causation allows one to conceive a causal world free of laws no matter to what theory of laws one is committed. This would render false the claim that there is a conceptual connection between causation and nomicity. By consequence, the claim of *a priori* necessity would be discarded. The second step would be an appeal to the highly debatable principle that conceivability entails possibility. Plausibly, the only reason one might have for taking the second step is for

rejecting the more fundamental claim that there is a necessary connection between there being causal relations and there being nomic relations.

As we shall see, both steps are unwarranted. Let us take them one by one. For the sake of the argument, let us consider that a defender of the essentialist account of laws gives us sufficient reasons to embrace its theses: firstly, what causal powers the particulars have are determinant or essential for their identity (the affiliations of the particulars to different natural kinds are made in accordance with the causal powers they possess) and hence the causal powers with which a particular is endowed are intrinsic to it; next, the causal behavior of any particular is the exercise of its causal powers; finally, there are "causal laws" describing the way of acting of causal powers.⁵⁷ A direct consequence of these theses is that any case of causation falls under certain laws, but these laws are nothing more than descriptions of the exercise of the intrinsic causal powers of the causal relata. Thus, given certain assumptions about the nature of the laws involved in the causal phenomena, it follows that the causal relations meet both the intuition of intrinsicness - they are simply manifestations of the intrinsic causal powers - and the intuition of being lawful - they fall under laws describing the respective manifestations.⁵⁸ Therefore, in order to be warranted, the implication from intrinsicness of the causal relations to the conceivability of their holding without the obtaining of any laws needs a supplementary assumption. Probably, something like this should be granted: the application of the laws of nature in a case of causation is made true by circumstances

⁵⁷ The view, that laws describe the ways of acting of causal powers of particulars, is originated in Harre and Madden (1975). It is retaken in other recent papers like Bigelow, Ellis and Lierse (1992), A. Chalmers (1999) and Ellis (1999).

⁵⁸ My argument is inspired by Howard Sankey's speculations about the possibility of bringing essentialism and singularism into a closer proximity. See Sankey (1999: xv).

extraneous to intrinsic properties/relations of causal relata. The reason would be the belief that laws are no more than regularities in the behaviour of things.

It is interesting to consider the position held by Armstrong and Heathcote (1991). They argue for the conceivability of a nomic free, but causal world. However, unlike other singularists, Armstrong and Heathcote do not rely upon the assumption that causal relations are intrinsic. Their reason is that our commonsensical conceptions on causation and on nomicity are not restrictive enough to prevent us from conceiving a world where causation is not regulated by any laws:

Consider a world that is a world of causes and effects – where, perhaps, every event has a cause and in turn is a cause – but where causes of the same sort do not give rise to the same sort of effects. It sounds quite wrong. But it can be shown purely by analysis of the concepts involved that such a world is impossible? We do not think it can. (p. 66)

My objection is that even if we admit, which I do not, that there are no sufficient conceptual and logical constraints against the conceivability of a causal world without causal laws,⁵⁹ this is not a genuine ground for the respective conceivability. The possibility of conceiving a certain situation is not given by what the involved concepts do not exclude, but rather by what they support. A conceived situation is rather a construction grounded on what our concepts support. Or, better, we should take a situation as being conceived only if it is grounded on what our concepts support. I do not see how our concepts of laws and of causes support a coherent scenario where the events act as causes, but do not obey to causal laws.

⁵⁹ In § 3.4, against Armstrong and Heathcote, I shall try to bring briefly some reasons sustaining an inference from certain conceptual aspects to the impossibility of such a world.

Interestingly, even though Armstrong and Heathcote accept the conceivability of a causal, but anomic world, they stop short of endorsing the second step of the argument under analysis, holding that:

Not everything that is conceivable is possible. For we can conceive of water as a nonatomic, undifferentiated stuff, much as a pre-Socratic Greek philosopher might have done, without conceding that it is possible for water to *be* that non-atomic, undifferentiated stuff. For it is not possible: water is H_2O , differentiated and atomic. (p. 67)

Actually, Armstrong and Heathcote undertake themselves a plea for the idea that the relationship between causes and laws is necessary, but, in their view, the conclusion derives not from conceptual analysis, but rather from empirical considerations. The idea that grounds the argument brought by Armstrong and Heathcote is that, in speaking about the relationship between causes and laws, a better guide than our available concepts (and what we can conceive using such concepts) is the success of the experimental investigations.

Just as investigation shows that water is H_2O and can be nothing else, so investigation shows that causal sequences are essentially nomic. (*Ibid*)

The argument brought by Armstrong and Heathcote in favor of the idea that "causal sequences are essentially nomic" is supported, on the one hand, by an anti-Nominalistic view of causes and by the "strong" theory of laws,⁶⁰ and, on the other hand, by empirical investigations:

- The causal relations are instances of relations between the properties of the causal relata;
- ii) Laws are contingent relations between universals;

⁶⁰ Such a view of laws is developed among others by Armstrong (1983). It is a strong theory of laws in the sense that it stands against the Regularity theories by taking regularities to be phenomena having underlying laws rather than phenomena constituting laws.

- iii) It is a "meta-inference" from the success of repeatable experiments that the causal relations are identical with the instantiation of laws;
- iv) Therefore, it is empirically true that any causal relation is type-identical with the application of a law or of some laws, and hence it is an *a posteriori* truth that the causal relations are necessarily law-governed in the sense that the causal relations are instantiations of nomic connections.

The experimental method by which the identity between a causal connection and a

nomic connection is picked out is characterized as follows:

Suppose that one has a causal connection between A and B ... and one wishes to establish that it falls under a law, what would one do? The answer seems immediate and obvious: experiment. To identify the nomic connection one must isolate the property of the causal event that necessitates the effect. Thus one will begin by repeating the event but with slightly altered conditions ... [W]hat underwrites this process is a search for salient properties that are nomically related to the effects. (p. 68)

This leads Armstrong and Heathcote to the remark that:

There is nothing conceptual or logical that would guarantee that the causal connection between A and B is at bottom a nomic connection, although in fact it is. That it is in fact a nomic connection is the ontological justification for the experimental method. (p. 69)

Even though I tend to agree with the above argumentation, I still want to make the comment that the downgrading addressed by the authors to the conceptual analysis is unfair. Any defender of a conceptual connection between causality and nomicity would claim that a compelling conceptual analysis is able to provide good reasons for requiring that any causal relation falls under some laws in the sense that there should be some laws such that their application is identical with, or determines, the causal relation. This is not the same as the promise that, by conceptual analysis, we shall identify the respective laws. In other terms, the necessity one may be tempted to support by conceptual analysis regards the governance of causation by laws; it

should be distinguished from the necessity one might empirically prove, that holds in virtue of finding that there is identity between some causal relation and the instances of *certain* laws. The former is a conceptual or logical necessitation; the latter is a metaphysical necessitation. The latter is stronger because, once being established, the former follows.

A further point that I want to make is that the above argument brought by Armstrong and Heathcote for reducing causality to nomic instances, if correct as I believe, should teach us the lesson that what theory of laws one is in favor of is determinant for what theory of causes one is committed to. For example, making use of his theory of laws which takes nomic connections to be relations instantiated by fundamental universals tokening in elementary particles and hence to be universals themselves, Armstrong (1999) deduces from the identity of causal relations with instantiations of such nomic universal the intrinsicalness of the causal relations.⁶¹

Armstrong is aware that, if one adopted a regularity theory of laws, the same argument would deliver the conclusion that the instantiation of the law is extrinsic in character, which, by the identity thesis, would further commit one to a non-singularist view of causation. To make a distinction between the two views of laws, let us consider the following truism: an instantiation of a law exhibits a certain correlation, namely the "antecedent" conditions of the law are attended by the "consequent"

⁶¹ I take the following quotation from Armstrong (1999) to be relevant: "Suppose that laws link properties; suppose that the ultimate properties, the properties that divide the world along its ultimate joints, are universal; suppose that the ultimate laws, at least, are linkages of universals. Such a linkage of universals will itself be a universal. As a result, each instantiation of a fundamental law will be the instantiation of a universal. But, famously, each instantiation of a universal is complete in itself, so the law will be present *completely* in each instantiation. So where singular causation is the instantiation of such a law it will be a completely instrinsic relation." (p. 184)

conditions. Contrary to Armstrong's theory of laws, in which all it takes for something to be an instantiation of a law is to involve the obtaining of the antecedent conditions (even if those conditions proves to be met just for once, i.e. even if their meeting is unrepeatable), a regularity theory is a Hume-inspired theory in the sense that it preaches that the instantiation of a law is nothing else than a sequence (or an exemplification) of a regularity extended over the space and time.

3.3 Visiting the non-singularist camp

It should be no surprise for a friend of the regularity theories of laws and of causes that there are empirical considerations supporting the idea that there is an identity holding between the causal relations and instances of some laws. For one thing, once you take laws to be regularities, and causal relations to be instances of some regularities, it is an analytical truth for you that any causal relation is an instance of one or more laws. For instance, Davidson (1976), assuming both regularity theories, holds that any singular causal statement entails that there is a law – which should not be taken for the claim that the respective law is also known.

As far as I see, there are at least two objections that challenge such an approach. One objection is a technicality, another is more of a principle. I shall present and discuss first how Clendinnen tries to answer the former, and next I shall make some considerations by which one might hope to overcome the latter objection.

If one embraces a regularity theory of causes, an objection one has to face is the difficulty of distinguishing between some cases of genuine causation and the various

CEU eTD Collection

cases of non-causal correlation, occasioned by, and co-instantiated with, the respective cases of genuine causation. David Lewis takes this issue in his well-known article 'Causation' (1986 [1973]). He starts by considering the following general form of the contemporary regularity theories of causes:

In present-day regularity analyses, a cause is defined (roughly) as any member of any minimal set of actual conditions that are jointly sufficient, given the laws, for the existence of the effect. More precisely, let *C* be the proposition that *c* exists (or occurs) and let *E* be the proposition that *e* exists. Then *c* causes *e*, according to a typical regularity analysis, iff (1) *C* and *E* are true; and (2) for some nonempty set *L* of true law-propositions and some set *T* of true propositions of particular fact, *L* and *T* jointly imply $C \supset E$, although *L* and *T* jointly do not imply *E* and *T* alone does not imply $C \supset E$. (pp. 159-160)

Then, the main reason for which Lewis finds such regularity theories of causes as being unsatisfactory is their failure in distinguishing between the genuine causal relations and "other causal relations".

Thus, according to Lewis, there are at least three questions remaining unanswerable by any regularity approach. Firstly, *"causing – being caused" discrimination* or *asymmetry in causation*: how does such an approach discriminate a cause from any of its effects? Next, *"causation – preempted causation" discrimination*: given *e*, an effect of *c*, and *p*, one of its preempted causes, in what are *c* and *p* different? (In what are a cause of an effect and a preempted cause of the same effect different?) Finally, *"causation – non-causal effect-effect correlation" discrimination*: given *c*, *e*₁, and *e*₂, a cause and, respectively, two of its effects, how is a distinction to be made between the *c*-*e*₂ causal relation and the *e*₁-*e*₂ non-causal correlation?⁶²

⁶² Note that it is assumed a case where there is no causal connection between the effects e_1 and e_2 . However, it does not mean that there could not be cases where, besides c, e_1 is one of the causes of e_2 . Such cases will make the object of a detailed analysis in the last chapter of the present dissertation.

Here is how Lewis argues for the claim that the regularity theories cannot account for the asymmetry of the causal relations, i.e. for the difference between *causing* and *being caused* which holds between causal relata:

If *c* belongs to a minimal set of conditions jointly sufficient for *e*, given the laws, then *c* may well be a genuine cause of *e*. But *c* might rather be an effect of *e*: one which could not, given the laws and some of the actual circumstances, have occurred otherwise than by being caused by *e*. (*Ibidem*)

Some comments are in order here. I shall begin by noting that any law instantiation leads to the existence of some relation of *nomic necessity*. Further, I hold that Lewis's argument depends on what nomic necessitation is taken to mean by the assumed theory of law. In particular, my claim is that the argument is not conclusive unless nomic necessitation is taken in the *weak* sense, i.e. unless nomic relations are seen as being symmetrical.⁶³ This is because, in a nutshell, the argument points to an unwanted import of nomic symmetry to any causal relation identified with some nomic relation, and concludes against such an identification.

Let us consider that x and y are two nomical relata. It follows that there is a relation of *nomic necessity* holding between x and y, in short, xNNy. In the weak sense, 'xNNy' means that it is nomically necessary that x and y occur together. Here, nomic necessitation is not taken to involve a (logical) "sense of necessitation": both $X \supset Y$

⁶³ No surprise since Lewis endorses himself a systemic regularity theory of laws that involves such symmetry. His best-system analysis of laws is articulated in Lewis (1973: 73-5), where it is advocated that what makes a true regularity-statement – i.e. a statement whose truthmaker is a global pattern of regularity – to be a law is that it fits into some "integrated system of truths that combines simplicity with strength in the best way possible". The point that laws are the most systematic set of the regularities that the world contains is retaken in Lewis (1986 [1980]): "This is a kind of regularity theory of lawhood; but it is a collective and selective regularity theory. Collective, since regularities earn their lawhood not by themselves, but by the joint efforts of a system in which they figure either as axioms or as theorems. Selective, because not just any regularity qualifies as a law. If it would complicate the otherwise best system to include it as an axiom, or to include premises that would imply it, and if it would not add sufficient strength to pay its way, then it is left as a merely accidental regularity" (p. 122).

and $Y \supset X$ are true at our world and at any other world where the same laws that underlie the nomical relation hold. (*X* and *Y* are taken to stay for '*x* occurs/exists' and '*y* occurs/exists', respectively.)

In the *strong* sense, '*xNNy*' means that it is one of the relata (say, *x*) that nomically necessitates the other relatum (say, *y*). Here, nomic necessitation is taken to involve a (logical) sense of necessitation, say, from *x* to *y*: $X \supset Y$ is true at our world and at any other world where the same laws that underlie the nomical relation hold. Note that lack of sense of necessitation amounts to a kind of symmetry. Instead, in speaking about nomic necessitation, if the one employed is its *strong* sense no such symmetry is guaranteed, that is, nothing about the truth value of $Y \supset X$ follows from the existence of *xNINy*.

I still think of the Lewisian argument as having a flaw: *there is no direct argument for the claim that all nomical relations are symmetrical.* As I have said, such symmetry would be implied by the weak sense of nomic necessitation. Rather, I am tempted to adopt its strong sense because, by containing the idea of "agent-patient" relation, it sticks to the original meaning of 'relation of necessity', where relata are divided into those doing the necessitation (the agents) and the necessitated ones (the patients). Moreover, if there are causal laws – i.e. laws grounding causal relations –, the necessitation they impose is expected to involve a logical asymmetry.

The reference to causal laws is not intended to play any role in any argument against nomic symmetry, but rather to point out a difference between my starting point and Lewis's. Lewis, in accord with his theory of laws, is committed to nomic symmetry

and hence, on pain of symmetric causal relations, he is bound to reject any nomic account of causality. But, if the account of causality that one favors – as I do – relays on laws of nature, on the contrary, one is committed to causal laws and hence one is bound to reject any theory of laws that takes nomic necessitation in the weak sense, implying nomic symmetry. Thus, Lewis's task is to work out an account of causality without making use of nomic considerations, whereas the task for a friend of the nomic account of causality is to answer the last two Lewisian discrimination challenges.

As I said in the beginning, contrary to Lewis, I do favor a nomic approach to causation in the sense that I think of any causal relation as being a relation that holds in virtue of some nomic necessities.⁶⁴ However, beyond my reserve about the counterfactual interpretation Lewis gives of causal dependencies, I think of the technique of dependencies he makes use of in dealing with the three questions mentioned above as being valuable. He ingeniously reduces causal relations to chains or series of *causal dependencies*; given an event *e*, any of its causes, *c*, should be found backtracking along one of the chains of causal dependencies that end up in *e*.

Let c, d, e, ... be a finite sequence of actual particular events such that d depends causally on c, e on d, and so on throughout. Then this sequence is a *causal chain*.

⁶⁴ Usually, relations of necessity are supposed to hold between propositions. However, sometimes we also say that such-and-such property necessitates such-and-such property in the sense that any instantiation of the former *requires* an instantiation of the latter. Thus, in speaking about necessities, I shall move freely between the reign of propositions and that of property instantiations.

Any relation of nomic necessity is said to hold in virtue of some laws of nature that become active by instantiation of some proper properties. Furthermore, *nomic necessitation* is often distinguished from *metaphysical necessitation*. Granting that both are modal notions, the distinction could be easily understood if we make use of the apparatus of possible words. A particular relation of metaphysical necessity holds in every possible world, where the proper properties, which occasion the respective necessitation, are instantiated. Instead, the identity of any relation of nomic necessity depends on a certain set of underlying laws and hence it holds in all possible worlds that share the respective laws.

Finally, one event is a *cause* of another iff there exists a causal chain leading from the first to the second. (p. 167)

Lewis defines causal dependence between two events, *c* and *e*, as a counterfactual dependence between the same events. Considering two counterfactuals, $O(c) \square \rightarrow O(e)$ and $\sim O(c) \square \rightarrow \sim O(e)$ (where O(x) is the proposition that *x* actually occurs), if they are each non-vacuously true, the family O(e), $\sim O(e)$ depends counterfactually on the family O(c), $\sim O(c)$.⁶⁵ For simplicity, I shall say that the two counterfactual are each non-vacuously true iff *e* depends counterfactually on *c*.

On the base of such a counterfactual reading (interpretation) of causal dependency,

Lewis calls our attention on two characteristics of causal dependencies. On the one

hand, any causal dependence is an asymmetric relation.⁶⁶ Given that any causal

⁶⁵ Lewis explains further: "There are two cases. If *c* and *e* do not actually occur, then the second counterfactual is automatically true because its antecedent and consequent are true: so *e* depends causally on *c* iff the first counterfactual holds. That is, iff *e* would have occurred if *c* had occurred. But if *c* and *e* are actual events, then it is the first counterfactual that is automatically true. Then *e* depends causally on *c* iff, if *c* had not been, *e* never had existed." (p. 167)

 $^{^{66}}$ Lewis derives the asymmetry of causal dependences from the "irreversibility" of counterfactual dependences: "Just such irreversibility is commonplace. The barometer reading depends counterfactually on the pressure – that is as clear-cut as counterfactuals ever get – but does the pressure depend counterfactually on the reading? If the reading had been higher, would the pressure have been higher? Or would the barometer have been malfunctioning? The second sounds better: a higher reading would have been an incorrect reading." (pp. 168-9)

The irreversibility of counterfactual dependences expressing causal dependences is assured by Lewis by his imposing that the counterfactual involved must not be "backtracking" conditionals. (By the assertion of such a conditional it is claimed that, for an event not to have occurred, the past would have had to be different.) "Suppose that *c* causes a subsequent event *e*, and that *e* does not also cause *c*... If *e* had been absent, it is not that *c* would have been absent Rather, *c* would have occurred just as it did but would have failed to cause *e*. It is less of a departure from actuality to get rid of *e* by holding *c* fixed and giving up some or other of the laws and circumstances in virtue of which *c* could not have failed to cause *e*, rather than to hold those laws and circumstances fixed and get rid of *e* by going back and abolishing its cause *c*. (p. 170)

Note that, by using counterfactuals that are not backtracking, there is the advantage of answering directly the first question raised by Lewis: considering two events standing in the relation of causal dependence, the cause is that event of whose occurrence is asserted or denied in the antecedents of the two counterfactuals meeting the requirement of non-backtracking. Moreover, we are able now to distinguish cases of causation from correlations due to a common cause. (According to the above order, this is the third question.) It can be established a chain of counterfactual dependences from an effect to another only if it is made use at a certain point of backtracking counterfactuals. For simplicity, let us assume that c and e_1 stand in the relation of causal dependence, and the same is the case about c and e_2 . Let us also consider that e_2 occurs latter than e_1 . According to Lewis's theory of

relation is a chain of causal dependences, it follows a result consistent with our basic intuitions, namely, that any causal relation is an asymmetry. On the other hand, a causal dependence holding between two events implies that the events stand in a causal relation, but the converse is false because, unlike causation, causal dependency is not transitive. Consider that e_2 is causally dependent on c_2 , and e_1 (= c_2) is causally dependent on c_1 . (See Figure 3.1.) It follows that c_1 is a cause of e_1 , and c_2 (= e_1) is a cause of e_2 . By the transitivity of causation, c_1 is a cause of e_2 . However, it may be the case that e_2 is not causally dependent on c_1 .



Unfortunately, there are some difficulties with Lewis's taking causal dependences to be counterfactual dependences. Beyond the plausible objection that it is improper to

causation, e_1 is a cause of e_2 iff it can be established a chain of counterfactual dependences from e_2 to e_1 . (See Figure 3.2.) Backtracking from e_2 , by definition, e_2 is counterfactually dependent on *c*. Further, e_1 is the following on the chain of counterfactual dependences iff the following counterfactual is non-vacuously true: "If e_1 had not occurred, then *c* would not have occurred." But, then, counterfactual "If *c* had not occurred, then e_1 would not have occurred" (i.e. the counterfactual assuring the counterfactually dependence of e_1 on *c*) would be backtracking, and this would be against to Lewis's definition of counterfactual dependencies.

ground an account of causality on a semantic of possible worlds,⁶⁷ Lewis's analysis of causation does not overcome some cases of causal preemption.⁶⁸

An example of an author whose approach to causality comes close in some regards to the theory that I favor is John Clendinnen. Thus, before spelling out my own position, it might be useful to state briefly his stance on causation. Clendinnen, like Lewis, assumes a regularity view of laws, but, unlike the other, he argues for a regularity theory of causes. Given that Clendinnen takes the step of acknowledging that necessarily any causal relation falls under laws,⁶⁹ his consent to a regularity theory of causes is rather a consequence. To Lewis's reticence that such a theory does not allow for a distinction between causal relations and non-causal correlations holding in virtue of some proximal causal relations, Clendinnen opposes a technique of election (selection) based on what I would call the principle 'the more fundamental, the better'. It is this technique that I present next since it really matters for the approach I take to causation.

⁶⁷ I have in mind a point made by Armstrong (1999: 175-6), and retaken by Clendinnen (1999: 199). They say that what appears to be fishy about the account Lewis gives for causation is that, according to his counterfactual approach, the truth conditions for our causal claims should be ultimately the truth conditions for certain counterfactual conditions. But given that it is our evidence about the actual world that guide us in assessing causal claims (relative to the same actual world), it follows that some counterfactual truth – as it is the case with those defining a causal dependence – are true only in virtue of features of this world. (Supposedly, those features would provide us with the necessary evidence in appraising causal claims.) Armstrong witnesses that Lewis made apparent his support for such a conclusion. However, the problem with it is that one can hardly see *precisely* the way features of our world settle the truth-value of certain counterfactuals, which is a matter of ranking different possible worlds in terms of overall resemblance. True, Lewis gives us some instructions regarding the assessing of such resemblance – keeping the laws as constant as possible, and minimizing the size of miracle (see, e.g., Lewis, 1986 [1979]: 47-8). However, this does not make too much for diminishing the lack of precision.

⁶⁸ In literature, there are plenty of exemplas of cases of preemption complicated or tricky enough for making Lewis's analysis inefficacious. See, for example, L.A. Paul's works.

⁶⁹ It should be noted that Clendinnen, contrary to Armstrong, establish the necessary correlation between causation and nomicity on conceptual grounds, since, in his view, our concept of causation involves nomic connections between events.

Following Lewis, Clendinnen treats causal relations as chains of causal dependences. However, his persuasion about the existence of a conceptual connection between causality and nomicity, conjoined with Lewis's difficulty in interpreting causal dependences as counterfactual dependences, lead Clendinnen to a sort of nomical account of causality by reading causal dependences as *nomic* dependences.

All we need to know here about the Lewis-Clendinnen definition of nomic dependency is that they both agree that, given two events, *c* and *e*, the family O(e), $\sim O(e)$ depends nomically on the family O(c), $\sim O(c)$ iff the material conditionals $O(c) \rightarrow O(e)$ and $\sim O(c) \rightarrow \sim O(e)$ are true in virtue of some laws⁷⁰ and conditions that obtain. For simplicity, I shall say that, if the two material conditionals are each true, *e* depends nomically on *c*.

Note that the material conditionals $O(c) \rightarrow O(e)$ and $\sim O(c) \rightarrow \sim O(e)$ are true iff the material equivalence $O(c) \leftrightarrow O(e)$ is true. It appears then that if a causal relation is a chain of causal dependences, and, as Clendinnen holds, any causal dependence implies a nomic dependence, the way nomic dependence is defined leads to the idea that for any cause *c* and any of its effects, *e*, if *e* is nomically dependent on *c*, the occurrence of *c* is both a necessary and sufficient condition for the occurrence of *e*. But, if a cause is both a necessary and sufficient condition for its effect, then the converse is true too: the effect is both a necessary and sufficient condition for its effect.

⁷⁰ Actually, Clendinnen speaks about "projected generalizations" instead of laws. However, this difference is not relevant for the present discussion.

cause. Thus, it seems that such a conception allows for an objectionable symmetry in causal relation.

However, Clendinnen does not take account of the above worry. It might be that, in his view, the *logical* symmetry holding between causal relata does not affect the *metaphysical* asymmetry we usually assume to stand between a cause and its effect. Even though not made explicit, Clendinnen assures the causal asymmetry by identifying the sense of causation with the positive temporal direction. Yet, such a solution is legitimately criticized by Lews:

One might be tempted to solve the problem of effects by brute force: insert into the analysis a stipulation that a cause must always precede its effect (and perhaps a parallel stipulation for causal dependence). I reject this solution. ... It rejects a priori certain legitimate physical hypotheses that posit backward or simultaneous causation. ... It trivializes any theory that seeks to define the forward direction of time as the predominant direction of causation. (1986 [1973]: 170)

Instead, Clendinnen takes a serious interest in another challenge raised by Lewis to any regularity theory of causation: how does it distinguish between the genuine causal relations and the non-causal correlations holding between the effects having a common cause? Let us consider a sequence of events in the following temporal order: c, e_1 , and e_2 , such that e_1 and e_2 are each causally dependent on c, but e_1 does not cause e_2 . (See Figure 3.2.) In this case, if Clendinnen's view on causal dependency is correct, namely, that causal dependences are nomic dependences, there are two nomic dependences $-e_1$ on c, and e_2 on c. Moreover, given the definition of nomic dependency, there is also a nomic dependence holding between e_2 on e_1 . The difference between the first two nomic dependences and the third is that only the formers are causal, because only the formers holds between events

causally related. What is missing in the third one which makes it non-causal? How might we establish that e_1 is not a cause of e_2 ?



The solution that Clendinnen gives is helpful to, and in the spirit of, my idea that a cause is linked to any of its effects by a chain of causal dependences, each causal dependence implying the existence of a primitive causal relation.⁷¹ Clendinnen argues that, if a causal dependence is nothing else than a nomic dependence, in a case like that under consideration, we should discriminate causal nomic dependences from non-causal ones by imposing the requirement that the former hold in virtue of certain laws more basic than those underlying the latter. Clendinnen's proposal is that we should deem a nomic dependence as being non-causal if its underlying laws are derivable from the laws underlying other nomic dependences. In the given case, making use of the above notations, the laws underlying the nomic dependence of e_2 on e_1 are derivable from the laws underlying the nomic dependences of e_1 and of e_2 on c. In Clendinnen's own terms, in such a case, the

⁷¹ See Clendinnen (1992: 346-8, 352-3; 1999: 196-7).

nomic dependence of e_2 on e_1 is parasitic *inter alia* on other nomic dependence of e_2 , namely that of e_2 on c, the latter having a better claim to be *causal* than the former. In conclusion, a nomic dependence that links events in the forward temporal direction is a causal dependence provided there is no other nomic dependence ending in the same effect-event, whose underlying laws are constituents of the laws empowering the former dependence.

3.4 No laws, no causation

As I have mentioned before, once the regularity theories of laws and of causes are assumed, it seems to be a conceptual truth that any causal relation requires the application of some laws. However, not everybody is happy with the regularity theories. Presumably, a genuine argument for the conceptual connection between causes and laws, if there is any such connection, should not make use of concepts heavily loaded with assumptions requiring additional argumentation and being susceptible of falsification. It is our commonsensical notions of cause and of law that should be employed in one's analysis, if one expects a large addressability.

I hold that there is an a priori correlation between causes and laws. The reason is that there are at least two features inbuilt in our concept of cause, which would be affected in a world devoid of laws. In what follows, I shall make a brief review of these two ingredients featuring in our primary notion of causation.

First, the common sense view is that it is a sort of *necessitation* that is involved in bringing about an effect in the sense that, given the occurrence of its causes, it is

CEU eTD Collection

compelling for the effect to occur. Secondly, it seems to me that the common sense notion of causes excludes absence of regularity: by acknowledging a case of causation, we expect that, with other occasions, like causes produce like effects. Is it really conceivably a situation where *A* causes *B* and another time something similar to *A* causes something unsimilar to *B*? Is it really conceivable a situation where the heating of water makes it to boil and, another time, the heating of water makes it to freeze? Presumably, the commonsensical answer is negative. Hence, it seems that, our commonsensical conception on causation takes a causal relation to be an instance of *recurrent necessity*. But, in the empirical domain, even at the common sense level, such a relation of necessity is the very necessitation involved by the notion of law.

Given the above considerations, we are naturally led to say that *the recurrent constraint similar causes exert on similar effects to occur is nothing over and above the sort of necessitation to which we are committed when we acknowledge the laws of nature.* That is, there is a pretty irresistible temptation of admitting the following twofold explanation:

- (a) the occurrence of an effect is compelled by the occurrences of its causes in virtue of some laws triggered among others by the existence of the causes;
- (b) the recurrence in producing like effects from like causes is the result of the reactivation of the same underlying laws, which come in action on each occasion the same antecedents come into existence.

Even though the above reflections do not amount to a fully-worked out argument, it seems to me that they make up a reason enough for claiming that it is not

CEU eTD Collection

conceptually open that a causal relation can hold in the absence of any covering law. By contrary, it does not make any sense to speak about a world where there are causal processes, but no laws or generalities governing the causal correlated events. It is not a matter of lack of imagination, but rather a conceptual blockage in admitting a world where a brick's hitting a window is the cause of breaking it, but its hitting next time a window of the same kind as the first leaves it unbroken. I agree with Armstrong that a causal process is nothing over and above the activation or, in his terms, instantiation of some laws, but this truth does not need to be established by empirical investigation. In my view, it derives from conceptual analysis.

True, the case is still not clear-cut. I have held that causation cannot take place in a world devoid of any laws. The reason, as I have said, is that, in such a world, the absence of empirical necessitation (i.e. the necessitation in virtue of the triggering of some laws), which is a sort of necessitation required by any notion of causation, excludes the occurrence of any causal relations. But, the question raises, how come that the concept of empirical necessitation is inbuilt in our primary concept of causation?

Hopefully not very speculative, my answer is strongly based on considerations about concept formation. There are some remarks that I would like to make here. My guess is that the causal phenomena we encounter in our daily experiences provide us with the evidential stuff needed for the formation of the concept of empirical necessitation in a very rudimentary, but essential form. Presumably, the cognitive process mainly responsible for acquiring the concept of empirical necessitation is that of categorizing the causal evidences in classes in accordance with the *rule* 'like causes, like

effects⁷² (probably, without conceiving any of them as being either cause or effect). Further, once the causal relations are grouped in classes of similarity, it is a process of induction that comes into stage two: for any event, x, we met in our experience, there is the expectation that it is correlated with an event, y, such that y is similar to those events correlated to events similar to x. The expectation may be called the expectation of non-exception. In a way, it marks an epistemological step from particular experience to universal.

Now, my speculation is that, since the concept of empirical necessitation is acquired in virtue of our causal experiences and the concept of causation regards the very items – i.e. causal relations –sorted by cognizers to form the notion of necessity operating in their evidential environment, it makes sense to call a relation causal only if it is governed by a certain relation of empirical necessity, i.e. by some laws of nature.

3.5 Leading conditions for an account of causation

There are two theses to which I am committed and which lead me to the account of causation that I favor.⁷³ The first thesis concerns the relation between causation and nomicity (lawhood). The features of necessity and of generality constituting our concept of causation are to be naturally explicated as deriving from the necessities and generalities involved by the application of some laws of nature. Thus, an effect-

⁷² One may object that the causal evidence are not sorted out in classes by noticing degrees of similarity, but rather degrees of differences. According to the latter view on categorization, the cognizers would group two items as member of a class because they are less different from one another when contrasted against a third item present in the context of the assessment. But in response to this it may be urged that, since, on both views, the result of categorization is the obtaining of similarity classes, it is not crucial for the present discussion to take a stance in the issue of categorization.

⁷³ See the next section.

event is nothing else than an event whose occurrence is necessitated by some cause-events in virtue of certain laws triggered by the same cause-events.

Secondly, from my point of view, the crucial idea to which a theoretician of causation should be committed is that a cause is linked to any of its effects by a chain of primitive causal relations, where 'primitive' is synonym to 'without proper parts/constituents'. The second thesis is based on the reason that, given an event, *e*, one can get to a cause of it, *c*, only if one "breaks" the causal relation holding between the two events (or the way from *c* to *e*, that usually is mediated by some intermediary cause-events) in primitive or unmediated causal relations. In conclusion, any causal relation is either a primitive causal relation or a chain of such primitive causal relations.

Thus, as a consequence of the second leading thesis, if one agrees to follow Lewis's good intuition of defining a causal relation as a chain of causal dependencies, one should conjoin it with the proviso that any causal dependence corresponds to some primitive causal relation. My bet is that what might solve a good deal of issues in causation, including the three Lewisian discrimination challenges, is to find a proper understanding of causal dependency to the effect that it marks the most intimate causal relations. That is to say that *the causal dependence holding between any two given events is correlated to some primitive causal relation holding between the same events*.

But, then, what should causal dependency look like? Even though Lewis, who fathered the notion of causal dependency, seems to use it as a reliable guide in

CEU eTD Collection

backtracking from an effect to its genuine causes, he does not take causal dependency as being a mark for primitive causation.⁷⁴ Moreover, if one takes into account the difficulties encountered by Lewisian analysis in dealing without equivocation with some cases of preemptive causation, causal dependency, if understood in a counterfactual key, is not such a reliable guide for genuine causation.

As I shall point out in the next section, there is a notion of causal dependency coming close to that of Clendinnen's that stands as a serious candidate for the job of signaling primitive causal relations. But, before deploying the specification of the respective candidate, we need to have a look at what would make a causal relation primitive.

Let us begin by noting that a causal relation, as it should be understood here, is a relation holding between only two events, a cause-event and an effect-event. The *identity* of any causal relation is given by the events among which it makes a connection and the route followed by the respective connection.

Usually, we assume that any causal relation is characterized by some asymmetry consisting in a kind of logical (if not even metaphysical) precedence the cause-event has relative to the effect-event; the occurrence of the cause-event contributes to the fulfillment of a sufficient condition for the occurrence of the effect-event. Thus, it makes sense to say that causal relations are *orientated relations*, the sense of any causal relation being from the cause to the effect. When one talks about a causal relation, one takes the cause to be point of reference, looking from that "point" to the

⁷⁴ Moreover, Lewis does not even encourage the notion of primitive causation.

effect. Instead, if there is also a causal dependence holding between the same events, we are dealing again with an oriented relation, its sense running this time from the effect to the cause. The difference between a causal dependence and the correspondent causal relation holding between two events is similar to the difference of accent between two sentences expressing the same thing, but under a different voice. A direct example would be the passive voice sentence 'e is caused among others events by c' and the active voice sentence 'c along with other events cause e'.

Let us consider the case illustrated in Figure 3.3. The c_1 - e_2 causal relation is correlated to a chain of causal dependences going from e_2 to c_1 . But, any causal relation that is correlated to a causal dependence featuring in the respective chain features in the constitution of the c_1 - e_2 causal relation. Thus, for example, the c_2 - e_2 causal relation is a constituent of the c_1 - e_2 causal relation. The direct conclusion is that a causal relation correlated to a chain of causal dependences is composed, and hence not primitive.





Then, seemingly, it would remain only for those causal relations corresponding each to a single causal dependence to be primitive. As I have mentioned above, this claim is strengthen by the considerations that, on the one hand, the notion of causal dependency has been introduced to ensure a mean for backtracking from an effect to its real causes, and, on the other hand, the safest way for such a backtracking is to "jump" each time from an event to one of its most intimate causes and so on.

As I see it, given two events standing in a causal relation, the relation is primitive if and only if it is a direct or intimate connection holding between the two events in the sense that there is no need of another event to mediate the causation. Therefore, granting that any primitive causal relation is a correlate of some causal dependence, causal dependency should be construed such that, whenever it is the case, it reveals the feature of being unmediated that each primitive causal relation has.

It should be noted that the existence of a chain of causal dependences linking two events does not entail that there is no primitive causal relation holding between the same events. Even though the chain of causal dependence is a correlate to a composed causal relation, there is no reason for ruling out the possibility of two distinct causal relations holding like two parallel connections between the same two events.

Note that Figure 3.3 illustrates a case where, besides being causally dependent on c_2 , e_2 is also causally dependent on c_1 . Therefore, c_1 is a cause of e_2 in virtue of two

CEU eTD Collection

distinct causal relations holding between the two events, one of them mediated by $e_1(=c_2)$, an intermediary cause, the other connecting c_1 to e_2 in an unmediated way.

3.6 The nomic account of causation

Now, in the end of this chapter, I shall make a sketch of what I have called the nomic account of causation. In this view, the notion of causation should be glossed in the light of the concept of fundamental nomic necessitation.

In the spirit of Lewis's analysis of causation, the nomic account of causation urges that an effect is linked by any of its causes through some chain of causal dependences, but, contrary to Lewis, a causal dependence is not taken to be a counterfactual dependence, but rather a nomic dependence.

Following Clendinnen, I hold that just some of the nomic dependences, namely, those empowered by fundamental laws, should be thought of as being causal. Recall that, according to Clendinnen, e_k depends nomically on c_i iff the material conditionals $O(c_i) \rightarrow O(e_k)$ and $\sim O(c_i) \rightarrow \sim O(e_k)$ are each true in virtue of some laws and conditions that obtain. In my view, a nomic dependence coincides with a part of a relation of nomic necessity in the sense that, by saying that e_k depends nomically on c_i it is meant that there is a relation of nomic necessity holding between c_i plus some other necessitating events, and e_k plus some other necessitated events. See Figure 3.4.



Fig. 3.4 c_i and c_j are causes of $e_k \Rightarrow \{c_1, \ldots, c_i, c_j, \ldots\}$ nomically necessitates $\{e_1, \ldots, e_k, \ldots\}$

Any nomic necessitation is empowered by some set of laws whose application is occasioned by the occurrence of the necessitating events. It should be considered that $C = \{c_1, ..., c_h, c_j, ...\}$ nomically necessitates $E = \{e_1, ..., e_k, ...\}$ iff the events belonging to *C* necessitate – in other terms, are sufficient for – the occurrence of each events belonging to *E* in virtue of some set of laws, L.⁷⁵ Iff, on the one hand, in any given possible world where the laws of *L* hold, the occurrence of $c_1, ..., c_h, c_j, ...$ is a sufficient condition for the occurrence of any of $e_1, ..., e_k, ...$, and, on the other hand, in our world, but also in any of the worlds where the laws of *L* hold, the occurrence of the same sort of events as $c_1, ..., c_h, c_j, ...$ is a sufficient condition for the same sort of events as $e_1, ..., e_k$, In short, *C* nomically necessitates *E* iff in any given possible world where the laws of *L* hold, the laws of *L* hold, the

⁷⁵ It does not mean that there could not be a case where nomic necessity runs between individual events, say, c_i and e_k . It is only that I prefer to take the most general case – i.e. when more than one event necessitate more than one event –, because I try to avoid the threat of the widespread prejudice that such a necessity runs *only* between individual events and hence, if the nomic account of causation is granted, that any effect has only one cause.

occurrence of *C*'s events or of other events of the same sort is a sufficient condition for the occurrence of each of *E*'s events or of other events of the same sort.

The last equivalence is made under the assumption that: 1) two events are of the same sort when each of them implies the instantiation of the same property,⁷⁶ and 2) the application of any law is a case of a recurrent association of property instances.

Before discussing in greater detail the nomic account of causation, it is important to note two aspects about it. In the first place, it is a reductive approach to causation that moves the burden onto a theory of laws. Accordingly, the way laws are understood – either as relations between universal, as regularities or as something else – reflects how we should take causal relations to be: either intrinsic or extrinsic. Another related point to notice here is that, fortunately, it is not necessary for my purposes to decide between the competing theories about laws; and I shall try to remain neutral between them. Therefore, I shall use the concept of law of nature more as a primitive notion to the effect that the other useful notions like nomic necessity, nomic dependency, causal dependency, causal relation and causal role are introduced in a derivative manner.

There are cases where, relative to an effect, e, an event, c, is only one of its *partial* causes. In such a case, beside c, the occurrence of e requires additional events to cause it. This is coherent with the nomic account of causation in which the occurrence of an effect may be due one or more cause-events triggering the

⁷⁶ This is in agreement with my conception that the occurrence of an event implies the exemplification of a property. See chapter 4.

application of some causal laws. As I have mentioned, the nomic approach to causation sketched here takes any causal relation as being part of a nomic necessitation. If this is right, given that a relation of nomic necessity may be a many-to-many or many-to-one relation, it follows that the same effect-event may be causally related with more than one cause-event. In other words, it might be the case that there are more than one causal relations holding between the same effect-event and some cause-events, and, importantly, none of the respective causal relations features in the constitution of another causal relation from the same group.

I should add that I am not very enthusiastic either about the claim that a cause is necessary for its effects. It seems to me plausible enough that there are cases where an effect could have had a different cause, if its actual cause had not been. By acknowledging that any case of causation is part of some nomical necessitation, that is, a case where the occurrence of some effect-events is necessitated by some cause-events in virtue of triggering the activation of certain laws, it is not also implied that in the absence of some of the actual cause-events, there would have not occurred other events ready to play the same role in triggering the nomic necessitation.

In what follows, before closing this chapter, I want to advance an argument designed to show that the nomic account of causation meets the constraints I have stressed in the previous section. According to the discussed account of causation, any causal dependence is a nomic dependence empowered by some fundamental laws. Then, given that, by definition, any nomic dependence coincides with a part of a relation of nomic necessity, it follows that any causal dependence coincides with a part of a

relation of nomic necessity empowered by some set of fundamental laws. From that, it follows further that causal dependency is a mark of unmediated connectivity holding in virtue of fundamental nomicity.

Moreover, given the thesis of the nomic account of causation that any causal relation coincides with a chain of causal dependencies, it follows that any causal relation is part of a relation of nomic necessity in the sense that its causal relata are among the relata of the necessitation. Hence the first condition (constraint) for any account of causation is met.

Next, given that causal dependency signals an unmediated connectivity, if any causal relation coincides with some chain of causal dependencies, a certain causal relation is primitive if and only if it coincides with a causal dependence. The further consequence is that any causal relation is either primitive or decomposable in a chain of primitive causal relations, and hence that the second condition is met.

Chapter 4: A Sketchy Metaphysical Theory of Events

Why should there be a chapter in the present work devoted to events? The main reason is that the analysis is built up on the belief that there is no causation without events. The identity of such entities is crucial for the identity of the causal network involved by CEP.

As § 4.1 will show, following Davidson, one may base the positing of events on the reason that they serve us to account for the entailments between some action sentences. However, § 4.2 points out that there is a problem when it comes to saying what events are: their nature cannot be derived or deduced from the role they play in accounting for certain entailments. The reason is that, as § 4.3 exemplifies, there are at least two different metaphysical views on events, each rendering a different account of the same action sentence entailment.

Given that a key notion responsible for the differences is that of metaphysical structure, § 4.7 discusses in some detail what it is to be understood by 'having a structure' when it is about an ontological entity. Exemplifications are given using Kim's view on events. The discussion proves to be crucial in approaching the issue of event identity, because it is highly dependent on what structure (if there is any) events have.

In § 4.8 and § 4.9, I shall present the theory of events that I favor in a sketchy manner. Starting from the common idea that an event is a change in properties, I

shall suggest a view that takes events to be temporal structures of states. Naturally, if I am right, the identity of an event depends also on the temporal profile of transition through the intermediary states, as it is shown in § 4.10.

Along the chapter, I shall address various objections to the competitor theories. § 4.4 emphasis a conflict among Kim's claims regarding events, and § 4.5 points out a major difficulty for any current theory of events. Finally, in § 4.10, I shall argue for the superiority of the view I defend in comparison to those discussed above.

4.1 Davidson's argument for the existence of events

Lowe suggests that "we don't really need an argument to justify our belief in the existence of events quite generally, any more than we need an argument to justify our belief in the existence of things quite generally, because we can't really make sense of the idea of there being no events or things whatever." (2002: 219) However, against this suggestion it may be urged that, by contrary, the existence of events has the same need of proof as the existence of any metaphysically interesting kind of entities.⁷⁷ The moral seems to be that, no matter whose existence is talked about (whether objects/things or events or other kind of entities), the considered level of requirement should be the same.

Many times, metaphysicians posit certain entities under the guidance given by the need of achieving a better understanding in some particular matter. Thus, a credible plea for including events in our ontology may start from the following question: do

⁷⁷ Lombard (1986: 9), for example, takes a similar stand.
they consolidate our understanding? What good are having events in ontology? Or, the other way round: what is it that would be lost by expelling events from ontology?

Perhaps the most famous answer is proposed by Donald Davidson. He points out that it seems highly unlikely to "give a cogent account of action, or explanation, of causality, or of the relation between the mental and the physical, unless we accept events as individuals" (1980b: 165).⁷⁸ Without events, he argues, it does not seem possible to give a proper account of the logical form of simple sentences about actions.⁷⁹ An action sentence typically consists of a subject term denoting an agent, a verb of action, and one or more adverbs or adverbial phrases expressing the time, manner or place at or in which the agent acted.

Let us consider the following sentence from Davidson (1980a: 106):

(1) Jones buttered the toast slowly, in the bathroom, with a knife, at midnight. From such adverbially qualified sentence it can be validly inferred a sentence obtained from the former by detaching one or more of the adverbial qualifiers. For example, (1) entails:

(2) Jones buttered the toast slowly, with a knife, at midnight.

Further, (2) entails:

(3) Jones buttered the toast slowly, with a knife.

⁷⁸ Davidson's detailed arguments for this claim are expended in some of his essays collected in *Essays Actions and Events* (Oxford: Clarendon Press, 1980).

⁷⁹ This argument was presented by Davidson for the first time in his 'The Logical Form of Action Sentences' (1967). What is actually offered is an argument for actions' existence. But Davidson takes actions to be events. It is of no direct concern for me here whether this belief is true. Rather, I am interested in the question of what and how much such an existential proof could reveal us about a certain kind of entities and what and how much it is left out.

Davidson's proposal for explaining this kind of entailment consists in taking each action sentence as involving a covert existential quantification over events. Thus, according to Davidson, the three sentences could be rewritten as:

(1D) $(\exists e)$ (*e* was a buttering of the toast by Jones, and *e* was done slowly, and *e* was done in the bathroom, and *e* was done with a knife, and *e* was done at midnight)

(2D) $(\exists e)$ (*e* was a buttering of the toast by Jones, and *e* was done slowly, and *e* was done with a knife, and *e* was done at midnight)

(3D) $(\exists e)$ (*e* was a buttering of the toast by Jones, and *e* was done slowly, and *e* was done with a knife)

Thereby, the above entailments can be explained in virtue of the truth-functional properties of the conjunction 'and'.

According to Quine's criterion of ontological commitment,⁸⁰ having expressed action sentences in the language of first-order predicate logic, their truth requires the existence of events as possible values of the variables of quantification in the logical form of such sentences. For instance, the truth of (1D) requires the existence of some entity to which predicate 'is a buttering' applies. Davidson takes that entity to be an event.

⁸⁰ 'To be is to be the value of a variable'. See, for example, W. V. Quine, 'Existence and Quantification', in his *Ontological Relativity and Other Essays* (1969).

4.2 Competing candidates for taking a role

One deficiency of Davidson's argument appears to be that it shows only that there should be some entities, namely, events, as one may call them, leaving unanswered the question of what these entities are. Since Davidson's argument relies only upon his semantical theory, it is short of being a proof for a *qualified* existence. Without providing some qualification, we get only an argument for the existence of a certain role to be fulfilled, but the nature of the entities fulfilling the role remains unrevealed.

However, the objection I have in mind requires some expansion. To see better what it is all about, let us consider the following train of thought. Even if we admit that Davidson's semantical theory for sentences describing events wins the competition with other theories he takes into account, we should still note that Davidson's theory is winning only relative to the players present around the game table. Thus, one could reasonably claim that the status of the theory under consideration does not exclude the possibility of another theory capable of accounting for the entailments between event sentences⁸¹ without constraining us to admit events into our ontology. Moreover, it does not exclude the possibility of a very different theory which would also show that event sentences contain a covert existential quantification over events.

Let us consider two sentences, s and q. It seems fair to say that s entails q if and only if the information expressed by q is contained in the information expressed by s. So, as a general strategy, given that any entailment is nothing else but information

⁸¹ Henceforth, instead 'action sentences' I shall use the phrase 'event sentences', taking actions to be just a kind of events.

containment, accounting for an entailment between two sentences amounts to providing the *right* logical form of each action sentence such that the information containment becomes logically obvious and hence justified. Only a proper logical form that discloses the real jointure among the bits of information (i.e. the logical underlying structure of information) expressed in each sentences qualifies us to discern genuine information containments from illusory ones (on which more later).

Let us assume that *s* and *q* are two event sentences. Then, granting that each event sentence expresses some information about an event, it is plausible that, by making explicit the logical underlying structure of the information in each sentence, it would become clear whether or not *s* entails *q*. The logical underlying structure of the information expressed by *s* reveals one about all sentences that are entailed by *s*. The point is that, by knowing how the information is structured in *s*, one implicitly knows how one can operate on the structure, by cutting it out, to obtain another sentence such that *s* entails it.⁸²

Davidson's argument for the existence of events strongly relies on a certain particularization of the above general strategy of accounting for entailments between event sentences. However, his manner of revealing the information containment engages a certain view about the jointure among the *bits* of information in event sentences. As long as there is more than one such view, the things are far from being settled.

⁸² Note that the respective information is not always about the initial event as a whole. For instance, it could be information about a constituent (if any) of the event. In such a case, we could witness an entailment from an event sentence to a sentence of another kind (e.g., a sentence about the substance that undergoes the event).

Let us take a look at the considerations made by Oliver (1996) about a general way of arguing for the existence of the entities of a certain ontological category. He correctly remarks that "some ontological categories can be associated with offices which entities hold, or roles which they play. Very often, the role can be characterized as explaining a purported fact or solving a problem." (p. 11) Thus, says Oliver, by associating a certain role with a certain ontological category and by arguing that the role must be played by something, it is proved that there should be some entities which qualify as entities belonging to the given ontological category. In other words, he acknowledges the validity of the following ontological inference: 'For any given theoretical role⁸³ to be necessarily played, if the entities of a presumed ontological category is not empty.'

However, what worries Oliver is that one would reasonably want to know more than whether a certain ontological category is empty or not. It is not enough to say that there should be some entities in an ontological category, because there is a certain worth playing role to be fulfilled by those entities. Not much theoretical progress will be made unless more is said about the nature of entities in the category.

Oliver shows that different players could play the same theoretical role (i.e. explaining the same fact or solving the same problem). He calls a group of entities able to play the role 'system of candidates'. The argument is that, when there is more than one system of candidates that could fulfill a certain role, the necessity of that role is a

⁸³ In fact, usually, there could be more than one role which could be associated to one ontological category, but for the simplicity of the argument I formulate it as if there were only one role to be played by the entities of one category.

sufficient rationale for claiming the existence of a system (i.e. the system of entities belonging to the ontological category responsible for the role), but it is not a sufficient means for fixing on a unique system. What it is required here is to find a reason in virtue of which, among the possible systems, one is able to pick up the system that actually fulfills the role:

In the case of [equally serviceable candidate systems], if we want to claim that a certain system exists, then we must do more than argue that the given system plays a necessary role. We must proceed to dismiss the other candidate systems on grounds other than their ability to play the role since, *ex hypothesi*, all the candidate systems play the role equally well. (p. 12)

It is one thing – a right thing – to claim that *events exist because there is a worth playing role* as long as by 'events' it is meant those entities, whatever they may be, that actually play the role. It is as if you claim that since there is a *patient*, i.e. something on which it is acted, there must be something that is *its agent*, i.e. something who acts on it. Nothing more grounded than the analytical inferences – e.g., if there is a son, there is a father –, yet not much progress in knowledge is made.

It is quite another thing – a wrong thing – to claim, as Davidson does, that *events exist because there is a worth playing role* as long as by 'events' one referres to entities of a certain kind that could play the role. It is as if, besides the reasonable claim that since there is a *patient* there must be something that is *its agent*, it is added that since *A* could be the agent, *A* really is the existing agent. There is nothing more ungrounded than the inference from possibility to actuality.

Davidson correctly infers from the worth playing role of accounting for event sentence entailments that there must be entities of a certain kind that play the role. Moreover, there is no problem with naming those entities 'events'. But there is a problem when it comes to saying what events are. Their nature cannot be derived or deduced from the role they play in accounting for certain entailments.

To my mind, in order to come closer to the nature of events, the progress in knowledge must take the following two steps: 1) identifying a system, *S*, of candidates capable of playing the role initially associated to the category of events; and 2) proving that *S* is the same as the system that actually plays the role. How is (2) to be taken? One has to show either that *S* is the only system of candidates capable of playing the role – there is something *special* in the nature of the entities belonging to *S* that makes them the only eligible candidates for playing the role –, or that there are further reasons for positing *S* instead of other system of candidates. What these further reasons would be is the subject matter of a subsequent section.

4.3 With or without internal structure?

CEU eTD Collection

Let *S* be the system of candidates proposed by Davidson. We should note that there could be at least one more way of rendering an account of the entailments under discussion – hence, a competing system, *Q*, of candidates. Someone like Kim, and unlike Davidson, may argue that any event has a certain internal structure whose constituents belong to other ontological kind than events do.⁸⁴ If one took this line, it

⁸⁴ It is not that Davidson denies that an event may have more *ordinary* constituents – i.e. parts which are nothing else than other events –, thus having a structure in a trivial sense. Rather, the point is that Davidson would deny that any event has constituents which are not events. That is, he would deny that any event has – what I call – a *metaphysical structure*. Plausibly, Davidson takes any event, if composed, as being ultimately composed of simple, *monolithic* events – i.e. as being entities devoid of metaphysical structure.

It is true that Davidson does not exploit anywhere within his account of the event sentence entailments the fact that some entailments are true in virtue of some compositional structures (that is,

would not be surprising to find out that one distinguishes between information about constituents of a certain event and information about the event as a whole. I shall follow Kim (1993a [1976]: 42) in using 'intrinsic descriptions of event e' to denote those descriptions which can be used to express information about the constituents featuring into the presumed structure of e. Such descriptions should be distinguished from those – denoted by 'extrinsic descriptions of event e' – used to express information about event e as a whole.

Of course, the identification of these two kinds of descriptions depends on what one takes to be the general structure of events. Such a structure is delivered by the metaphysical theory of events to which one is committed. In other words, the way the information about an event is structured heavily reflects the way the nature of events is metaphysically understood. Thus, each metaphysical theory of events renders both a particular system of candidates for the true nature of events and a particular account of the event sentence entailments, according to a particular way of writing event sentences in the language of the first-order predicate logic.

what I call *categorial structures*). Nor he explicitly rejects anywhere the idea that events would be metaphysical constructions constituted of entities of other metaphysical kinds. However, it should be noted, there are some reasons for attributing him the monolithic view. For instance, the underlying logical structure he urges for any event sentence suggests that, from his perspective, the information is composed of bits of information such that each of them is about the event as a whole. (In the present usage, something is deemed to be a bit of information only if there is an intelligible sentence to express it.)

But the adopted conception on the structure of information about events (to choose arbitrarily the name of entities belonging to a certain kind) reveals a great deal about the adopted metaphysical conception on events, including on whether they have or not a metaphysical structure. The reason: how an entity is structured determines how the information about this entity is (logically) structured; or, maybe better said, how an entity is thought to be structured constrains how the information about that kind of entity is thought to be structured. Thus, since, in Davidson's case, the logical form he proposes for an event sentence does not speak about anything which could feature in the structure of the event, but only about the event as a whole, it is reasonable to claim that, for him, events are either monolithic or homogeneous entities. This conception should be distinguished from another according to which events are heterogeneous entities. An example is Kim's conception on events. The whole distinction is grounded on what entities are thought to be present in the structure of an event. In the next section I shall discuss at large what could be meant by 'having a structure.'

Given that there are (or there could be) more than one metaphysical theory of events or, at least, *various conceptions about the general structure of events*, the above remarks suggest that there are (or there could be) *more than one proposal of logical underlying structure of information* expressed by an event sentence. But, if that is right, the possibility emerges that the discussed strategy of accounting for the event sentence entailments could be pursued in various ways.

Even though Davidson's system of candidates, *S*, is capable of playing the role of accounting for the entailments, besides *S*, there are (or there could be) more other systems (e.g., *Q*) capable of playing the same role. Therefore, a further reason is needed for electing the winning system – i.e. the system whose entities actually play the required explanatory role. Thus, the competition moves from a semantical level to a metaphysical one, because the combat is now given among theories accounting for the nature of different candidates aspiring to the same semantic role.

For illustration, let us consider Kim's metaphysical theory about events. Unlike Davidson, who takes events to be monolithic entities or, at most, homogeneous entities, Kim advances a metaphysical conception which attributes to every event a heterogeneous or metaphysical structure. Given that, according to Kim, events are exemplifications by substances of properties at a time, it follows that "an event is a structure consisting of a substance (an *n*-tuple of substances), i.e. the *constitutive object*, a property (an *n*-adic relational attribute), i.e. the *constitutive property* or *generic event*, and a time." (1993a [1976]: 34)

110

Davidson's conception about the nature of events allows only extrinsic descriptions, whereas a conception like Kim's opens the possibility of intrinsic descriptions. The difference becomes more visible if one looks at what logical form an event sentence receives from each perspective. Perhaps, if events have a metaphysical structure as that proposed by Kim, one is entitled to assign to the sentence below:

(1) Jones buttered the toast slowly, in the bathroom, with a knife, at midnight. the following logical form:

(1K) ($\exists < x, y, t, R >$) (< x, y, t, R > is an exemplification of *R* by *x* and *y* at *t*, and *x* was Jones, *y* was the toast, and *R* = *a* butters *b* slowly, with a knife, and *t* was midnight, and < x, y, t, R > was done in the bathroom).

On this particular approach, the information about the structure of the event is expressed by the following intrinsic description: '<x, y, t, R> was an exemplification of R by x and y at t, and x was Jones, y was the toast, and R was a butters b slowly, with a knife, and t was midnight'. Whereas phrase '<x, y, t, R> was done in the bathroom' represents an extrinsic description of the event, i.e. a description that gives information about the event as a whole. Compare with (1D), Davidson's proposed logical form, which treats sentence (1) as conveying only information about the event as a whole.

Now, remaining under the same conception, the sentence below:

(2) Jones buttered the toast slowly, with a knife, at midnight. receives the following logical form:

111

(2K) ($\exists < x, y, t, R >$) (<*x*, *y*, *t*, *R*> is an exemplification of *R* by *x* and *y* at *t*, and *x* = Jones, and *y* = the toast, and *R* = *a* butters *b* slowly, with a knife, and *t* = midnight)

Thus, it can be observed that the information expressed by (2K) is contained in the expressed by (1K). What (1K) has complementary is information referring to the event as a whole, namely, that the event has the property of *occurring in the bathroom*. It is obvious, then, that (1K) entails (2K) and hence why (1) entails (2).

Finally, in the same vein, the appropriate logical form of the sentence below:

(3) Jones buttered the toast slowly, with a knife.

would be:

(3K) ($\exists < x, y, t, R >$) (< x, y, t, R > is an exemplification of R by x and y at t, and x

= Jones, and y = the toast, and R = a butters b slowly, with a knife)

Unlike the (1)-(2) case, where the entailment between the sentences is justified by a relationship of inclusion holding at the level of information regarding the whole event, (2) entails (3) in virtue of another kind of inclusion, namely, that holding at the level of information about some constituent featuring into the structure of the event.

Summing up, I have presented two different logical forms for the same set of sentences, namely (1)-(3), each logical form seeming to account for event sentence entailments according to what metaphysical conception on events is taken.

4.4 Some difficulties for Kim's view

It is true that Kim has not used anywhere the logical form ascribed above to his view on the nature of events. Moreover, it is not at all sure that Kim would agree that the logical forms of sentences (1)-(3) may be (1K)-(3K). The worry arises mainly when one thinks about Kim's pretence of compatibility between his metaphysical theory of events and Davidson's semantical theory of action sentences. Let us consider the following event sentence:

(i) Flora dried herself with a towel on the beach at noon.

According to Davidson's semantical theory, the underlying logical form is:

(ii) $(\exists e)$ (e was a drying of Flora by Flora, e was done with a towel, e occurred

on the beach, and e occurred at noon.

Kim holds that (ii) is compatible with his metaphysical view:

there seems to be no reason why the variable 'e' cannot take as its values the event structures of the property-exemplification account: in fact, no reason why the particular event structure [(Flora, Flora), a dries b, noon] isn't just the value of 'e' that makes (ii) true. (1993a [1976]: 39)

Let us call this the *compatibility claim* (cc).

In these circumstances, the question is whether (cc) is in conflict with other two claims made by Kim on events. It is said that predicate modifiers indicating meansmanners-methods – e.g. 'with a towel' – do not indicates properties of individual events – e.g. *Flora's drying herself* –, but do modify generic events (i.e. the constituent properties) – *e.g. drying*.⁸⁵ Call this the *modifying claim* (mc). Also, it is said that, since *drying* and *drying with a towel* are different properties, any two events

⁸⁵ See *op. cit.*, pp. 44-45.

like *Flora's drying herself* and *Flora's drying herself with a towel* are different.⁸⁶ Call this the *proliferation claim* (pc).

According to (pc), there are at least two distinct events, *Flora's drying herself* and *Flora's drying herself with a towel*, whose existence sentence (i) is talking about. Also, if (cc) is right, any of the two events is a value of variable '*e*' that makes (ii) true. Then, the following two sentences are true:

The particular event of *Flora's drying herself with a towel* was done with a towel,

and

The particular event of *Flora's drying herself* was done with a towel.

The consequence is that 'with a towel' indicates a property of some individual events, coming in contradiction with (mc). For one thing, by asserting (mc), Kim holds explicitly that modifier 'with a towel' does not indicate a property of any individual event.

Kim's interest in (mc) comes mainly from the idea that, if the predicate modifiers indicating means-manners-methods were taken as signaling properties of some individual events, one would miss what initially stood as reason for the commitment to a structured complex view of events:

[Events] were intended to be entities that enter into causal relations with one another, and that can be objects of explanations. But it is clear that we may want to explain not only why Sebastian strolled, i.e., Sebastian's stroll, but also why he strolled leisurely, i.e., his leisurely stroll. Under the approach being considered [namely, the one in which predicate modifiers indicating means-manners-methods stand for properties of some individual events], the second explanation would be of why Sebastian's stroll

⁸⁶ See *op. cit.*, p. 43.

was leisurely; we would be explaining why a certain event had a certain property, not why a certain event occurred. (p. 45)

Actually, letting aside (cc), (mc) appears to have problems by itself. It seems kind of analytically true that the event of *Flora's drying herself with a towel* was done with a towel. It follows the puzzling conclusion that, whenever a predicate modifier indicating means-manners-methods modifies the *constituent* property of some individual event, the modifier also says something (in fact, the same thing) about the individual event taken as a *whole*.

It is also (pc) that brings some unwanted difficulties by itself. Let us consider two sentences:

- (3) Jones buttered the toast slowly, with a knife.
- (4) Jones buttered the toast slowly.

There is no problem in seeing that sentence (3) entails sentence (4). Unfortunately, how the account of the entailment is to be done is not obvious to the same degree.

Applying Davidson's semantical prescription, (3) and (4) are to be rewritten as:

(3D) $(\exists e)$ (e was a buttering of the toast by Jones, and e was done slowly, and e was done with a knife)

and, respectively,

(4D) ($\exists e$) (*e* was a buttering of the toast by Jones, and *e* was done slowly) Then, it is obvious that (3D) entails (4D), accounting for why (3) entails (4). In accordance with (pc), Kim holds that property *stabbing* is not the same as property *stabbing with a knife*,⁸⁷ and hence neither *buttering slowly* is the same property as *buttering slowly, with a knife*. It follows that the predicate '*a* butters *b slowly*' picks up a different property than that picked up by the predicate '*a* butters *b* slowly, with a knife'. But how could Kim account for the entailment from (3), a sentence informing about the occurrence of *Jones' buttering the toast with a knife slowly*, to (4), a sentence informing, according (pc), about the occurrence of a different event, namely, *Jones' buttering the toast with a knife*?

According to the semantical account of event sentences that I take to follow from Kim's view on events, the covert logical forms of (3) and (4) could be explicitly written as follows:

(3K) ($\exists < x, y, t, R >$) (<*x*, *y*, *t*, *R*> = the exemplification of *R* by *x* and *y* at *t*, and *x* = Jones, and *y* = the toast, and *R* = *a* butters *b* slowly, with a knife) and, respectively,

(4K) ($\exists < x, y, t, R >$) (<*x*, *y*, *t*, *R*> = the exemplification of *R* by *x* and *y* at *t*, and *x* = Jones, and *y* = the toast, and *R* = *a* butters *b* slowly).

Then, (3K) and (4K) are simply two quite distinct sentences as logically unrelated to one another as (4K) is with:

 $(\exists < x, y, t, R >)$ (< x, y, t, R > = the exemplification of *R* by *x* and *y* at *t*, and *x* = Jones, and *y* = the toast, and *R* = *a* burns *b*)

In these conditions, if one is committed to Kim's property-exemplification account of events, one may find reasonable to adopt the conjunctional view of properties. This view takes, for example, the "*buttering - buttering slowly, with a knife*" relation to be one holding between a conjunct and its conjunction. Accordingly, (3K) and (4K) can be rewritten as follows:

(3K*) ($\exists < x, y, t, R >$) (<*x*, *y*, *t*, *R*> = the exemplification of *R* by *x* and *y* at *t*, and *x* = Jones, and *y* = the toast, and *R* was a buttering, and *R* was done slowly, and *R* was done with a knife)

and, respectively,

 $(4K^*)$ ($\exists < x, y, t, R >$) (< x, y, t, R > = the exemplification of R by x and y at t, and

x = Jones, and y = the toast, and R was a buttering, and R was done slowly).

Then, it is obvious that $(3K^*)$ entails $(4K^*)$, eventually accounting for the entailment from (3) to (4).

The proposed solution, however appealing it may be, is held to miss its target. For one reason, it is hardly plausible that *being done slowly* and *being done with a knife* are genuine properties. Fortunately, the next sections will advance progressively an alternative, hopefully more tenable, proposal.

4.5 Troubles brought, this time, by complex events

In this section, I shall try to emphasize a challenge for any theory of events; it concerns the case when an action is done by more than one agent. Let us consider the following event sentence:

(6) Ilf and Petrov wrote "The twelve chairs."

It entails the falsity of the following sentence:

(6*) Ilf wrote "The twelve chairs."88

Instead, sentence

(7) Ilf and Petrov listened to the latest news in the dining room, at evening does not entail the falsity, but rather the truth of the following sentence:

(7*) Ilf listened to the latest news at evening.

Both (6) and (7) describe actions whose authors are two agents: both 'writing <<The twelve chairs>>' and 'listening to the latest news' are predicated in each sentences of two agents. However, whereas (7) speaks about a complex event – *the listening to the latest news by IIf and Petrov* – whose constituent events are *IIf's listening to the latest news* and *Petrov's listening to the latest new*, (6) speaks about a complex event – *the writing of "The twelve chairs" by IIf and Petrov* – which does not have any proper constituent event.⁸⁹ It seems, then, that the constitution of an event might be

⁸⁸ There is a similar relationship between sentences 'Ted and Mary conceived Johnny' and 'Ted conceived Johnny' or between 'Dan and Christian lifted the 150 kg weight' and 'Dan lifted the 150 kg weight.'

⁸⁹ As will become clear in the next section, the former event, unlike the latter, has a categorial structure.

crucial for what kind of entailments hold from the event sentence speaking about the event.

At this point, intuitively, it seems that, in the case of certain actions (e.g., writing a novel, playing a symphony, etc.) that occur by virtue of the cooperation among certain agents, the sentence expressing the action entails the falsity of the sentence attributing the action only to some of these agents. However, there are also event sentences that speak about actions that are undertaken individually by an agent (e.g. listening in the news, wishing for a glass of vodka, etc.), even though more than one agent is mentioned as undertaking the same action.

Perhaps, it is not only the logical form of the event sentence in question, but also the *meaning* of the expression denoting the collective event-type that imposes constraints on what information can be abstracted from the initial content. Taking the Davidsonian route, the entailment from sentence (7) to sentence (7^{*}) would be obvious, if one were able to reveal that sentence (7) has the following logical form:

(7D) $(\exists e_1)$ (e_1 was a listening to the latest news by IIf, and e_1 was done in the dining room, and e_1 was done at evening) and $(\exists e_2)$ (e_2 was a listening to the latest news by Petrov, and e_2 was done in the dining room, and e_2 was done at evening).

Unfortunately, the Davidsonian analysis does not enable one to go deeper than the first level, where the event of listening to the latest news by IIf and Petrov is treated as a whole.

119

The Kimian account would not do better. Probably Kim would say that (7) talks about two exemplifications of the relation a *listens to* b, one by (IIf, the latest news), the other by (Petrov, the latest news), whereas (6) speaks only about one exemplification of the relation a *write(s)* b.⁹⁰ Yet, these claims are in need of a theory of property exemplification to illuminate, for instance, why the exemplification of *listening to* by IIf and the exemplification of *listening to* by Petrov.

4.6 Hunting for a metaphysical account of events

As I have held before, I believe that the best way of making plausible the claim that events exist is to try to give a coherent metaphysical theory of events. According to the considerations made in the previous sections, such a theory should be able to provide an answer to each of the following problems:

- how it is to be understood the nature of events whether they are structured or not; and, if structured, what structure they have;
- accordingly, how events are to be individuated externally or internally –, and what distinguishes them from other similar entities such as states;
- 3) what logical form accounts better for the entailments emphasized above.

In the next sections, I shall try to advance some solutions⁹¹ to the questions adumbrated above.

⁹⁰ It is a = (IIf, Petrov) and b = "The twelve chairs" that instantiate the relation a write(s) b.

⁹¹ By that, I do not want to imply that my perspective on events is entirely original. Rather, my intention, mostly, is to make clear with what of actual theories my own sympathies lie.

4.7 Categorial structures and metaphysical structures

This section is devoted to the following preliminary question: what is to be understood by 'having a structure' when it applies to an ontological entity? Before considering any answer, it is useful to make some opening clarifications. First, I shall work under the assumption that there is not only one ontological category, i.e. that not all entities have the same nature.⁹² Second, there is a special sort of relations that are often referred by 'metaphysical relations'. Consider, for instance, *being a cause of ..., being a part of ..., being an instantiation of* (or, more neutral, *being an exemplification of*) ... and *being identical with* They make by themselves separate subject matters for metaphysics. Thus, in what follows, I shall reserve 'relation' for referring only to a relation that does not belong to so-called category of metaphysical relations.

Let us consider *x*, an entity. It could have at most two kinds of structures. A structure of *x*, if it has any, is *categorial* if the entities featuring into the structure are only entities belonging to the same ontological category as *x* does. The constituents of some categorial structure could be either proper parts of x^{93} , or entities involved by *x*, although they do not compose x^{94} .

⁹² Note that the assumption that there is more than one ontological category should not be confused with the stronger claim that there is more than one basic ontological category.

⁹³ In the literature, one can meet various kinds of entities – substances, properties, states, events, etc. – that could be seen as being structured in the sense of having proper parts. If such an entity has proper parts it is tempting to call its structure *compositional*. (Thus, a compositional structure is a species of categorial structures).

⁹⁴ See, e.g., those theories which are committed to complex universals and, in particular, to structural universals. Lewis characterizes the structural universal (if any) as involving other universals: necessarily, something instantiates a structural universal if and only if it has proper parts such that these parts instantiate certain universals. The example he considers is that of the structural universal *methane*. This universal involves the monadic universals *carbon* and *hydrogen*, and the dyadic universal *bounded*: "necessarily, something instantiates *carbon*, each of the *h*s instantiates *hydrogen*, and each of the *c-h* pairs instantiates *bounded*." (1998: 200) It seems that each structural universal is characterized by a pattern. Or, better, any particular instantiating a structural universal has a

Yet, *x* might have one more structure, of another kind, a *metaphysical structure*. Its constituents belong to another ontological category than *x* does. For instance, those theories that give a substrate-attribute account of substances, take any substance to be metaphysically structured. Another example is Kim's theory of events. He views any event as a metaphysical structure characterizable by the triplet <object, exemplified property by the object, time of exemplification>.

Interesting enough, on this account, there are events which are both categorially and metaphysically structured. For example, a Kimian may claim that such an event as *the chess match between Kasparov and Karpov that happened ten years ago* has both a categorial structure, being an event composed of the chess moves made by challengers during the match, and a metaphysical structure in the sense that the happening of the *big* event is constituted of certain exemplified properties, certain exemplifying objects, and certain times of exemplifications.

compositional structure organized according to the pattern specific to the structural universal. I shall say that the structure a structure universal has is a *pattern structure*. (Thus, a pattern structure is the other species of the categorial structures.)

Presumably, a Kimian committed to structural properties would say that there are also structural events in the sense that a structural generic event involves other generic events.

It appears that there is a connection between the fact that a structured universal has a pattern structure and the fact that the particular instantiating the respective universal has a compositional structure: if *U* is a universal having a pattern structure such that it involves universal *V*, then any particular P_U that instantiates *U* has a compositional structure having as a proper part some particular, P_V , such that P_V instantiates *V*. In the same vein, if one takes the Kimian view of events and is universalist regarding properties, one must hold: if the event E_U is the exemplification of property *U* having a pattern structure such that property V is involved, then E_U is a compositional structure having as a proper part the event E_V , which is the exemplification of *V*.

Note, however, that pattern structure of an universal is not always exported to a particular that instantiates, or is the exemplification of, it. We should understand the relation between universal U having a pattern structure and universal V featuring in U's structure in the following terms: (1) an instantiation of U necessitates an instantiation of V. Extrapolating this to a particular, P, that has a pattern structure, one obtains: (2) the existence of P necessitates the existence of any of its actual parts, that is, all parts of P are essential part of P. Further, supposing that (1) is true and P instantiates U, and a part of P, namely, P_V , instantiates V, it is not true that (1) entails (2). In order to assure the truth of (2), beside the truth of (1) it is needed something extra, namely, (3) V can be instantiated only by P_V . Unfortunately, it is hard to see what an example can be given in this sense. I shall come back to the issue of pattern/compositional structures later.

Generally speaking, by saying that an entity, x, has a structure, it should be understood that *something* – some ontological stuff – is structured in a certain way such that x is obtained. In the Aristotelian style, by 'matter', I shall call what is structured, and by 'form', how the matter is structured. The matter of x could include entities of different ontological categories. Also, the matter of x can be structured in at least one form, namely the form of x. However, in some cases there are at least two forms in which the same matter could be structured. A suggestive metaphor is the following: the H₂O molecules of a flake of snow are the same as the H₂O molecules of the drop of water obtained after the melting of the respective flake of snow.

The form of *x* is that ontological constraint over the entities featuring into the matter of *x* that renders *x* as it is, with a certain structure. For instance, for a Kimian, there are events which have the following metaphysical form: *the exemplification of* p_1 *by* p_2 *and* p_3 *at* p_4 , where p_i is a *place* to be filled up (occupied) by some entity of a certain ontological category. Hence, if the Kimian is realist about the universals and substantialist about time, the matter that is structured in such a form in the case of the event John's laughing at Tom at 2 p.m. consists of the following entities: the substances *John* and *Tom*, the relation a *laughs at* b, and the time 2 *p.m.*

It should be noted that one has to resist the temptation of claiming:

(a) A structured entity is nothing over and above its matter and the form that structures that matter.

The reason is the following. If (a) is true then the direct consequence would be:

123

(b) Two structured entities, x and y, are identical if and only if the following conditions are met: the matter of x = the matter of y, and the form of x = the form of y.

Let us consider two events: *John's laughing at Tom at 2 p.m.* and *Tom's laughing at John at 2 p.m.* Given that the one could have existed without the other, the two events are distinct. Yet, the matter of the first is the same as the matter of the second; the form of the first is identical with the form of the second.⁹⁵

But, then, besides the matter and the form corresponding to a certain entity, what else is a structured entity? To answer to this question we have to consider the notion of structuration (or structuralization): *the structuration of a matter by a form is the particular occupation of the form by the entities of the matter*. I take a form to be a sort of pattern or configuration of "places" to be occupied by the entities of a matter; the respective "places" are disposed in accordance with a "syntax" of the metaphysical relations specific to each kind of entity and type of structure. Thus, it seems correct to say that the structuration of a matter by a form is simply the allocation of a certain place in the form to each entity of the matter.⁹⁶ The consequence is:

A structured entity, x, is nothing over and above its matter, a form of that matter, and the structuration of the matter in that form. That is, a structured entity is nothing besides some entities, some places in a form, and the matching between the entities and the places.

⁹⁵ This argument was partly suggested to me from one of Armstrong's arguments for the claim that the complete constituents of a state of affairs are capable of being the complete constituents of a different state of affairs. See Armstrong (1999a: 204).

⁹⁶ That is to say that the structuration of a matter by a form is the particular occupation of the places of the form by the entities of the matter.

I shall adopt the following notations:

- $[p_1, p_2, ..., p_i, ...]$ stands for a categorial form, F_c , having as places $p_1, p_2, ..., p_i, ...,$ where *i* shows the indexation of the place p_i relative to F_c ;
- < p_1 , p_2 , ..., p_i , ...> stands for a metaphysical form, F_m , having as places p_1 , p_2 , ..., p_i , ..., where *i* shows the indexation of the place p_i relative to F_m ;
- * p_1 , p_2 , ..., p_i , ...* stands for a form in a generic sense;
- $\{m, n, ...\}$ stands for some matter, *M*, where *m*, *n*, ... are the entities of *M*;
- m_i stands for the fact that place p_i is allocated to the entity m_i .

Hence a formal expression of the above consequence is:

If *x* is characterized by matter $M = \{m, n, ...\}$, and by form

 $F = *p_1, p_2, ..., p_i, ...*$, then x is nothing over and above the structure:

 $S = *...m_s, ..., n_r, ...*.$

A further consequence appears to be:

Two structured entities are identical if and only if their structures are identical.

That is to say that, considering that x and y are structured entities, x = y iff:

- the matter of x = the matter of y = M;
- the form of x = the form of y = F;
- the structuration of *M* in *F* in *x*'s case is the same as that in *y*'s case.

Let us consider a structured entity, *x*, such that *x* can be truly described as having a structure, S_i . As I have shown above, S_i is obtained by a certain structuration of the matter of *x*, namely M_i , in a certain form of *x*, F_i . Further, let us suppose that M_i is made of some entities having each a certain structure. Of course, such an entity

belonging to M_i is obtained by a certain structuration of its matter. Consequently, x can be said to have, at various levels of details, various structures. Note also that, corresponding to each level of detail, there are specific entities featuring in the respective structure. The constituents belong only to those ontological kinds admitted at the respective level.

At the next level of detail after i,⁹⁷ the corresponding structure, S_{i+1} , has a matter, M_{i+1} , made of the entities existing in the structures of the entities belonging to matter M_{i} , which is structured as S_{i} . This is the reason for which I take M_{i+1} to be a more fine-grained matter than M_i is. Moreover, the structuration of M_{i+1} need a more complex form than that needed for the structuration of M_i . The deeper level of detail is, the more complex the form of an entity is. Form F_{i+1} in which M_{i+1} is structured is a ramification (an elaboration) of form F_i . If one goes on from level *i* to level *i*+1, some new forms will be revealed within the old form, F_{i} , it is about those forms of the structured entities belonging to M_i that are point-like forms at the level *i*.

I believe that there is a suggestive metaphor for the idea of revealing progressively various structures of a structured entity that has, once one goes on to the next level of refinement and complexity, more and more complicated forms and more and more fine-grained entities filling up those forms. It is the metaphor of a slide-show of pictures taken from satellite and presented in the order of increasing detail and complexity. Thus, watching such a slide-show, one can see, at the beginning, a continent, say, with its forms of relief and their disposition; then, going closer, a

⁹⁷ In many cases, "the next level of detail" is only a way of saying, because there could be many levels of detail very close to each other and going from i to i + 1 could be a jump over some of them.

region of that continent; then, a town of which one can see only the main (major) streets, parks, lakes and buildings. The level of detail can reach such a scale that one can notice cars, people, trees and even small rocks.

A last question that I would like to approach here is what relation is between categorial structures and metaphysical structures when they coexist. Let us consider an account that embraces the idea that the kind of entities to which *x* belongs could have both categorial and metaphysical structures, and *x* actually has structures from both types. For instance, let x_c be an entity featuring in one of *x*'s categorial structures, and x_M an entity featuring in one of *x*'s metaphysical structures. I call an entity such as x_c a C-constituent of *x*, and an entity such as x_M a M-constituent of *x*.

On the assumption that both x_c and x belong to the same ontological category, it follows that x_c could have both types of structures and hence both types of constituents. Let us suppose that relative to x_c , x_{cc} is a C-constituent, and x_{cM} a Mconstituent. Let us also suppose that x_M is such an entity that has an internal structure, and x_M * features in the respective structure. For an illustration of the above ontological scenario, see Figure 4.1.

Now, it seems natural to say the following:

- x_{CC} is a C-constituent of x,
- x_{CM} is a M- constituent of x, and
- x_M^* is a M- constituent of x.

It is pretty obvious that the above three elements figure in one or another structure of *x*. Since x_{CC} belongs to the same ontological category as x_C , and x_C to the same

CEU eTD Collection

ontological category as *x*, it follows that x_{CC} and *x* are from the same category, and hence x_{CC} features in a categorial structure of *x*. Relative to *x*, the difference between x_C and x_{CC} is that they feature in categorial structures corresponding to different levels of detail.



Fig. 4.1

CEU eTD Collection

For the sake of illustration I shall use the Kimian view of events. Let us consider the event of *the chess match between Kasparov and Karpov*.⁹⁸ On the one hand, the event of *Kasparov's moving his knight from F3 to G5* is an event-part and hence a C-constituent. On the other hand, the property *chess matching* exemplified by Kasparov and Karpov is an M-constituent.

⁹⁸ For simplicity I shall ignore the temporal component.

Relative to the event of *Kasparov's moving his knight from F3 to G5*, the event of *Kasparov's grasping his knight at F3* in order to move it to G5 is just a event-part and hence a C-constituent. Since a part of a part is a part of the whole, it follows that *Kasparov's grasping his knight at F3* is a C-constituent of *the chess match between Kasparov and Karpov*. Instead, *Kasparov* is only an M-constituent of the part-event *Kasparov's grasping his knight at F3*; therefore, *Kasparov* is also a M-constituent of *the chess match between Kasparov and Karpov*.

The property *chess matching* exemplified by Kasparov and Karpov is an ordered conjunction of various properties exemplified by the chess movements of the challengers during the match.⁹⁹ So, *opening the match* is a property involved by the property *chess matching*, and is exemplified by Karpov. Given that *opening the match* figures in the structure of a property which figures in the internal structure of the event *the chess match between Kasparov and Karpov*, it follows that *opening the match* is a constituent of the respective event. Of course, according to the above distinction, it is an M-constituent. In the same vein, *moving Kasparov's horse from F3 to G5* is a property figuring in the structure of *chess matching* and hence is an M-constituent relative to the event of *the chess match between Kasparov and Karpov*.

The above discussion suggests that, on the considered account, an entity such as an event could have a series of categorial structures. The series starts with the event, *e*, and goes top-down to more and more detailed categorial structures, until it reaches a categorial structure with constituents having no categorial structure. Call the structure

⁹⁹ In this sense, it is not wrong to say that *chess matching* is a (temporally) structured property.

of such constituents *basic categorial structure*. Moreover, given that, on Kimian account of events, any event has a metaphysical structure, it follows that events existing in any categorial structure of *e* have their metaphysical structure. The consequence is that, corresponding to any categorial structure of *e*, there is a metaphysical structure of *e*. The respective metaphysical structure of *e* is the result of categorially structuring of certain metaphysical structures. Note that the metaphysical structure corresponding to the basic categorial structure is not necessarily the basic metaphysical structure. For it could be constituted of structured entities.

For illustration let us consider an event, *e*, having a macro-metaphysical structure, $\langle s, P \rangle$, in which *s* is a substance or a complex of substances, and *P* is a property instantiated by *s*. Assume further that *e* is composed of events e_{11} , e_{12} , ..., and, also, let us admit that any event e_{1i} has a metaphysical structure, $\langle s_{1i}, P_{12} \rangle$. Note that s_{1i} could be even *s*. So, on level₁, the categorical structure of *e* is $[e_{11}, e_{12}, ...]$, and its metaphysical structure is $[\langle s_{11}, P_{11} \rangle, \langle s_{12}, P_{12} \rangle, ...]$. If there are compounded events among e_{11} , e_{12} , ..., each of them will be treat in the same way as above. Thus, the next level, namely, level₂, will be reached: *e* has a categorical structure, $[e_{21}, e_{22}, ...]$, and a metaphysical structure, $[\langle s_{21}, P_{21} \rangle, \langle s_{22}, P_{22} \rangle, ...]$. Of course, it could be the case that e_{2j} is identical with e_{1t} if e_{1t} is a simple event (i.e. it has not any event-parts). In other words, simple events from one level are preserved to the next level. The same operation of revealing a deeper structure will be repeated until it is obtained the most basic categorial structure of *e*, i.e. a structure in which only categorially unstructured events (if any) feature. Going on deeper than this level, and nearer to the microlevel, one will find the same categorial structure, but a more complex

130

metaphysical structure. The metaphysical structures become increasingly complex as the structures of various M-constituents are uncovered. See Table 4.1.

Level	Categorial	Metaphysical structure
	structure	
the macrolevel	E	<s, p=""></s,>
1	[e ₁₁ , e ₁₂ ,]	$[< s_{11}, P_{11}>, < s_{12}, P_{12}>, \ldots]$
2	$[e_{21}, e_{22}, \ldots]$	$[< s_{21}, P_{21}>, < s_{22}, P_{22}>, \ldots]$
:	:	:
:	:	:
n =	[e _{n1} , e _{n2} ,]	$[< S_{n1}, P_{n1}>, < S_{n2}, P_{n2}>, \ldots]$
the basic categorial		
level		
n + 1	[e _{n1} , e _{n2} ,]	$[<[s_{n1}^{(1)}, s_{n1}^{(2)}, \ldots], P_{n1}>, ,$
]
:	:	:
:	:	:
the microlevel =	[e _{n1} , e _{n2} ,]	$[<[s_{n1}^{(1)}, s_{n1}^{(2)}, \ldots], [P_{n1}^{(1)}, P_{n1}^{(2)}, \ldots]>,$
the basic metaphysical]
level		

Table 4.1

4.8 Events as changes in properties

I am committed to the idea that the world contains both substances and events – two ontological kinds of particulars. It is a familiar point that, in the most cases, events are things which happen to substances. Also, it seems entirely correct to say that substances do have a different kind of existence than events do. In this vein, Crane (2001a: 36-7) rightly remarks that, since substances persist through time through changes in *their* properties, they are wholly present at every time at which they exist. Instead, events "consist of changes in *other* things" and have temporal parts. Thus, one may say that events have a sort of *existence distributed in time*. Lowe (2002) makes a very similar point:

We can remark that events occur at times or last through periods of time, whereas things persist over time or from one time to another; that events happen to or befall things, whereas things participate in events; and that events may uncontentiously have temporal parts or stages – earlier and later phases – whereas things apparently do not. (p. 244)

It is a matter of debate whether events are reducible to substances. According to Davidson's view, substances and events belong respectively to two basic and hence mutually irreducible ontological categories.¹⁰⁰ But, it seems fair to object, as Lowe (2002) does, that:

[such a non-reductive pluralist position] leaves inherently mysterious the relationship between an event and the things which participate in it (...) [the relationshion] must itself be something basic and incapable of analysis or explication. If so, this is deeply unsatisfactory feature of the non-reductive pluralist's position. (pp. 244-5)

Thereby, it is tempting to adopt a view on events similar to Kim's in the sense of making the event-substance relationship intelligible in virtue of positing for any event the existence of a metaphysical structure. Thus, by claiming that an event is

¹⁰⁰ See Davidson (1980b [1969]).

characterized by triplet $\langle s, P, t \rangle$, in which *s* stands for a substance undergoing a property change, *P* stands for the property acquired or lost by *s*, and *t* stands for the moment when the change occurred, Kim (1993a [1976]), for example, explains the event-substance relationship as being one that holds between an entity having a certain metaphysic structure and a constituent filling up one of the places of the form of the metaphysic structure. More specifically, in Kim's view, the participation of some substance in some event should be understood as the participation had by that substance in some change of its properties:

The term 'event' ordinarily implies change, and most changes are changes in a substance. A change in a substance occurs when that substance acquires a property it did not previously have, or loses a property it previously had. (p. 33)

Similarly to Kim, Lowe (2002) considers that "any event is a change in the properties of, or in the relations between, some thing or things" (p. 237). He views his proposal as leading to a *thing-ontology*, that is, "a reductionist or eliminativist stance towards events in favor of the ontological primacy of things" (idem).¹⁰¹ His *monistic ontology*, as Lowe notices, opposes both to any *event-ontology*, the other kind of monistic ontology,¹⁰² and to any pluralistic ontology.

¹⁰¹ Lowe (2002: 241) says that his proposal may be given either a reductionist or an eliminativist gloss. The reductionist will say that events are "ontologically dependent entities, depending for their existence and identity upon things, properties, and relations in terms of which they are to be explicated". Instead, the eliminativist will hold that "we can, and in principle should, replace all of our talk about events and their properties and relations by talk of change in the properties and relations of things". Lowe mentions that an example of recent defense of eliminativism regarding events is Michael Tye's *The Metaphysics of Mind* (Cambridge: Cambridge University Press, 1989), ch. 2.

¹⁰² Lowe (2002) emphasizes that, typically, modern physics prescribes an event-ontology. In this vein, one could be committed to a reductionistic event-ontology: "A complete description of the physical world can in principle be supplied by specifying what is happening at each such point of space-time, that is, in terms of the events occurring at all such points. What we ordinarily think of as being a persisting object is just a series of events (a process) which occupies that object's so-called 'world-line'." (p. 233) As examples of such a scientifically inspired event-ontologie, Lowe mentions W. V. Quine's 'Identity, Ostension, and Hypostasis' in his *From a Logical Point of View*, and Alfred North Whitehead's *The Concept of Nature* (Cambridge: Cambridge University Press, 1920). According to Lowe, another kind of event-ontology one can meet in modern physics is an eliminativist one: "Talk of persisting objects should be replaced by talk of events and processes, at least in a fundamental scientific description of the physical world." (p. 234)

Since Kim takes the constitutive substance of an event to be essential to the identity of that event, it may not be inappropriate to say that his account of events implies a thing-ontology too. But, Kim insists that it should not be given an eliminative gloss:

[The advanced account] doesn't attempt to show that events are in some eliminative sense "reducible" to substances, properties, and times. It attempts to tell us something about the metaphysical nature of events by relating them to such other ontological categories as substances, properties, and times. (p. 36)

I endorse his claim for the simple reason that, by admitting that entities of a certain kind has a certain kind of metaphysical structure, one is not committed to the idea that the talk of those entities should (or could) be replace by a talk of the constituents featuring in their structures. Besides its constituents, there is one more thing that is crucial for the identity of a metaphysical structure, namely, *how* its constituents occupy its form. The idea is that the same constituents might occupy various places in the same form to the effect that there might be two distinct entities of the same kind, having the same form and the same constituents.

In what follows, I assume that an event is occasioned by a change in some properties. I believe that Kim justly notes that, if events involve changes, states seem be "unchanges". Examples of events are: my changing my mind, your opening the door, her raising the right hand; whereas, examples of states are: my having a belief, your being in front of the door, and her keeping the right hand hanging loose. If there were no property change in a substance, there would be no event happening to that substance. For example, a piece of metal's melting is an event only because there are certain changes in some of the properties of that piece of metal like its temperature, shape, atomic structure, etc. Unfortunately, Kim chooses eventually not to take the dichotomy of events and states too seriously. His most substantial reason, I believe, is that "some properties already imply changes in the substance that has them; for example, fading in color, falling, and freezing" (p. 34). Therefore, he concludes, "a change need not necessarily be characterized as a losing or acquiring of some property it may simply be the *having* of some property at a time" (*Idem*).

One may be tempted to make the following distinction. *Prima facie* there seem to be two kinds of properties – let us call them *static properties* and, respectively, *dynamic properties*. Examples of dynamic properties would be: falling, exploding, pushing, ripening, walking, ageing, dying, laughing, thinking, counting; whereas, examples of static properties would be: having a certain position in space, being red, weighting, being extended, believing, being dead, lying. Perhaps, dynamic properties would be exemplified on the occasion of some changes in properties. Moreover, a dynamic property would not be wholly exemplified at a given moment, but rather its exemplification would take as long as would be needed for the respective change to be done. For example, the dynamic property of ripening is exemplified on the occasion of, and during the change of the property of having a certain biological structure and chemical composition.

Thus, Kim would be able to preserve his healthy intuition of identifying events with changes and of opposing them to states, by slightly completing his theory of events: an event would be the exemplification of a dynamic property by a substance during a period of time. Instead, a state would be the exemplification of a static property by a

CEU eTD Collection

135

substance at a moment of time. In this way, it would become clear in what sense events are distinct from states, yet the two ontological categories being related: an event would happen to some substance when and only when that substance would undergo some changes in his states.

However, I believe that it would be wrong to regard verbs signifying the occurrence of certain changes in properties as indicating properties because, if we take that option, we should accept that any change in properties occasions the exemplification of some property (probably, a dynamic property). Such an acceptance would lead to disastrous consequences. For one thing, since, as Kim points out, any acquiring of a property is also a property change,¹⁰³ by construing any property change as a property exemplification, it follows that any property acquiring (i.e. any exemplification of a new property) is attended by the exemplification of one more property. Following this line, one ends up with an unacceptable infinite regression in the sense that any exemplification of a property involves the exemplification of an infinite number of properties.

¹⁰³ It is a property change simply because the respective substance has not had the property and, now, has it.

4.9 Events as temporal structures of states

I believe that what Kim takes to be an event, namely the exemplification of a property by a substance at a moment of time, is not really an event, but rather a state of the respective substance. Further, an event happens to a substance when and only when that substance undergoes a change in the exemplification of some of its properties¹⁰⁴ during a period of time. Thus, the whole thing comes to the idea that an event is a change happening to some states of a substance during a period of time.

Given that any event is a succession of states, it seems reasonable to hold that any event has a metaphysical structure characterized by a form that is temporal in nature, its places being occupied by states. Briefly, any event has a temporal metaphysical structure made of states: $\langle S_{initial}, ..., S_{i}, ..., S_{final} \rangle$, in which $S_{initial}$ stands for the initial state, i.e. the state that is firstly changed, S_i stands for an intermediary state, and, of course, S_{final} stands for the final state, i.e. the result of state change. But any state has its own metaphysical structure, $\langle s, P, t \rangle$, in which *s* stands for the substance

¹⁰⁴ Lowe (2002: 238) rightly notes that only intrinsic changes in objects amounts to genuine events. An example given by Lowe of such change is a change in the color: an object is green prior to *t* but it begins to be red at *t*. Thus, purely relational changes (also called Cambridge changes) do not counts as events. An object undergone a relational change inasmuch as something to which it is somehow related has undergone an intrinsic change. According to Lowe, examples of Cambridge changes are the following: the change which occurred when Xanthippe became a widow upon the death of her husband, Socrates, and the change which occurred when Mary (who has stopped growing) becomes shorter than her brother, Tom, because Tom has suddenly grown (*idem*). Also, something getting older (the mere passage of time) is not an event. "We can perfectly easily conceive of an object which gets one day older without changing intrinsically in any way at all. So it seems that an object's merely getting older is not in itself a matter of that object's undergoing some kind of intrinsic change." (*Idem*)

In the same vein, regarding the properties of whose acquisition would matter as genuine events, Kim (1993a [1976]) takes a number of justified precautions: "It clearly will not do to count as an event the exemplification of any arbitrary property by an object. This becomes obvious if one thinks, as many do, that there is a property expressed by any open sentence, or if one thinks of properties in the way modal logicians do nowadays, namely as functions from possible worlds to sets of individuals. There is also the problem of properties ordinarily considered generic events (e.g. becoming a widow), which give rise to <<Cambridge events>>." (p. 37) According to Kim, what should guide us in picking out properties whose change are events are scientific theories: "They are among the important properties, relative to the theory, in terms of which lawful regularities can be discovered, described, and explained." (*Idem.*)
being in the respective state, *P* stands for a property exemplified by the substance, and *t* stands for a moment of time of the exemplification. Naturally, then, at a higher level of detail, any event has a temporal metaphysical structure consisting of a substance (i.e., an entity persisting through changes through time), and a range of properties the substance exemplifies successively; the form of the structure is given by a certain *timing* of property exemplification during the period of the considered change.

Any change is time consuming; it involves a period of time during of which a substance usually loses properties in favor of acquiring others. Such a change in properties is made, in the case of each event, in a specific manner. In most of the cases, for characterizing (individualizing) an event, it is not enough to provide the initial state and the final state; it is also necessary to mention the profile of the transition between those two states. This kind of profile is given both by the intermediary states and by the timing of transition.

Before closing this section, one remark is in order. If any event is a change in property exemplification from an initial property to a final one, a question arises: except their being exemplified by the same substance, what else, if any, do those properties have in common? The question is mainly motivated by the fact that a substance can undergo more than one property change in the same period of time, that is, more than one event can happen in a substance in a certain amount of time. Let us say that a substance has properties P_i and P_i^* at time t_i , and, later, instead those properties, it gains other two properties, P_f and P_f^* . The question is: which

property change, that from P_i to P_f or that from P_i^* to P_f^* , amounts to the occurrence of an event?

Let us consider that a child brings a snowball in a house, where is quite warm. After a while, the snowball melts leaving a small pool of water. Some property changes happen on that occasion: the solid, white, spherical substance becomes a liquid, transparent substance having the shape of a small pool. How is an event to be identified here? Do not say that the event in question is *the melting of the snowball*. Of course, that is an event quite all right, but it is the *whole* event. We are interested here in those "smaller" events that compose *the melting of the snowball*. It seems obvious that there are a number of events that happen to the snowball in the same period of time: the snowball changes simultaneously its consistency, its color, its shape, etc. It would seem weird to say that *the snowball's having white color* and *the snowball's having the shape of a pool* may be taken as the initial state and the final state of some event happening to the snowball, respectively

Speaking about properties, it would be useful here to recall the notorious distinction between determinables and their determinates. Further, it seems obvious that a determinable cannot but be exemplified as one of its determinates. For example, a substance can exemplify determinable *having weight* only as determinable *having a certain weight*. Likewise, *having shape* is a determinable which is always exemplified as *having a certain shape*: for instance, *being a cube* (that is, *having the shape of cub*), or *being a sphere*.

A consequence: the state had by a substance in virtue of the exemplification of a determinable at a moment of time is the same as the state had by the substance in virtue of the determinate the determinable is exemplified at that moment of time. In other words, a proposition conveying information about the exemplification of a certain determinable by some substance at a certain time is about the same state as that referred by another proposition that conveys information about the exemplification of the determinate by the substance at that time.

Let us consider again the question 'In the case of some property change, what does the lost property and the acquired property have in common that makes their exchange in exemplification to be an event?' The answer, as it follows from above considerations, seems to be obvious: those properties are determinants of the same determinable. For example the event of changing the color of somebody's face is a change in property exemplification from bright red to lavender, for instance, where *being bright red* and *being lavender* are determinants of the same determinable, namely, *having a color*.

4.10 Criteria of identity

Different concepts of event tend to suggest different criteria for event individualization. It seems that what identity criterion one adopts for events is highly dependent on what view of events one is committed to. Of particular importance for such a criterion is whether events are taken to have metaphysical structure. If one considers events as monolithic entities (i.e. entities without any metaphysical structure), any event gains its identity only from its relations with other entities. Instead, if events are seen as structured entities, the individuality of any event is given both by its constituents and by how constituents are structured.

Davidson (1980b) is representative for the former case. By assuming that all events have causes and effects, he claims that events are individuated by their causes and effects. That is, if x and y are events, then x is identical with y if and only if x and y have the same causes and the same effects. But, given that a further assumption of Davidson is that the protagonists of causal relations are events, the causes and the effects of any event are events as well. Therefore, it can be objected that Davidson's criterion leads to a vicious circularity, because, on his proposal, the identity of any event presupposes the identity of other events.

Since Kim (1993a [1976]) holds that events are structured entities, he, as I have noticed above, is able to use their metaphysical structure in giving an identity condition:

[E]vents are identical just in case they have the same constitutive property, object, and time. (p. 36)

Yet, Kim claims that it is the constitutive object that is the crucial element in establishing event identity:

The fact that someone other than Sebastian could have taken a stroll in his place does not make it the case that the very stroll that Sebastian took could have been taken by someone else. If Mario had been chosen to stroll that night, then there would have been another stroll, namely Mario's. (...) Only Socrates could have died *his* death. (p. 48)

Like Kim, I find it plausible that events are structured entities. But, as I have held in the previous sections, I have a different perspective on how they are structured, which leads me also to a different claim regarding their identity. In the previous section, I have argued for the thesis that any event has a temporal metaphysical structure consisting of a substance (or a system of substances), and a range of properties that substance exemplifies successively, in accordance with a certain *timing*, during a certain period. In other words, an event is characterized by $\langle s, P_i, P_f, (t_i, t_i), p \rangle$, in which *s* is a substance (or a system), P_i is an initial property had by *s*, P_f is the final property *s* acquires instead of P_{i_i} (t_i , t_f) is the period needed for change to be accomplished, and *p* is the profile of change. Profile *p* consists of the intermediary properties *s* acquires and looses in going from its initial state to its final state, plus the respective moments or periods when *s* has each intermediary property.

In consequence, within a given world, an event is uniquely referred by $\langle s, P_h, P_h, (t_h, t_i), p \rangle$, that is, by phrase 'the change obtained in substance *s* by replacing the exemplification of property P_i with the exemplification of property P_f during period (t_h t_i) by following a transition of profile p'. (See Figure 4.2.) Accordingly, the criterion of event identity is: *two events are identical if and only if they have the same constitutive substance, the same initial property, the same final property, the same period of time, and the same profile of property change.* As a corollary: *two sentences are about the same event if and only if they are about the same substance that undergoes a change in exemplification between the same properties in the same period, following the same profile of change.*



Recall that, in virtue of his (pc), Kim claims that property *buttering* is different from property *buttering slowly*.¹⁰⁵ I have shown that such a distinction puts Kim in difficulty when it comes to accounting for entailments of a certain kind. Let us consider two sentences:

- (5) Jones buttered the toast slowly.
- (6) Jones buttered the toast.

Obviously, (5) entails (6). However, Kim's theory of events does not seem fit to account for this entailment. For one thing, given his criterion for identity event, neither (mc), nor (pc) allows Kim to acknowledge that both sentences speak about the same event, which, obviously, they really do.

¹⁰⁵ See § 4.4.

I find this a reason enough for rejecting both (mc) and (pc). Some clarification should be made at this point. As I have remarked already, *buttering* is wrongly taken as being a property. Rather, 'a butters b' picks up a property change in the sense that system [a, b] loses a property in favor of gaining another one, both properties being determinants of the same determinable.¹⁰⁶ Consequently, against Kim, I believe that 'slowly' does not modify a constituent property of the event, *buttering* not being such a thing. Rather, 'slowly' modifies the profile of changing from one to another constituent property. Predicate modifiers indicating means-manners-methods – e.g. 'with a knife' or 'slowly' – characterize how events evolve from their initial states to their final states.

I have held that a determinable cannot but be exemplified as one of its determinates. Likewise, the profile of the change in constituent properties of some event, if it is qualifiable, it cannot be but exemplified as qualified. Thus, in action sentences, predicate modifiers indicating means-manners-methods express information regarding such qualifications. Therefore, sentences (5) and (6) speak about the same event, the difference between them being that (5) is more informative than (6) regarding the event. This should lead one to the conclusion that there is a *relation of informational containment* between (5) and (6). It is this relation that grounds the entailment from (5) to (6).

¹⁰⁶ What properties? What determinable? Simply put, the determinable is *the distribution of butter*. Initially, let us say that system [Jones, the toast] exemplifies the determinable as a distribution of butter such that the whole amount of butter is on Jones' right hand as a piece of butter. In the end, the system exemplifies the determinable as a distribution of butter such that a part of butter is widespread on the surface of the toast.

Note that, even though Kim takes the event of *Jones' buttering the toast slowly* to be different from the event of *Jones' buttering the toast*, he says that they are not entirely distinct since the former *includes* the latter. Yet, Kim refrains from making explicit what he means by the relation of inclusion between events.¹⁰⁷ Interestingly enough, Kim is not unaware of the alternative solution of informational containment: 'Jones's buttering the toast slowly' and 'Jones's buttering the toast' are "two descriptions of the same event, one being somewhat more detailed and more informative than the other" (p. 44).¹⁰⁸ However, given his (pc), Kim chooses eventually to reject it in favor of the idea of different, but included events.

Granting (pc), Kim has to admit that any event is attended by an overwhelming number of included events.¹⁰⁹ For example, according to Kim, there were indefinitely many butterings done by Jones that night. How does he manage to deal with such a proliferation of events? Even though, as I have said, he does not provide an account of inclusions between events, Kim insists again to play blindly on the card of event inclusion. He tries to assure us that his proliferation of events "is not in itself serious", the sense in which there would be indefinitely many *butterings done by Jones* being "as harmless as the sense in which there are indefinitely many tables in this room":

We normally count this as *one* table; and there are just so many (a fixed number of) tables in this room. However, if you believe in the calculus of individuals [in the sense of Nelson Goodman, *The Structure of Appearance*, Harvard University Press, Cambridge, 1951], you will see that included in this table is another table – in fact, there are indefinitely many tables each of which is a proper part of this table. For consider the table with one micrometer of its top removed; that is a table different from this table; and so on. (p. 46)

¹⁰⁷ See Kim (1993a [1976]): "I will not try to give a characterization of 'inclusion' for events here; a completely general characterization gets, as far as I know, to be very complicated without being philosophically interesting..." (p. 45)

¹⁰⁸ Actually, there Kim speaks about other two descriptions, namely, 'Brutus's stabbing Caesar' and 'Brutus's stabbing Caesar with a knife'.

¹⁰⁹ See *op. cit.*, p. 46.

Admittedly, on the occasion of Jones' buttering the toast, there are indefinitely many butterings done by Jones, but, against Kim, I believe that this proliferation should not be accepted in virtue of (pc). Moreover, Kim's example of harmless proliferation of entities – namely, the existence of indefinitely many tables, each of them being a proper part of some table – is brilliant, but, I hold, it is not analogical with his claimed (alleged) proliferation of events. Rather, given that the proliferation of tables is possible because of the part-whole relationship, an analogy with the proliferation of events is legitimate only if the events stand in the same kind of relationship. Thus, the trouble is that *Jones's buttering the toast slowly* and *Jones's buttering the toast* do not stand at all in a part-whole relationship.

Since events have an existence distributed in time, their parts are temporal. Thus, an example of event proliferation would be the following series of events: *Jones' buttering the toast during the first 2 minutes after midnight, Jones' buttering the toast during the first 2 minutes and 10 seconds after midnight, Jones' buttering the toast during the first 3 minutes and 11 seconds after midnight, etc. All those events are temporal parts of <i>Jones' buttering the toast at midnight, during 4 minutes and 30 seconds.*

In what follows, I shall present how my view on events accounts for the entailment from (5) to (6). The logical form of each of these sentences is:

(5C) (\exists <*s*, *P*_{*init*}, *P*_{*fin*}, (*t*_{*init*}, *t*_{*fin*}), *p*>) (<*s*, *P*_{*init*}, *P*_{*fin*}, (*t*_{*init*}, *t*_{*fin*}), *p*> = the change of profile *p* occurring in system *s* by replacing the exemplification of property *P*_{*init*} with the exemplification of property *P*_{*fin*} during period (*t*_{*init*}, *t*_{*fin*}), and *s* = [Jones,

the toast], and P_{init} = the initial distribution of butter, and P_{fin} = the final distribution of butter, and *p* was done slowly)

and, respectively,

(6C) ($\exists < s, P_{init}, P_{fin}, (t_{init}, t_{fin}), p >$) (<*s*, *P_{init}, <i>P_{fin}*, (*t_{init}*, *t_{fin}*), *p*> = the change of profile *p* occurring in system *s* by replacing the exemplification of property *P_{init}* with the exemplification of property *P_{fin}* during period (*t_{init}*, *t_{fin}*), and *s* = [Jones, the toast], and *P_{init}* = the initial distribution of butter, and *P_{fin}* = the final distribution of butter).

Thus, given the truth-functional properties of the conjunction 'and', it is obvious that (5C) implies (6C), accounting for the entailment form (5) to (6).

Having considered the identity of events within the same world, we can now turn to the question of transworld identity of events: how are events to be individuated across possible worlds? Or, slightly different: given two events, e_1 and e_2 , occurring to the same substance, but happening each in a different world, what should it be enough for securing their (transworld) identity?

Note that, in the second question, it is already assumed a necessary constraint for having an identity between events: the sameness of their constituent substance. Given an event in the actual world, like Kim, I take the constitutive substance of the event to be essential to its identity across possible worlds. Borrowing an example from Kim, only Socrates could have died his death.

No doubt, that the substance to which an event is happening is crucial for the identity of that event is largely shared, but, if it is not carefully handled, it may lead us to serious confusions. Let us begin here by noting that a sentence like 'He could have died a different death' is perfectly intelligible, and, in some cases, it might be true.¹¹⁰ When it is true, it is crystal clear that the identity of the individual (i.e. the substance) is preserved across some possible worlds, but the identity of the event happening to that individual is not preserved. Instead, one could wrongly take a sentence like 'His death could have been much more quick than it was' as speaking about a case when there is a transworld identity both of the individual and of the event of his death. Such a wrong understanding of the above sentence leads some respectable philosophers like Lowe (2002) to claim that it is not crucial for the transworld identity of an event whether its temporal characteristic are preserved exactly:

It would be implausible to contend that the precise time of an event belongs to its 'individual essence' – that is, the one and the same event could not have occurred slightly earlier or later than it actually did occur. It may absurd to suppose that the Battle of Hastings – that very conflict – could have occurred during the reign of Queen Elizabeth I instead of in 1066, but it doesn't seem absurd to suppose that it could have occurred a day earlier or later than it actually did. (p. 181)

By contrary, in my view, sentence 'His death could have been much more quick than it was', if true, informs us about a particular case when sentence 'He could have died a different death' is true too.¹¹¹ A quicker death of some individual than his actual death is a different death. The reason is that, in my view, given any event, the profile of the transition between an initial state and a final one is crucial for its identity across the worlds.

¹¹⁰ When? Probably, *only when* other causes than those that actually caused him to die can be envisaged. But, attention, that individual would not have died a different death *only because* his death had had different causes.

¹¹¹ In other words, the former sentence implies the latter one.

I ground the latter remark on a particular view of events that I have made known above. I have argued that any event has a certain temporal metaphysical structure in the sense that any event is a change evolving in a certain manner, in time, from an initial state to a final state of a certain substance. Therefore, given any event, the profile of the transition between an initial state and a final one is crucial for its identity across the worlds. As I have mentioned, the profile of change is given by the states of transition and by the timing of their occurrences. By all means, the period of time of the event reflects in the timing of transition, and hence in the identity of the event.

My intention is only to emphasize a difference in perspective on the matter of the individuation of events, even though in both cases events are seen as changes. My analysis makes use of an approach that takes any event to be a temporal structure of states through which a substance goes during some period. Thus, any event is seen as a change evolving in a certain manner, in time, from an initial state to a final state of a certain substance. Accordingly, two events are identical if and only if they have the same constitutive substance, the same initial property, the same final property, the same period of time, and the same profile of property change.

Therefore, contrary to Lowe, given any event, I believe that the profile of the transition between an initial state and a final one is crucial for its identity across the worlds. As I have mentioned, the profile of change is given by the states of transition and by the times of their occurrence. By all means, the duration of the event reflects itself in the timing of transition, and hence in the identity of the event.

One reason for taking this view, against Lowe's, is Davidson's claim that if two events differ in their causes or in their effect, they are different events.¹¹² It seems to me that an earlier death must have different causes than the actual death; consequently, the two deaths are different.

Let us consider the following objection to my criterion of event identity: it is not a sloppy manner of talk to say that the party was a great event, perhaps the greatest event of the summer, but it would have been even better had John came as well (different substance); it would have been better had it started with a different dance, and ended with fireworks (different initial and final properties); and of course different things could also have happened during the party and it could have lasted longer (different profile of transition).

Even though the above phase is not a sloppy manner of talk, it does not force us to take "it would have been even better..." to mean the same as "the same event would have been even better had John came as well ...", because the original phrase might mean as well that there would have been a greater event, a different (!) event anyway, had John came as well...

¹¹² I endorse it without saying that it is also a criterion of event identity.

Chapter 5: The Constituents of Causation

This chapter will provide some useful notional support for the discussion to follow on overdetermination. A series of causal notions referring to some of the constituents of causation will be defined, explained and illustrated by proper examples.

I open the chapter by arguing that in any case of causation there always must be found at least an event which acts as a cause. As I shall show, causes are linked in causal chains that further generate the causal history of any effect. § 5.2 will present the phenomena – namely transitivity, convergence, and divergence – in virtue of which such a generation is possible. However, in giving causal explanations causal processes are more workable than causal histories. § 5.3 will define what it is for some causes to be linked in a causal process.

The notion of a causal process is elaborated further in § 5.4, where the relations holding between various causal processes will be discussed. Thus, as will be seen in § 5.4.1, two causal processes could be connected to each other either in a serial or in parallel manner. § 5.4.2 will investigate what it is for a causal process to be part of another one. Also, it will introduce the notions of parallel part and serial part. The elaboration is closed with an analysis of two causal processes, each bringing the same ending effect, share one of their parts.

The chapter ends up with the notion of causal role. It is intended to play a key part in my compatiblist attempt at dealing with CEP. The notion of causal role makes room

for discriminating among the different ways the same effect is or could be directly caused. § 5.5 will make explicit what gives the identity of a causal role, but also what gives the fragility of causal roles across the possible worlds.

5.1 Causal relata

In Chapter 3, I have suggested an account of causation that takes any causal relation to be part of some relation of nomic necessity in the sense that the causal relata are among the relata of the necessitation. Thus, in this view, if c_i is a cause of e_k , then there is some relation of nomic necessity holding between two sets, *C* and *E*, such that c_i is a member of *C* and e_k is a member of *E*.

Usually, each of the two sets has more than one member. Let us consider such a case where sets *C* and *E* are $\{c_1, ..., c_i, c_j, ...\}$ and $\{e_1, ..., e_k, ...\}$, respectively. I hold that each individual belonging to set *C* is a cause of each and all individuals belonging to set *E*.

Let us consider the following line of reasoning. That c_i is a cause of e_k boils down to the fact that they are involved in a relation of necessity holding in virtue of certain laws. Importantly, the laws underlying the necessitation of e_k are not triggered only because of c_i , but rather because of the whole bunch of individuals belonging to set *C*. It appears that, comparative to other members of set *C*, c_i has no special status in bringing about e_k . Therefore, it would be reasonable to take each of the members of set *C* to be a cause of e_k . A similar reasoning could be used to justify that each element of set *E* is an effect of each and all elements of set *C*. Besides the claim that the existence of any causal relation takes place in virtue of being part of some relation of nomic necessity, I have also held in Chapter 3 that any causal relation is decomposable in some chain of primitive causal relations. Moreover, in the case of a primitive causal relation, the underlying laws empowering the corresponding nomic necessity are fundamental.

Now, granting all these, a natural question is whether or not, whenever a case of nomic necessity takes place, there is also a case of causation. In other words, given that causation requires nomic necessitation, is the reciprocal true? Does any relation of nomic necessity involve some causal relation? If we allow a positive answer, the consequent view sneaking here is one according to which the gravitational attraction between Earth and Moon, the occurrence of an electrical field around a particle having a charge, a picture hanging on the wall would all count as cases of causation.

It seems to me that it is too much a dilution of the notion of causation to include gravitational attractions, say, as instances of causation. Even though one accepts that causation is grounded on a lawful network, the usual intuition is that the cases of causation require a feature that not all nomical bonds have.¹¹³ In my view, such a specific feature required by causation could be phrased like this: there is a case of causation only when *a change takes place in the world because of another change*.

¹¹³ The same point is also made by A. Chalmers (1999: 12-4), who argues that some laws like the laws of thermodynamics and the law of conservation of energy are not causal laws. Instead, Armstrong (1999) is tempted to consider all fundamental laws as being causal: "Not all laws are causal, or at least are not obviously so. So it seems that all I am entitled to say is that singular causation is instantiation of *a certain sort of* law. ... One very satisfactory solution of this problem from my point of view would be that all the *fundamental* laws are, or can be fairly represented as being causal laws." (p. 184)

Here, by 'change' it is meant a property change in some substances or objects, and such a change might amount both to getting new properties and to loosing previous properties.

Now, if I am right in saying that there is a case of causation only when a change takes place in the world because of another change, and granting that any change involves the occurrence of some event, the point of the above considerations is that *any causal phenomenon is initiated by at least one event*.¹¹⁴

Note that, the above considerations do not lead to the idea that the notion of cause is legitimately applicable only to events. Rather, the point is that, *in any case of causation, there always must be found at least an event which acts as a cause.* This is not to be confounded with the claim that in any causal "show" the real "actors" are only events.

Thus, the approach does not conflict with the commonsensical intuition that, in some cases of causation, besides events, there are also states that play a role in bringing about the effects. As I have mentioned already, sometimes we are inclined to say things like 'besides the event of striking the match, dryness and the presence of the

¹¹⁴ We should also take into account that we cannot say that it was an object that caused such and such. For, as Davidson rightly pointed out, an object has many properties, and not all of them are relevant for the respective causation; thereby, it would be improper to say that the object as a whole did the causation. Neither we can say that it was a number of properties (understood here as universals) that did the causation. Given that we are dealing in each case of causation with a particular phenomenon, a causal connection is a relation that should hold among individuals, not properties if they are understood as universals. But, on the other hand, we cannot say that objects or all of their properties do not have relevance for causal phenomena, because causations occur in virtue of the instantiation by some objects of certain properties. In the light of the above considerations, it seems plausible to say that an "actor" of any causal relation is made of an object and a property, the two being "brought togeder" by what is called instantiation or exemplification.

Instead of saying that, in virtue of having property C, object x caused object y to instantiate property E, I shall simply say that event c caused event e, where c is x's being C, and e is y's being E.

oxygen in the room were also causes of the event of lightning the match'. As far I can see there are two possibilities for accommodating causal events with causal states: either we simply call them all causes or we make use of the distinction between causes and causal conditions¹¹⁵, reserving the status of causes only for events.

I believe that, between causes and causal conditions, there is no metaphysically grounded distinction to be made. Rather, it is only an epistemic bias that would plea for it: for human cognizers, changes (to be read *causal events*) are *usually* more preeminent than static determinations (to be read *causal states*) in a causal picture. In a case of causation, the relevant events *trigger* the causation, whereas the relevant states *permit* or *house* it or *give a framework* to it. Again, these terms should be read in the cognitive-epistemological key.

Turning now to the question of what makes the relations of nomic necessity eligible for underlying causal relations, it seems that a plausible answer would be the following. Any relation of nomic necessity underlying causal phenomena occurs in virtue of certain laws triggered, among others, by the transition from a property exemplification to another.

¹¹⁵ It is important to be careful to avoid taking 'causal condition' for causally sufficient condition. On my usage, *c* is a causal condition of *e* if *c* is a state that matters causally for *e*, whereas *p* is a causally sufficient condition of *e* if '*p*' refers to a bunch of causes and causal conditions which are sufficient for *e*.

5.2 Causal histories

In his paper *Causal Explanation* (1986a), Lewis justly points out that the causation of any event is actually a meeting of a huge (maybe infinite) number of causal chains, where by 'causal chain' it is meant a *line* of ordered events such that there is some causal relation holding between any two successive events. Given any effect, there is some causal network into which it features as a node and where its causes make up its causal history. Here is what Lewis says:

Any particular event ... stands at the end of a long and complicated causal history. (...) An explanandum event has its causes. These act jointly. We have the icy road, the bald tire, the drunk driver, the blind corner, the approaching car, and more. Together, these cause the crash. Jointly they suffice to make the crash inevitable, or at least highly probable, or at least much more probable than it would otherwise have been. (...) But these are by no means all the causes of the crash. For one thing, each of these causes in turn has its causes; and those too are causes of the crash. So in turn are their causes, and so, perhaps, *ad infinitum*. The crash is the culmination of countless distinct, converging causal chains. (p. 214)

Let us consider an oversimplified case of causation, which brings about an event, *e*, whose causal history, i.e. a linkage of relevant causal relations, is illustrated in Figure

5.1.



It is important to note that causal histories appear to be generated in virtue of certain phenomena – namely transitivity, convergence, and divergence. Any causal "wave" is initiated with the occurrence of some event, and travels on some causal chain by passing intermediary events until it reaches a certain effect. For instance, in Fig. 3.1, one may see that a_1 is a cause of e in virtue of causal transitivity leading from a_1 to e through a_2 and a_3 on the way of the causal chain [... a_1 , a_2 , a_3 , e, ...]. Given an event and any of its causes, there is always at least a causal chain to be drawn in virtue of causal *transitivity* from the causal ancestor to the effect-event.

Also, a case of causal *convergence* occurs whenever there is a meeting of two or more causal chains.¹¹⁶ By causal convergence, events belonging to different causal chains (that is, events being causally unconnected) could jointly bring about a certain effect. In Fig. 5.1, the convergence of [...*a*₁, *a*₂, *a*₃, ...] and [...*b*₁, *b*₂, ...] makes possible that *a*₁ and *b*₁ jointly cause *e*.

Moreover, according to Lewis, the chains may diverge as well as converge, where the causal *divergence* is the inverse of causal convergence. By divergence, one or more causal chains may end up in more than one effect.¹¹⁷ In Fig. 5.1, the causal

¹¹⁶ By converging with another causal chain, a given causal chain does not cease to be a causal chain. The meeting between two causal chains marks only the beginning of a part-chain shared by both convergent chains. In Fig. 3.1, causal chains [... a_1 , a_2 , a_3 , e, c_1 , ...] and [... b_1 , b_2 , e, c_1 , ...] share the part-chain [e, c_1].

¹¹⁷ Considering only one causal chain, a divergence would amount to splitting the causal chain in two or more chains at the point where more than one effects are brought about by the causes belonging to the given chain. It follows that the initial chain is nothing but a part of chain shared by the resulting divergent causal chains. In Fig. 3.2, causal chain $[e, c_1]$ splits in $[c_2, c_3, ...]$ and $[d_1, d_2, ...]$. Obviously, $[e, c_1]$ is a part-chain shared by $[...e, c_1, c_2, c_3, ...]$ and $[...e, c_1, d_2, ...]$.

Given that sooner or latter any causal chain diverges or, conversely, any causal chain is result of two or more converging causal chains, one may worry about the identity of any such causal chain. In order to avoid any such complication, by 'causal chain' I shall mean a *closed* causal chain, i.e. a chain that has a *prime* cause and an *ultimate* effect. In notes ... I have tacitly used '[x, ..., y]' for denoting a closed causal chain whose prime cause is x and ultimate effect y.

Considering any two causally related events, it could be proved that between them there is only one closed causal chain whose ends (extremities) are exactly the events under consideration. In Fig. 3.2, a_2 and d_1 are causally connected through the causal chain $[a_2, a_3, e, c_1, d_1]$.

Of course, usually, on the one hand, a closed causal chain is a shared part of other causal chains, and, on the other hand, it shares with other causal chains parts. For instance, in Fig. 3.2, $[a_2, a_3, e, c_1, d_1]$ is shared by $[a_1, a_2, a_3, e, c_1, d_1]$ and $[a_2, a_3, e, c_1, d_1, d_2]$; and it shares $[e, c_1]$ with $[b_1, b_2, e, c_1, c_2]$.

divergence occurring after c_1 makes possible that an event like e is a cause of both c_2 and d_1 .

Two or more causal chains could have a "spot of convergence&divergence" in the sense they point each to a number of events they jointly cause; the histories of those effect-events share the same causal chains. For instance, in Fig. 5.1, [... a_1 , ..., e, c_1 , ...] and [... b_1 , ..., e, c_1 , ...] realize a causal convergence which divergently brings about c_2 and d_1 .

In Figure 5.2, it is illustrated a case of causation where two causal chains diverge and converge in the same spot. Moreover, the chains converge more than once: aside the convergence occurring after a_3 and b_2 , there is also a convergence after a_2 and b_1 .



Fig. 5.2

Relative to a given effect, let us call *unrelated causes* any two events that are causes of the effect, but do not belong to the same causal chain.¹¹⁸ Further, relative to an event (or a set of events), let us call *unrelated effects* any two events that are caused by the given event (or by the given set of events) such that there is no causal chain passing each of them. In Fig. 5.2, relative to c_1 , events a_1 and b_1 are unrelated causes, and, relative to a_3 , events c_1 and d_1 are unrelated effects.

5.3 Causal processes

As it has been defined in the previous section, the causal history of an effect is the totality of causal chains converging in the respective effect. However, for practical reasons, in giving the causal explanation of the occurrence of some event, it is preferable to make reference to the most recent and preeminent part of the causal history of the event to be explained. Thus, it is the notion of a causal process that comes into play when a reasonable account is to be given for a certain phenomenon. The advantage of working with causal processes is clearly their finitude and metaphysical economy

By definition, the identity of a causal process is given by a triplet consisting in:

- i) the causes that started it call them the originating causes;
- ii) the effects with which the process ends up call them the ending effects;
- iii) the causal network that joins the originating causes to the ending effects call it the *linking network*.

¹¹⁸ Also, relative to an effect, let us call *unrelated sets of causes* any two sets of causes of the effect such that any two events, each of a different set, are unrelated causes. For instance, relative to effect c_1 , $\{a_1, a_2, a_3\}$ and $\{b_1, b_2\}$ are unrelated sets of causes.

For example, I shall use '<X; Y; Z>' for denoting a causal process whose originating causes are X, linking network Y, and ending effects Z. (see Figure 5.3).



One constraint is that, considering any causal process, for any couple (*i*, *j*) formed by an originating cause ($i \in X$) and an ending effect ($j \in Z$), there is a causal chain connecting *i* with *j*.

Note that the linking network of a causal process is not made up of all causal chains going through the originating causes and leading to the ending effects. It is possible that two different fascicles of causal chains leave from the same originating causes and arrive in the same ending effects, and each fascicle stands for a different causal process. For example, in Figure 5.4, there are two distinct causal processes, <X, Y,

Z> and < *X*, *Y**, *Z*>, having the same set of starting causes and the same set of ending effects, but different linking networks.¹¹⁹



Fig. 5.4

Therefore, as another constraint, in order to avoid the existence of vague borders in cases like the one mentioned above, we need to impose that, for any causal process, its linking network is closed to any causal input except the originating causes. In speaking about a causal process, I shall set under the umbrella of term 'node' any of its originating causes and any of the constituents of its linking network. Thus, considering any given causal process, any node of its linking network¹²⁰ has its direct causes exclusively among the nodes of the causal process; conversely, any direct cause of a node of the linking network is itself a node of the causal process. By

¹¹⁹ Of course, the containing fascicle – i.e. the one made of the two parallel fascicle – stands for the containing causal process, $\langle X, Y + Y^*, Z \rangle$ – i.e. the one having as parts the two causal processes corresponding each to one of the two thinner fascicles.

¹²⁰ The network does not include its heads (i.e. originates causes and ending effects).

definition, *c* is a direct cause of *e* iff there is a primitive causal relation standing between *c* and e^{121} That is, iff *c* causes *e* without mediating causes.

Note that the closure of a causal process is only relative to inputs, not to outputs. Relative to a causal process, any direct cause of a node of the linking network is itself a node in the same process; that is to say that there is no convergence within a causal process between its chains and chains coming from the outside of the originating causes. Instead, it may be that a node directly causes something located in the outside of the causal process.

For instance, in Figure 5.5, the triplet $\langle c_1; c_2; e \rangle$ is not a causal process. This is because c_3 is a direct cause of c_2 and hence, according to what is meant by causal process, c_3 should be taken to be a node of the causal process, which it is not. Since c_3 is not caused by any other element of the process, it is of the same kind as c_1 , namely, an originating cause; hence a triplet that really is a causal process is $\langle c_1, c_3; e \rangle$.



Fig. 5.5





¹²¹ The notion of primitive causal relation has been discussed at large in Chapter 3.

However, regarding causal processes, I believe that there is no reason for imposing a constraint like the following: there is no causal relation holding between its originating causes. In Figure 5.6, it is pictured a case identical with that corresponding to Figure 5.5, except that in the latter case c_3 is caused by c_1 . However, $\langle c_1, c_3, c_2; e \rangle$ is still a causal process alright. True, it is identical with $\langle c_1; c_3, c_2; e \rangle$. But, in choosing one or another symbolization, it is the causal sufficiency of the originating causes for the ending effect that matters most. Some relevant examples will be available in last chapter of the thesis.

5.4 Relations holding between two causal processes

5.4.1 Connecting two causal processes

As far as I can see, two causal processes may be connected in two possible ways. Two causal processes, Φ and Π , are *serially* connected if and only if Φ 's ending effects are identical to Π 's originating causes. See Figure 5.7. Instead, Φ and Π are connected in a *parallel manner* if and only if they are identical in their ending effects. See Figure 5.8.





For instance, in Figure 5.9, process $\langle c_1; -; c_2, c_3 \rangle$ is serially connected to process $\langle c_2, c_3; -; e \rangle$. There is also a serial connection between $\langle c_1, c_5; -; c_2, c_3, c_4 \rangle$ and $\langle c_2, c_3, c_4; -; e \rangle$.

In the same figure, process $\langle c_2; -; e \rangle$ is parallel to process $\langle c_3; -; e \rangle$. Note that the two processes have no originating causes in common. But this is not a requirement for a parallel connection. Process $\langle c_2, c_3; -; e \rangle$ is connected in a parallel manner to process $\langle c_3, c_4; -; e \rangle$. The former connection instantiates a separation case, whereas the latter an overlapping case.



It is important to note that, except the case when either one's originating causes are nodes of the other or vice versa, two processes connected in a parallel manner do not meet each other until they both reach their ending effects. In other words, if a parallel connection holds between two causal processes, Φ and Π , none of one's originating causes being a node of the other, their linking networks do not share anything. Why? Let us assume by *reductio ad absurdum* that there is *c*, a node shared by the linking networks of the two processes. See Figure 5.10. It follows that there is *c*', a constituent of Φ 's linking network, such that *c*' is a direct cause of *c*. But, given that *c* also belongs to Π 's linking network and *c*' is a direct cause of *c*, it follows, by definition, that *c*' features into in the constitution of Π 's linking network. Repeating the same rationale of backtracking, one must find eventually an originating cause, *c**, of Φ , for instance, such that *c** belongs to Π . This conclusion is against the hypothesis that Φ and Π do not share any of their originating causes. Therefore, the presupposition is false.



As we shall see in § 5.4.3, if some of the originating causes of Π are nodes of Φ , those causes originate a process whose linking network has nothing in common with the causal chains linking the rest of originating causes of Π with the shared ending effects. See Figure 5.11.



Fig. 5.11

5.4.2 Being part of a causal process

It should be uncontroversial that, by connecting two causal processes, Φ and Π , it is obtained a new causal process, *P*, such that Φ and Π are parts of *P*. Also, it is safe enough to claim that Π is a part of Φ if and only if Π is engaged with other processes in a series of connections, ending up with the obtaining of Φ .

Now, let us assume that Π and Φ are two processes having the same ending effect eand Π is a part of Φ . I propose the following series of connections which starts with Π and leads to Φ . Given that Π is a part of Φ , Π 's originating causes are included in Φ 's elements. Therefore, there is a process, Φ^* , having e as ending effect, such that Φ^* is obtained by cutting Φ through a section going through Π 's originating causes. See Figure 5.12. Obviously, Φ^* and Π are connected in a parallel manner. But, given that Π 's originating causes are included among those of Φ^* , it follows that Φ^* is the whole obtained from Π connected in a parallel manner with $[\Phi^* - \Pi]$, the process inside Φ^* , whose originating causes are the originating causes of Φ^* without the originating causes of Π . Note that, according to the argument presented at the end of the previous section, the only thing that Π and $[\Phi^* - \Pi]$ have in common is *e*, their ending effect.

It is also important to note that, even though ϕ is composed of two pieces, ϕ^* and "something" containing some of the ancestors of ϕ^* , that something is not necessarily a process as such. For it may not meet the constraint (that any causal process is bound to meet) that there should be at least a causal chain between any originating cause and any ending effect.



Let us take the example illustrated in Figure 5.13. Let us say that Φ is $\langle c_1, c_5; c_2, c_3, c_4; e \rangle$ and, Π is $\langle c_4; -; e \rangle$. In order to get the value of Φ^* , one should split Φ though a section meeting c_4 . Let us say that the section also meets c_2 and c_3 . Then, Φ^* is $\langle c_2, c_3, c_4; -; e \rangle$, that is a process obtained by connecting in parallel Π , i.e. $\langle c_4; -; e \rangle$, with

< c_2 , c_3 ; -; e>. Further, given that < c_1 , c_5 ; -; c_2 , c_3 , c_4 > is not a causal process (for example, c_5 is not a cause of c_3), ϕ cannot be obtained from the serial connection between ϕ^* and some leading process, but rather by adding some causal chains linking c_1 and c_5 to some of the originating causes of ϕ^* .



Let us fix the following terminology: relative to a given process, by 'parallel part', it is meant a part related to its whole in the same way Π is related to Φ^* ; by 'serial part', it should be understood a part hooked to its whole in the manner Φ^* is hooked to Φ .

5.4.3 Sharing a causal process

Let us turn now to the situation when Φ and Π share one of their parts. For some reasons, the interesting case for our study is that where the two processes have the same one ending effect, *e*. Let us focus the on the largest part shared by Φ and Π . Call it [$\Pi \cap \Phi$].

It can be argued that $[\Pi \cap \Phi]$ has its originating causes among those of Φ or among those of Π . Let us assume by *reductio ad absurdum* that there is c, one of the originating causes of $[\Pi \cap \Phi]$, such that c belongs neither to Φ 's originating causes, nor to Π 's originating causes. Since $[\Pi \cap \Phi]$ is a part of Φ , it follows that c belongs to Φ 's linking network. The same works for Π : c belongs to Π 's linking network. Thus, there is c^* such that c^* belongs to Π , c^* directly causes c, and c^* is not one of the originating causes of $[\Pi \cap \Phi]$. But, one of the constraints imposed on any causal process is that any node of its linking network has its direct causes exclusively among the nodes of the causal process. It follows that c^* belongs also to Φ . But this is to say that c^* is a node of $[\Pi \cap \Phi]$ and also a cause of some of the originating causes of $[\Pi \cap \Phi]$. It follows further that c^* is an originating causes of $[\Pi \cap \Phi]$. Such a conclusion is against the initial assumption. Thus, the initial assumption cannot be but false, and the originating causes of $[\Pi \cap \Phi]$ features among those of Φ or among those of Π .

Next, it can be easily argued that, given that Φ and Π have a single shared ending effect, e, $[\Pi \cap \Phi]$ exists only if the ending effect of $[\Pi \cap \Phi]$ is e. Let us assume by *reductio ad absurdum* that there is e^* , an ending effect of $[\Pi \cap \Phi]$, such that $e^* \neq e$. Given that e^* belongs to Φ , there is at least a causal chain going through Φ and linking e^* with e. The same works for Π : the same linking causal chain goes through Π . Consequently, there is at least a causal chain shared by Φ and Π , that ends up with e. But, then, the same causal chain goes through $[\Pi \cap \Phi]$, and, hence, contrary to our assumption, e^* is not an ending effect of $[\Pi \cap \Phi]$. Therefore, the initial assumption cannot be but false, and $[\Pi \cap \Phi]$ has e as its single ending effect.

If the above argumentation is correct, one thing that it shows is that, in the case of two causal processes, Π and Φ , having e as their single shared ending effect, if there is such a thing as $[\Pi \cap \Phi]$, $[\Pi \cap \Phi]$ is a causal process whose single ending effect is e, such that $[\Pi \cap \Phi]$ is either a parallel part of Φ (in the case when the originating causes of $[\Pi \cap \Phi]$ are exclusively among those of Φ), or a parallel part of Π (in the case when the originating causes of $[\Pi \cap \Phi]$ are exclusively and Π are exclusively among those of Π and Π (in the case when the originating causes of $[\Pi \cap \Phi]$ are exactly the originating causes shared by Φ and Π and Π and Π and Π are exclusively the originating causes of $[\Pi \cap \Phi]$ are exclusively the originating causes of $[\Pi \cap \Phi]$ and Π are exclusively the originating causes of Π and Π are exclusively the originating causes of $[\Pi \cap \Phi]$ are exclusively the originating causes of $[\Pi \cap \Phi]$ are exclusively the originating causes of $[\Pi \cap \Phi]$ are exclusively the originating causes of $[\Pi \cap \Phi]$ and Π are exclusively Φ and Π and Π are exclusively the originating causes of $[\Pi \cap \Phi]$ are exclusively Φ and Π and Π are exclusively Φ and Π



Let us assume that $[\Pi \cap \Phi]$ is a parallel part of Π and a simple part of Φ . See Figure 5.14b. Then, $[\Pi - \Phi]$ and $[\Pi \cap \Phi]$ have no event in common, except their single ending effect, *e*. Given that $[\Pi \cap \Phi]$ and Φ have the same ending effect, *e*, there is Φ^* , a serial part of Φ , such that $[\Pi \cap \Phi]$ is a parallel part of Φ^* . Of course, then, $[\Pi \cap \Phi]$

 Φ] and $[\Phi^* - \Pi]$ have no event in common, except their single ending effect, *e*. It follows that there are three process, $[\Pi - \Phi]$, $[\Pi \cap \Phi]$ and $[\Phi^* - \Pi]$, such that they are connected in a parallel manner and have no event in common, except *e*. Therefore, Π and Φ^* are connected in a parallel manner and they share the same process as that shared by Π and Φ , i.e. $[\Pi \cap \Phi]$ is exactly $[\Pi \cap \Phi^*]$.

In Figure 5.15, if the considered values of Φ and Π are $\langle c_1, c_3; c_4, c_5; e \rangle$ and, respectively, $\langle c_2, c_3; c_4; e \rangle$, their maximal shared part, $[\Pi \cap \Phi]$, is $\langle c_3; c_4; e \rangle$; here, $[\Pi \cap \Phi]$ is a parallel part of Φ and also of Π .



Fig. 5.15

5.5 The notion of causal role

According to the nomic account of causality for which I have pleaded in Chapter 3, any causal relation is part of some relation of nomic necessity. Thus, if there is a causal relation holding from c_i to e_k , the causal relation is part of some relation of nomic necessity holding from the events of a set, including c_i , to the events of another set, including e_k . Let is be $\{c_1, ..., c_i, c_j, ...\}$ and $\{e_1, ..., e_k, ...\}$ the two sets. It follows that, if c_i is a cause of e_k , the events belonging to $\{c_1, ..., c_i, ...\}$ necessitates (i.e. are sufficient for) the occurrence of each event belonging to $\{e_1, ..., e_k, ...\}$ in virtue of some set of laws, *L*.

Further, I have also defined a causal relation standing between any two given events as being primitive iff it is a direct or intimate connection holding between the two events in the sense that there is no need of another event to mediate the relation.¹²² Next, an argument has been deployed for the conclusion that a causal relation is primitive iff it coincides with a causal dependence. The crucial role in reaching the conclusion is that, according to the nomic account of causation, any causal dependency coincides with a part of a fundamental relation of nomic necessity, that is, a relation of nomic necessity empowered by some set of fundamental laws. The direct consequence is that a causal relation is primitive only if it is part of some fundamental relation of nomic necessity.

¹²² It is important to recall that there might be two distinct causal relations holding between the same two events such that one is primitive and the other is made of a number of enchained primitive causal relations enchained each other, two by two.
If the above considerations are right, it follows that, if there is a primitive causal relation holding from c_i to e_k (that is, c_i is a direct cause of e_k) the events belonging to $\{c_1, ..., c_i, ...\}$ necessitates (i.e. are sufficient for) the occurrence of each event belonging to $\{e_1, ..., e_k, ...\}$ in virtue of some set of fundamental laws, *L*. Relative to an effect, a direct cause is to be distinguished from any mediated cause: the latter is not linked with the given effect by a primitive causal relation, but rather by a chain of primitive causal relations.

Thus, if the set of laws grounding the above nomic necessitation were different, c_i would be a direct cause of e_k of a different sort. Also, supposing that c_i and c_j are each a direct cause of e_k , c_i causes e_k in a different way than c_j does. In order to capture what particularity c_i comparatively to some other, whether possible or actual, direct causes of e_i , I introduce the notion of *causal role*. The notion of causal role is intended to make room for discrimination among the different ways the same effect is or could be directly caused. Thus, ' c_i is a direct cause of e_k ' is taken to be synonym to ' c_i has a certain causal role with respect to e_i ', and ' $R(c_i, e_k)$ ' to denote the respective causal role.

Usually, in scientific theories, a law is expressed by a mathematical equation whose variables are determinable properties and whose solutions are determinates of those properties. Consider, for example, the equation that expresses Coulomb's law: $E = q/4\pi\epsilon_0 d^2$ relating the electric field *E* to the non-distributed charge *q* and the distance *d* (all three are determinable). Thus, the instantiation of a charge *q* by a particle nomically necessitates the instantiation of a field whose electric intensity *E* at distance *d* from the particle can be calculated by the above equation. If there are two

particles charged with q_1 and, respectively, q_2 , we can calculate the intensity of their resultant electric field in any point, according to the contribution of each charged particle. Such a contribution can be individuated by the "place" the value of the charge has in the numerical structure reflecting the application of Coulomb's law to this particular two charges case.

Unfortunately, not all causal laws have a mathematical expression and not all nomic necessitation can be represented by the numerical structure of an equation (or set of equations). However, the suggestion to be highlighted is the following: the causal role c_i has with respect to e_k , $R(c_i, e_k)$, is the particular "place" c_i has in the structure of the necessitation holding from $c_1, \ldots, c_i, c_j, \ldots$ to e_1, \ldots, e_k, \ldots in virtue of some set, L, of fundamental laws. The identity of such a place is given both by the property whose tokening occasions the occurrence of c_i^{123} and by those laws belonging to L, to whose activation the tokening of the property contributes.

Any case of causation is a case of acquiring certain properties (by effect-relata) whose tokens are nomically necessitated by the tokens of other properties acquired by cause-relata. Thus, to say that c had a certain causal role relative to e, R(c, e), is to say that c and e are the instantiations of properties C, and E, respectively, such that, given some fundamental laws, L, a relation of necessity holds between the tokens of C and of other properties, on the one side, and the tokens of E and of other properties, on the one side, and the tokens of E and of other properties, on the one side, and the tokens of E and of other properties, on the one side, and the tokens of E and of other properties.

¹²³ For simplicity, if c_i is an event by which is instantiated property C_i (i.e. the final property featuring into the constitution of c_i), it is profitable to consider only the case where the property transition is done from a state of no instantiation of any property of the same determinable as C_i . Accordingly, in what follows, in an abbreviated way, it will be said that c_i is simply the instantiation of C_i , by "instantiation" being meant the process, not only the result.

causal role (that is, a role in bringing about *e*) that gains its identity from both the corresponding property acquired by the cause and the corresponding fundamental laws whose activation the triggering property is (co-)responsible.

Let us retake the match example from § 5.1. Supposedly, in our world, there is a chemical law relating somehow the tokens of the properties *having a certain temperature, having a certain chemical composition*, and *being rich in oxygen* with the tokens of other properties *combusting*, and *flashing*. In virtue of the chemical law triggered by the tokening of the proper temperature, the chemical composition, and the concentration of oxygen, it is the occurrence of some combustion and flash that are necessitated. Accordingly, the state of the air's being oxygenated is one of the direct causes that brings about the event of lightning the match. Another cause is the raising of the temperature in the inflammable substance from the tip of the match. Each of these causes has a specific causal role in catching fire.

We may speak not only about different ways by which a given effect is directly caused in the same world, but also about differences across distinct possible worlds. The most interesting case for the present discussion is when the ways of causation to be compared are undertaken by the same cause, but in two different possible worlds. In my view, such a transworld discussion is possible (allowed) because causation is grounded on lawhood, and, furthermore, the notion of law of nature is a modal one. Therefore, the notion of causal role is to be treated as a modal notion as well.

Let us see what determine the endurance or, by contrary, the fragility of causal roles across the possible worlds. Simply put, given an effect *e*, what assure the transworld

preservation of a causal role in bringing about *e* is the same as what give the identity of the respective causal role: the tokening of some property and the activation (or the participation to the activation) of some laws. Suppose that there is a possible world, *i*, where, like in the actual world, *c* is a direct cause of *e*. Given the supposition of transworld identity of *c*, it is trivial that property *C* is instantiated in both worlds. It follows that $R_{@}(c, e) = R_{i}(c, e)$ amounts to the activation of the same fundamental laws in the both worlds on the occasion when *c* causes *e*.

But, the transworld identity of causal roles is not restrained to the cases of the transworld identity of causes. It can be the case that two different direct causes of an effect, one actual, *c*, the other possible, *c**, can have the same causal role with respect to effect *e*. If the occurrence of the each cause involves the tokening of the same property, *C*, such that *C* contributes to the triggering of the same fundamental laws involved in bringing about *e*, then $R_i(c^*, e) = R_{\odot}(c, e)$.

In the match example, the causal role the state of the air's being oxygenated has is given by the place the token of *being rich in oxygen* occupies in the relation of nomic necessitation created by the activation of a certain chemical law of combustion. The same causal role is played by a different state occurring in a possible world where *being rich in oxygen* is exemplified by other molecules of air, and the same chemical law of combustion holds. More spectacularly, the same causal role in lightning the match has the state of some molecules of liquid that instantiates the property of *being rich in oxygen* in a possible world. Probably, the raising of the temperature in the inflammable substance from the tip of the match would not be acquired by friction, but rather by another manner – say, by contact with a heating device.

Now, let us clarify what is meant by the saying that, with respect to *e*, *c* could have had a different causal role. Let us consider a possible world, *j*, where both *c* and *e* occurred, and ask in what conditions the causal role *c* had in bringing about *e* within *j* is different from the causal role had by *c* within the actual world.¹²⁴ Briefly, when is the case that $R_j(c, e) \neq R_{@}(c, e)$? What is a world *j* where $R_j(c, e) \neq R_{@}(c, e)$? To my mind, relying upon the above considerations, an eligible answer should be the following: $R_j(c, e) \neq R_{@}(c, e)$ if and only if the structure of nomic necessitation leading to the occurrence of *e* within *j* is different from the actual one. That is, iff the causal laws underlying the actual causal relation from *c* to *e* are not activated within *j*. This could happen when either there are no such laws within *j* or, in spite of there being the same relevant laws in world *j*, the elements needed for their activation did not all occur.

Let us consider a world, *j*, such that, in *j*, there are no laws identical with those laws actually supporting the causal relation from *c* to *e*. Even though *c* does not preserve its actual role in bringing about *e*, *c* still has some causal role (i.e. $R_j(c, e) \neq 0$) if and only if there are some laws triggered by the occurrence of *c* in world *j* such that a primitive causal relation holds from *c* to *e*.

Next, let us consider a world, *h*, such that, even though the laws actually supporting the causal relation from *c* to *e* also govern in *h*, the elements needed for their activation did not all occur. In world *h*, *c* has no role in bringing about *e* (i.e. $R_h(c, e) = 0$) if and only if the configuration of the instantiated properties that nomically

¹²⁴ I also include the case where c had not causal role relative to e.

necessitates the occurrence of e does not contain the token of C. That is, if and only if, in world h, it is not merely that the structure of nomic necessitation responsible for the occurrence of e is different from the actual one, but also the token of C had no place in the respective structure. The occurrence of c is entirely irrelevant for the occurrence of e.

Let us take the match example one more time. A good illustration for the former case would be a possible world where, like in the actual world, the air's being rich in oxygen is a direct cause of the match's lightening, but the chemical law backing up the relevant relation of necessity is different. Besides the tokens related by the actual law, the possible law would require for its triggering the presence of light. In such a possible world, you cannot lighten the match just like in ours if you are in a dark room. Then, all the causes of the match's lightning, including the air's being rich in oxygen, change their causal role. The requirement of a "new" constituent, namely the presence of light, for assuring a nomically sufficient antecedent shows that the structure of nomic necessitation is different within *j*; other law governs the combustion of the match in the world *j*.

The law would also be different (changed) if an actual element looses its place within the possible nomic structure (that is, the nomic structure triggered in the possible world). For instance, for the lighting of the match, instead the rising of its temperature, the chemical laws holding in *j* requires its being exposes to a sound of a certain pitch. Again, such a difference in the structure of nomic necessitation makes the air's being oxygenated to cause the match's lightning in a different way – in other

words, to be a cause of the match's lightning in other sense than it is seen to be in our world.

Interestingly, the above result is pretty coherent with our common intuitions. If we are told that *c* causes *e* in a very strange scenario (e.g., a fairytale-like scenario), we will deem the respective case of causation to be different from what happens when *c* causes *e* in the actual world. Plausibly, we shall say that even if there is causation in a fairy tale abounding in witches' deeds, the causal events have different roles than they have in the actual world. Why that? I believe that what keeps us from accrediting them with the same causal roles is the lawhood underlying causation in a fairy tale world; such a lawhood is very different from that regulating cases of causation in the actual world.

In order to illustrate the case when, relative to some given effect, in a possible world, an event that is an actual direct cause has no causal role, let us consider a situation where, even though the law of combustion is unaltered, the inflammability of the substance from the tip of the match looses its causal role in lighting the match. You may concede that such a situation is possible if you concede the possibility of something's preventing the raising of the temperature in the inflammable substance. In the respective situation, given the incompleteness of the nomic antecedent, the law of combustion acting in the actual world would not be activated. However, this would not preclude the lightning of the match, because other laws of lightning would be triggered by the occurrence of other events.

I close this section by extending the notion of causal role such that to apply it to causal processes. Such an extension will play a crucial part in the following chapter, where I shall analyze the notion of causal overdetermination. Naturally, given a causal process having a single ending effect, the identity of the causal role such a causal process has is given by the identities of all causal roles played by the nodes of the causal process relative to their unmediated effects within the process. If the causal process has more than one ending effect, its causal role relative to one of these effects, *e*, is made up of all causal roles its nodes have in respect to their direct effects on the way of those causal chains leading to *e*. As a matter of notation, R(P,e) stands for the causal role the process *P* has in bringing about its ending effect *e*.

Let us consider the following causal process: the match's striking causes the match's rising its temperature, and the latter, "helped" by the state of air and the chemical composition of the substance from the tip of the match, causes the match's lighting. The causal role the process has in bringing about the lighting is to be individuated by the elements that individuate the causal role of the match's striking relative to its rising its temperature and the causal role played by each of the direct causes of the lighting.

Addendum 5.1 The metaphor of causal power

When I say that an individual has a certain causal power, I mean it in a metaphorical sense. As I shall try to show in what follows, when it comes to the issue of which – either the object or the events – are endowed with causal powers, the attribution of causal powers, even though only metaphorically, is to be done to objects. To say that an object has the actual causal power to do such-and-such is simply to say that it instantiates certain properties whose tokens nomically necessitate other property instantiations. I emphasize the fact that the usage of 'causal power' is only metaphorical because there is *no general agreement on how we should understand necessitations*. For something to have a power means *stricto sensu* to be able to make such-and-such to happen in virtue of its nature. But, according to the Humean conception, there is nothing in worldly entities which is metaphysically responsible for the occurrence of other entities; the nature of entities is not responsible for their spatial and/or temporal arrangement.

In the common usage, properties are what endow something with causal powers. But not everybody agree on which entities – whether the objects or the events – are empowered for causal jobs? On the one hand, one is tempted to favor the events because they are causes and not the objects featuring in events' constitution. On the other hand, one could reply that, given that the causal properties are exemplified by the objects, the objects are endowed with the causal powers.

My tentative account is the following. The change by which some causal properties are exemplified by an object is just the endowing of the object with causal powers,

CEU eTD Collection

but it is also an event; thus, the causal event is basically the endowing of the object with a certain causal powers.

But, what is for an object to be endowed by a property with a causal power? It is simply the actualization in the object of a power which was already there, but in a dispositional way (i.e. *in potentia*). In saying that the object has a causal power as a disposition, it is meant that the object *could instantiate a certain causal property*. That is, I equate the object's disposition to instantiate a property to the object's disposition to actualize a causal power by the instantiation of the property. So, the suggestion is that a cause of an effect is the actualization in an object of its causal power to bring about the effect.

It seems to me that the following chain of equivalence comes up analytically: the cause of e is what brings about e, which is the actualization of the power of bringing about e, which is, in turn, an event. Do not say: 'Still, if the actualization of the power of bringing about e is in the object, what brings about e is the object.' For such a claim would leave you exposed to the question 'why the object did not bring about e before?'. If you answer: 'Because the respective causal power of the object was not yet actualized', you already admit that it is not the object, but rather the actualization of its power – i.e. an event – that caused e.

Simply put, an event is the change by which an object acquires a property. Thus, any causal event is the actualization of some causal power the final property occasioned by the event provides to the object. Maybe an example would shed some light on this matter. Let us consider a bottle made of glass, which is full with water. Suppose

further that the water in the bottle is in a liquid state, and the bottle is sealed tight. If the temperature goes down, the water in the recipient freezes; by freezing, its volume raises and causes the breaking of the bottle. The event bringing about the breaking of the bottle is the freezing of the water, which is a change in properties consisting *inter alia* in changing in the volume.

The causal property that endows the water in the bottle with the causal power of breaking the bottle is *having a volume bigger then the interior of the unbroken bottle*. (Actually, the causal property is that of *having a certain water mass with a certain structure corresponding to a temperature below 0* $^{\circ}$ C.) So, the freezing of the water in the bottle is the actualization in the water of the power of breaking the bottle. Initially, the water had the power in a dispositional way.

Chapter 6: Overdetermination

This chapter is an analysis of the concept of causal overdetermination. § 6.1 will try to argue that if there is any way of answering the exlusionist's attack from CEP against interactionism, the first step to be taken is to look for a necessary condition on overdetermination. It will be pleaded for the idea that such a condition cannot be found by appealing either to intuition, or to broadly agreement. The main reason is that 'overdetermination' is a derivative notion, and hence its mastering firstly requires a conceptual analysis. Thus, § 6.2 will open a research whose outcome is the conclusion that the understanding of 'overdetermination' should be grounded on the grasping of other two notions: 'causal process sufficient for an effect' and 'excess in bringing about an effect'. The clarification of these notions is the task for § 6.3. Then, in the next section, using, on the one hand, the results obtained in Chapter 5, where the notions of causal process and causal role has been elaborated and, on the other hand, the analysis undertaken in the previous sections of Chapter 6, I shall propose a set of necessary conditions on overdetermination. § 6.5 will discuss two more useful notions, 'unnecessary cause' and 'overdetermining cause', in order to argue against the idea that the former does not necessarily involve the latter.

§ 6.6 will review a number of causal scenarios that are recurrently cited as cases of overdetermination in today's literature. In § 6.7, it will be made explicit some general kind of structure usually employed when one tries to provide an example of overdetermination. Such a structure will be used in § 6.8 to develop an argument that shows that some of classical examples are highly debatable, many of them

illustrating only cases of pseudo-overdetermination or partial overdetermination. A crucial role in the argumentation will be played by theory of events that is sketched in Chapter 4. § 6.9 will propose an argument intended to arise a serious challenge to the very idea of overdetermination. Finally, the last section of the present chapter is a study on what would ground our reluctance to admit widespread and systematic overdetermination.

6.1 Questionable intuitions

In § 2.6, I have called attention on two possibilities of blocking the exclusionist argumentative road that aims at showing that the CEP assertions, if they all were true, would amount to an exasperating systematic overdetermination. One way of refuting the exclusionist argument is to come with a counterexample that proves that the sufficient condition on overdetermination, (SC), used in the argument is false. Another way is to find a necessary condition on overdetermination, (NC), such that (NC) is not met when there are mental-physical relations of a certain kind. Given that the exclusionist argument is intended to show that, no matter what, the truth of the CEP assertions entails a case of overdetermination, the possibility such mental-physical relations makes the argument inconclusive.

Bennett (2003) emphasizes that some exclusionists relies on the claim that a sufficient condition on overdetermination, (SC), would be met when the effect has more than one sufficient cause. The counterexample she gives to refute the claim is that of two sufficient causes featuring in the same causal chain. But, my question is, how does Bennett know that, in such a case, the effect is not overdetermined?

Basically, according to Bennett, the reliable sources of this knowledge are *intuition* and *broad agreement*.

Many people have pointed out that events on a causal chain do not overdetermine their effects, even though both are causally sufficient for them" (p. 474, my italics).

And:

 c_1 is causally sufficient for c_2 , which is in turn causally sufficient for e - cases in which c_1 and c_2 are parts of the same causal chain. Our *intuitions* are quite univocal that these are not cases of overdetermination ... (p. 478, my italics)

Here, to my mind, the intuition works well and is correct, and the counterexample is faultless, serving its purpose to a certain extent. Unfortunately, it is perfectly possible that the exclusionist uses another alleged (SC), one having a very complicated content. Of course, the more complicate the proposed sufficient condition is, the more sophisticate the comeback counterexample is needed, and the less reliable the use of *intuition* and *broad agreement* is in deciding whether the effect is or not overdetermined.

The simple lesson is that if there is any way of answering the exlusionist's attack, it is not one relied exclusively on counterexamples. The better way would be to look for a necessary condition on overdetermination, i.e. the first step to be taken if one chooses the second comeback strategy.¹²⁵

Bennett, for instance, is committed to the following traditional necessary condition on overdetermination, **(TNC)**:

Happily, there is a very intuitive and widely accepted requirement available. It goes like this: e is overdetermined by c_1 and c_2 only if

¹²⁵ Then, as a second step, the research program to be followed by the compatibilist is look for a scenario – more precisely, a relation R holding between mental causes and physical causes – where (NC) is not met.

(O1) if c_1 had happened without c_2 , e would still have happened: $(c_1 \& \sim c_2) \Box \rightarrow e$, and (O2) if c_2 had happened without c_1 , e would still have happened: $(c_2 \& \sim c_1) \Box \rightarrow e$. (p. 276, my italics)

Then, although in the absence of a fully-worked out argumentation, we are invited by

Bennett to suppose - "at least for the sake of argument" - that overdetermination

would require not only the truth of both counterfactuals, but rather their non-vacuous

truth.

Note that Bennet invokes again broadly agreement and intuition in order to ground one's commitment to (TNC). In her note 8, we are presented with numerous citations

intended to witness that the above test is a commonplace:

Mills (1996, 107) uses the test very explicitly. In general, though, the counterfactuals are not put forward as an official test for overdetermination, but are rather mentioned in passing, given as a definition of 'screening off', or implicitly relied upon in explaining the problem overdetermination (or preemption) tends to pose for various theories of causation. To select some citations almost at random: Horgan 1987, 508–509; Kim 1989a, 252 and 253; Lepore and Loewer 1987, 639; Lewis 1973a, 193, and 2000, 183; McDermott 1995, 523–524; Mellor 1995, 101. Notice too that those who think that overdetermination is impossible because they think that events are extremely modally fragile implicitly rely on this sort of test.

But there is no need to say that, even though an idea is largely shared by a scientific community at a certain time, it still might be false. The history of ideas is full of examples. True, sometimes broadly acceptance may be helpful, but it cannot really serve in establishing whether something is the case.

It is true that Bennett takes her time to persuade us that the truth of (O_1) and (O_2) is

necessary for overdetermination. The main reason, she says, is:

[T]hey capture the reasoning we engage in when we want to distinguish cases of genuine overdetermination from cases of joint causation, or from cases in which one of the putative causes is not really a cause at all. (p. 477)

Of course, nobody would deny that a necessary condition on overdetermination, if it is to be a necessary condition at all, should reflect what distinguishes cases of genuine overdetermination from fake ones. However, admitting that the question is what reason we have for claiming that the non-vacuous truth of the two counterfactuals is a necessary condition on overdetermination, by saying that they do what any necessary condition on overdetermination should do, it is not answered the question. Rather, it pushes to a rephrase of the initial question: for what reason is it claimed that the non vacuous truth of the two counterfactuals does what any necessary condition on overdetermination should do – namely, discriminating overdetermination from pseudo- or non-overdetermination?

Considering the example of a man shot dead by two members of a firing squad,

Bennett says:

If we needed to decide whether or not the death was overdetermined, we would ask precisely whether these two counterfactuals are true. Would the victim have died if the first gunman had fired without the second? Would he have died if the second gunman had fired without the first? If the answer to both questions is 'no'—if both counterfactuals are false—then the death was not overdetermined, for it was jointly caused by the two gunshots. If only one of the counterfactuals is false, at most one of the gunmen is guilty. So the truth of the counterfactuals does play an important role in our willingness to say that some effect is overdetermined. (p. 477)¹²⁶

The above reasoning amounts to saying that the two counterfactuals "capture the reasoning we engage" in deciding whether or not something was overdetermined because the respective reasoning consists exactly in analyzing the truth values of the

¹²⁶ Note that the whole story is of any significance for the matter of overdetermination only if both shots hit the victim. Bennett's test is given under the assumptions that both c_1 and c_2 are causes of e. But, then, given that the two shots jointly damage the victim's body, against Bennett's view, if the victim's death is the result of the actual damage, them the two members of the firing squad are jointly guilt, *no matter the truth-values of the counterfactuals*. Unlike the causal sufficiency, guilty is not to be judged according to whether or not the victim would have died without the contribution of one of the gunmen. An argument showing that it is wrong to say that "if only one of the counterfactuals is false, at most one of the gunmen is guilty" is that, in the case when the two counterfactuals are false, if you are to be coherent with the first claim you have to say that none of the gunmen is guilty and this is hardly acceptable.

two counterfactuals. In other words, Bennett *justifies* the test she advances for overdetermination by simply claiming that the test is (implicitly or explicitly) used by any of us whenever one wants to find out whether or no an effect was overdetermined. I very much doubt the pretended *over*-usage of such a "reasoning".

Bennett's defence of (TNC) makes use again of a pretended intuitional support. She parades a number of cases where c_1 and c_2 are each causally sufficient for e, and where we are allegedly assisted by intuition or by commonsense in deciding whether or not e is overdetermined. And – guess what? – the decisions come out to be in accordance with (TNC).

I believe that it will be instructive to review shortly some of the cases mentioned in Bennett's paper. Let us consider an effect, *e*, that has two sufficient causes, c_1 and c_2 . Bennett discusses one by one various cases, each case being characterized by a certain relation between c_1 and c_2 :

1) a relation of causal sufficiency between c_1 and c_2 ; c_1 causes e only through c_2 (see

Figure 6.1_1):

Take, for example, the already noted issue about cases in which c_1 is causally sufficient for c_2 , which is in turn causally sufficient for e—cases in which c_1 and c_2 are parts of the same causal chain. Our intuitions are quite univocal that these are not cases of overdetermination ... (p. 478)

2) a relation of causal sufficiency between c_1 and c_2 ; c_1 directly guarantees e (see

Figure 6.1_2):

Suppose that there is a causal chain leading from c_1 through c_2 to e, and that c_1 is also directly causally sufficient for e. ... [This case] does not violate the counterfactual test; both (O1) and (O2) come out true. And this is all to the good, as it seems to me that if this sort of case is possible, the effect is indeed overdetermined. (pp. 478-9)

3) a relation of counterfactual dependency (e.g. $\sim c_2 \square \rightarrow \sim c_1$) (see Figure 6.1_3):

[T]he test predicts that a mere counterfactual connection between the causes is not enough to defuse overdetermination. This is as it should be, since it is easy to concoct clear cases of overdetermination in which the two causes are counterfactually connected. Such counterfactual connections might well be the norm in real firing squad cases—if either gunman were to shoot, the other would as well, and if one did not, the other would not either. (p. 479)

4) a relation of identity) (see Figure 6.1_4):

Both [counterfactuals] will come out vacuously true if c_1 is identical to c_2 , but it is a conceptual truth, if anything is, that one event cannot overdetermine another all by itself. (I owe this point to Mills 1996, 107). (p. 480)



Fig. 6.1

CEU eTD Collection

As I have remarked above, in the case illustrated by Fig. 6.1_1, Bennett's intuition seems to be correct. The same works for the case illustrated by Fig. 6.1_4. But let us focus on the case illustrated by Fig. 6.1_2. According to Bennett, in this case, counterfactual (O₁), namely, if c_1 had happened without c_2 , e would still have

happened,¹²⁷ comes out true. In what follows, I intend to find out whether there is a reason for making such a claim about (O_1) .

Following Bennett in granting the Lewisian analysis of counterfactuals, (O₁) is true if and only if, within the closest possible worlds where c_1 happens without c_2 , ehappens. Thus, let us consider the following answer: within a possible world where c_1 happens without c_2 , e should happen because c_1 is directly causally sufficient for e. In reply, one may argue that the answer is inconclusive because, without knowing what it is for a cause to be sufficient for a given effect, you cannot conclude that the effect happens within a possible world just because a cause, which is directly sufficient for the effect within the *actual* world, happens in the *possible* world. For one thing, within the same possible world, c_2 does not happen even though its actual direct sufficient cause, c_1 , happens.

But, would the above objection have the same appeal if the previous answer were completed with something like the following: '... in virtue of the fact that the possible world is one of the closest to the actual one'? *Prima facie*, it may seem that the completion provides us with a reason to conclude that *e* occurs in the respective possible world. The reason would be the similarity between the actual world and the possible worlds under discussion. Invoking the similarity, one might claim the preservation of the causal sufficiency between c_1 and e. That the causal sufficiency between c_1 and c_2 is broken may show, at most, that, unlike metaphysical necessitation, causal sufficiency may be altered across possible worlds. Instead, the

¹²⁷ Formally: $(c_1 \& \sim c_2) \Box \rightarrow e$.

alteration of the causal sufficiency between c_1 and e is prevented by the constraint of similarity between the two worlds.

Yet, I believe that the above comeback is rather misleading. Even though similarity is a crucial guidance in finding the true-value of the counterfactual, similarity does not support just by itself (no matter what) the truth-value attribution. A few remarks are in order here. Let us consider the following counterfactual: $p \square \rightarrow q$. If the truth-maker of *q* belongs to the actual world, the counterfactual conveys that, within all possible worlds "obtained" from the actual world by altering it in a minimal manner to render *p* true, the alteration does not affect the truth-value of *q*. Thus, if *q* is actually true, the counterfactual is true if and only if the most minimal alteration required by rendering *p* true does not render *q* false. Therefore, in judging the truth-value of such a counterfactual, it is not by default that the most minimal alterations rendering *p* true guarantee the preservation of *q*'s truth-value. Rather, the very matter at stake here is whether such alterations leave unchanged that *q* is true.

In conclusion, it cannot be an argument that, given that c_1 happens in the closest possible worlds, in the name of similarity, c_1 's status of direct sufficient cause of e is preserved.¹²⁸ Probably, a proper argument, if any, in favor of such preservation should insist on the idea that c_1 's *loosing its status of sufficient cause of* c_2 *does not require* c_1 's *loosing also its status of sufficient cause of* e and, hence, the latter alteration is not specific to the closest possible worlds. However, for making such an

¹²⁸ Note that, actually, if we do not want to commit ourselves to any theory of causality and, in particular, to Lewisian counterfactual analysis of causality, within the counterfactual (O₁), proposition *q* should be expressed not by 'e happens', but rather by ' c_1 makes by itself e happens' which is equivalent to ' c_1 is directly causally sufficient for *e*'. I take Bennett's preference of using 'e happens' to be motivated by the reasons of simplicity.

argument, one needs to know what is for a cause to be sufficient for some effect. Only on this *terra firma*, namely knowing the underlying conditions of being a sufficient cause, one has the possibility of establishing in each particular case whether or not there is a link between the conditions of being a sufficient cause of one effect, in our case c_2 , and the conditions of being a sufficient cause of another effect, in our case *e*. I take these considerations to be a plea for a deeper analysis of the notion of causal sufficiency.

But, the true-values of the two counterfactuals aside, what bother me most are not the assessments Bennett proclaims about the existence or non-existence of overdetermined effects in each of the above cases, but rather my lack of intuitions or my absent clarity in cases where she pretends to be fully assisted by largely shared intuitions in making the respective assessments. If I have an intuition in this matter, it is only the modest one that, contrary to Bennett, given any effect, it is *not* a matter of *intuition* in deciding whether or not the effect is overdetermined. Of course, 'intuition' is used here in order to refer to an epistemic situation where something strikes you as being the case, that is, a situation in which, without any further analysis, something seems to be such-and-such in virtue of its appearance. But, given that 'overdetermination' is a derivate notion, one would expect that its application relies firstly on conceptual analysis, that is, on a reduction to more primitive notions.

One may say that, in some cases, it is the very *meaning* of 'overdetermination' that enables us to see *clear* whether or not the effect is overdetermined. Maybe, but what is that meaning? Bennett does not let us know what her answer (if any) is; she does not make explicit anywhere what the general understanding of the notion is.

My belief is that one should not approach the question of whether or not some case is one of overdetermination by appealing to intuitions. We are not dealing with a matter of intuition, but rather with a matter of meaning of 'overdetermination'. Even in the simplest case, namely the case where the two sufficient causes feature into the same causal chain, it is not really intuition that guides us in correctly saying that the effect is not overdetermined. Rather, it is elementary reasoning that tells us that 'overdetermination' must refer to something (if any) particular in a process of causation.¹²⁹ If 'overdetermination' refers inclusively to regular cases of causation like that when the alleged overdetermining causes feature in the same causal chain leading to the effect, then the term would lose any particularity, i.e. there would be no difference between a case of overdetermination and any domestic case of causation. Thus, it is the full transparency of a gross conceptual fault that prevents us from mistaking here.

I admit that in some cases, intuition and broad agreement may serve as ultimate support in making certain judgments, but I consider that such kind of sources should be avoided as much as possible. Then, the question is what would guide us in finding an interesting condition that each and every case of overdetermination should meet.¹³⁰ My proposal, as it will become clear from the next section, is conceptual analysis.

¹²⁹ I shall try to reveal in the next sections what might approximate that particularity.

¹³⁰ Recall that we want to know whether the compatibilist has any chance to answer the CEP and hence an interesting condition would be one helping the compatibilist to show that not always, when a case conforms to the CEP assertions is also a case of overdetermination.

6.2 Going for something more basic

I have maintained not only once that it is not a matter of *intuition* whether or not the effect is overdetermined. Rather, given that 'overdetermination' is a derivative notion, one would expect that its application relies firstly on conceptual analysis, on a reduction to more primitive notions. In this section I shall undertake such a conceptual analysis to the effect of coming closer to a general understanding of overdetermination in more manageable (and intuitive) terms.

The verb 'overdetermine' belongs to the family of verbs whose general form is oververb, where verb is a variable taking as values verbs like 'buy', 'do', 'develop', 'grow', 'produce', 'pay', 'sell', etc., that is, verbs referring to human actions or unanimated processes. Henceforth, I shall use the term 'process' both for actions and processes.¹³¹ Thus, a verb of the same kind as 'overdetermine', when it is used, stands for a process such that the relation between the process and its result is to be characterized as an *excess*¹³². Such an excess may occur either on the level of the obtained result or on the generation level.

If there is an excess in the actual result, the excess comes from the difference obtained over the *reference* result. The reference result might be the expected result, i.e. the outcome we normally associate with the process referred by '*verb*'. The

¹³¹ There is the temptation of making a distinction by saying that actions appear to be occasioned by a goal (or by an intention to accomplish a certain goal), whereas, in the case of processes, the intentional component is not present. However, given that such a distinction would rather hinder than help us, I shall make reference to it only when it is strictly necessary, preferring to use the term 'process' to refer to a case where some changes happen and, in accordance, a consequence is brought about.

¹³² Sometimes, it is *exaggeration* that characterizes such cases (see, e.g., verbs like 'overreact'), but it seems to me that 'overdetermine' is closer to verbs which signal an *excess*.

reference result might also mean what motivates the initiation of the process. Within the class of processes referred by '*verb*', there is a sub-class referred by '*verb*'; by using '*over-verb*' refers to a process that can also be correctly referred by '*verb*'; by using an instance of '*over-verb*', it is conveyed that the obtained result goes beyond what would normally trigger any process referred generically by '*verb*'. For example, to overbuy means to buy more than is needed. That is, by overbuying, besides getting some useful stuff, one ends up with unnecessary stuff. Of course, the reference result for such kind of process is getting what it is needed. Other examples are 'overdraw', 'overproduce', 'overkill', etc.

However, 'overdetermine' does not signals the occurrence of a surplus relative to a particular reference result. Rather, confining the discourse to the sub-class of oververb family to which 'overdetermine' belongs, the excess is to be established taking into account how the actual result is obtained. The clearest cases are those where, by saying that the result, *e*, is over-*verb*-ed, it is meant that some process that actually produced *e* exceeds a non-actual process of the same kind, that is sufficient for producing e.¹³³

For instance, on a free market, if there is such a thing, if we take seriously some economical laws, for getting a certain kind of good there is an *objective* price to be paid. Thus, by saying that one overpaid for *e* we should understand that one paid for

¹³³ Whether or not a process is sufficient for the occurrence of the result does not depend on the actuality of the process. For example, we say that one would have got a certain car if one had paid the right price; one's paying the right price *is* a sufficient process for one's getting the car, but one's paying the right price is not an actual process. This is an example of a result that is not achieved: the person did not pay the respective price (s/he did not afford such a price) and could not get the car. Another example is that when the person got the car, but after s/he paid more than it is reasonable.

Moreover, I claim, sufficiency should be judged relative to a world, but it should not be understood as confined to a world. It should be treated as an attribute which may preserve across the relevantly similar worlds.

e too much, or more than the *due amount*. The same works for 'overcompensate' where you can assume that the expected result – namely that of establishing equity by repayment – must be obtained by a determinable process.

However, usually, a given result might be reached by various means; hence, among the various processes sufficient for its obtaining, it is too hard or even impossible to establish which is the minimal or the right or the proper one. Either there is no such precise correspondence or we seldom have at hand the necessary means for identifying it. Therefore, in those cases characterized by the existence of some excess relative to some non-actual process, any process sufficient for the result, but exceeded by the actual process, does the job. For instance, a falling brick is capable of smashing an egg; thus, relative to a possible falling brick, an actual exploding bomb which smashes the egg should be taken to be an excess in achieving the respective result.

Yet, overdetermination should be understood as corresponding to a different kind of excess. In any case of overdetermination, the excess in generating the result comes from the existence of, at least, two *actual* causal processes sufficient each for the same ending effect. It appears that an overdetermined effect is brought about by more than one sufficient causal process in a manner that involves redundant causal resources.

Thus, considering c_1 , c_2 , c_3 , ..., some causes of e, it is legitimately to say that they overdetermine e iff:

- 1) There are two processes, Π and Φ , each of them being causally sufficient for e, such that some events of c_1 , c_2 , c_3 , ... feature among the originating causes of the processes;
- 2) The existence of Π and Φ amounts to an excess in bringing about e^{134}

It is important to note that the idea of overdetermination, at least in this understanding, is not dependent on the idea of sufficient cause, but rather of that of sufficient causal process. The advantage of this approach is that, by knowing that *e* has, say, two causes, c_1 and c_2 , for answering whether *e* is overdetermined by c_1 and c_2 , one is not bound to establish whether c_1 and c_2 are each sufficient causes of *e*. Rather, what matters for the required answer is where there are two causal processes, each having c_1 or c_2 among their originating causes and each being sufficient for obtaining *e*. It is not causes, whether sufficient of partial, but rather or not there is some excess in bringing about the effect.¹³⁵

¹³⁴ Of course, assuming that c_i is one of the causes that do not feature among the originating causes of any of the two processes whose occurrences amount to an excess in bringing about e, the mentioning of c_i in the list of the overdetermining causes of e might be somehow superfluous. I choose not to be very strict in this matter, because the main aim of the present analysis is to provide a tool for searching whether an effect, in a given case, is overdetermined. Thus, if the initially listed causes of the effect are not exactly those that overdetermine the effect, some of them missing from the respective list, it is alright. The reason is that *the role of the mentioned causes is just to procure the base (even if it is incomplete or overloaded) where one should be looking for the originating causes of two processes causally sufficient for the effect. Highly precision is not needed here. Given the listed causes, the two sets of originating causes might be found either by selection, or by completion (with other, unmentioned causes), or by using exactly the whole list of causes. Whether the two identified sufficient causal processes amount to an excess is the crucial matter for deciding whether the effect is overdetermined.*

¹³⁵ True, as we shall see in the next section, a sufficient cause of an effect is nothing else than the only originating cause of a causal process sufficient for the same effect. However, the above stress on the idea of causal process should count later, when we shall analyze whether an effect is overdetermined by mental causation. For, even if some of the mental events are only partial causes of the considered effect, they feature among the originating causes (some of those being physical in nature) of some causal processes sufficient for the effect. Again, it is those processes that should come under scrutiny in order to find out whether there is some excess in bringing about the effect.

6.3 Clearing up two notions

In the previous section we have found out that the understanding of 'overdetermination' is grounded on the grasping other two notions: *causal process sufficient for an effect* and, given two such processes, *excess in bringing about an effect*. The clarification of these notions is the task for the present section.

Let us begin with the easier job, namely, that of answering what it is for a causal process to be sufficient for an effect. Recall that a causal process is characterized by a triplet consisting in some originating causes, some ending effects, and its linking network – i.e. a network made up of some causal chains going through the originating causes and leading to the ending effects.¹³⁶ Then, by saying that a given causal process is sufficient for one of its ending effects, it is meant that its originating causes make up a causally sufficient condition for the respective effect. Of course, the other way round, relative to some event, some events make up a causally sufficient for only if there is some causal process sufficient for the respective effect such that its originating causes are the very considered events. A sufficient cause of some effect is nothing else than the member of some causally sufficient condition having only one single element.

¹³⁶ It is also important to keep in mind that the notion of causal process requires that, for any causal process, Π , any ending effect of Π is a point of convergence for some causal chains going through all originating causes. The reason for imposing such a constraint is that the presence of two events, *x* and *y*, one being an originating cause of Π , the other being an ending effect of Π , such that *x* is not a cause of *y* would be a case in disagreement with the fundamental feature of any causal process, namely that of causally relating two sets of events – i.e. the heads of the process.



Fig. 6.2

According to the approach of causation to which I am committed, any causal relation is part of some relation of nomic necessity. Then, given that any causal process is a network of causal relations, it follows that any causal process is part of some nomic necessitation. For instance, let us assume the existence of a causal process, Π , whose originating causes and ending effects are c_1 , c_2 , c_3 , ..., c_n and e_1 , e_2 , e_3 , ..., e_m , respectively. (See Figure 6.2.) It follows that there is some relation of nomic necessity holding between n_1 , n_2 , n_3 , ... (i.e. the set composed of the necessitating relata) and n_1^* , n_2^* , n_3^* , ... (i.e. the set composed of some of the necessitated relata) such that *L*s, the laws supporting the necessitation, are causal in nature. Moreover, c_1 , c_2 , c_3 , ..., c_n feature among n_1 , n_2 , n_3 , ..., and e_1 , e_2 , e_3 , ..., e_m feature among n_1^* , n_2^* , n_3^* ,

A direct and natural consequence is that a process should be taken to be causally sufficient for one of its ending effects, *e*, if and only if, considering the corresponding relation of nomic necessity, the originating causes of the process are identical with the necessitating relata, and *e* features among the necessitated relata. Of course, some of the causes of *e* make up a causally sufficient condition for it if and only if there is some relation of nomic necessity such that the causes are identical with the necessitating relata, and *e* features among the necessitated relata.

Further, given *c* and *e*, some cause and its effect, *c* is causally sufficient for *e* if and only if *c* by itself nomically necessitates *e*; that is, when one talks about some sufficient cause of a given effect one should have in mind that the cause-event necessitates effect-event in virtue of some causal law of whose triggeration the cause-event is fully responsible.

In what follows, it will be analysed the notion of excess in bringing about an effect. In § 5.5, I have defined the causal role a process, Π , has relative to one of its effects, *e*, as consisting of the causal roles played by Π 's nodes featuring in the causal chains leading to *e*, in respect to their direct effects featuring in the respective chains. The causal role of such a node gains its identity from the final property featuring into the node and from the basic law governing its linking in the causal chain leading to *e*.

Therefore, the identity of the causal role actually played by process Π in respect to *e* is given, on the one hand, by the basic laws, L_{Π} s, underlying the nomic necessitation whose part process Π is, and, on the other hand, by the properties, N_{Π} s, instantiated in the constitution of Π 's nodes featuring in the causal chains leading to *e*. Thus, the identity of $R_{@}(\Pi, e)$ is given by the doublet $<L_{\Pi}$ s, N_{Π} s>.

Besides Π , let Φ be another causal process having *e* as ending effect. Supposing that each of the two processes is causally sufficient for *e*, the question is what would be for Φ and Π to be engaged redundantly in bringing about *e*. Note that the causal sufficiency each of them has in respect to *e* is not enough for claiming the existence of an excess in bringing about *e*. For example, it is no doubt that the case when one process is a serial part of the other is not a case of overdetermination. (See Figure 6.3.) To expertise this case as not being a case of overdetermination is to be coherent with the common-sensical thought that the doings of the *ancestors* of some occurrence cannot be deemed as an excess in reference to the doings of the occurrence. Or else, on pain of trivializing the notion of overdetermination, any effect is taken to be overdetermined. For one thing, given any sufficient process causally responsible for the occurrence of the effect, somewhere along the causal history of the effect, there is always another process to whom the former process is a serial part.



Fig. 6.3

To my mind, the existence of two processes causally sufficient for *e* amounts to an excess in bringing *e* if and only if the sufficiency of each process is not dependent on the sufficiency of the other. In other words, given that Φ and Π are each causally sufficient for *e*, there is an excess in causing *e* if and only if neither Π 's causal role with respect to *e* depends on Φ 's causal role with respect to *e*, nor *vice versa*.

A necessary condition for the independence required above is the truth of the following counterfactuals:

(α_1) if Π 's nodes and their causal ancestors were the only causes of *e*, then, relative to *e*, Π 's causal role would be identical with its actual causal role.

(α_2) if ϕ 's nodes and their causal ancestors were the only causes of *e*, then, relative to *e*, ϕ 's causal role would be identical with its actual causal role.

Given that Π is causally sufficient for *e*'s occurrence, by activating laws L_{Π} s, the tokening of properties N_{Π} s necessitates the tokening of *E*. Therefore, within counterfactual (α_1), to have Π 's causal role identical with its actual causal role is simply to be in a world where *e* occurs in virtue of L_{Π} s's triggeration by N_{Π} s's tokening. Similar considerations can be developed regarding counterfactual (α_2).

6.4 A necessary condition on overdetermination

Let us suppose that d_1 , d_2 ,... d_n are the nodes of Φ that are direct causes of e. Given that Φ is causally sufficient for e, it follows that there is a relation of nomic necessity holding in virtue of some fundamental law, L_{Φ} , and relating any given instances of properties D_1 , D_2 , ..., and D_n with some instance of property E. The former properties token respectively in the constitution of d_1 , d_2 , ..., and d_n and E tokens in e's constitution. (See Figure 6.4.)



Let *w* be a possible world where Π 's nodes occur and *e* is one of Π 's ending effects. Assuming that none of d_1 , d_2 ,... d_n is a node of Π , ¹³⁷ the antecedent of

¹³⁷ For simplicity, in the beginning, it will be considered the case where Φ and Π share no common part. In the case where Φ and Π share some common part, that part is a process that includes some of the nodes of Π that are direct causes of *e*. The latter case will be discussed in the end of this section.

counterfactual (α_1) is true at *w* if and only if none of $d_1, d_2, ..., d_n$ has any causal role on *e* in *w* and none of the events occurring in *w* replaces d_1 or d_2 ...or d_n in directly causing *e*. Given that the tokening of properties $D_1, D_2, ..., D_n$ actually activates law L_{ϕ} in virtue of which *E* is instantiated, and granting that the laws holding within a world supervene on the distribution of property instances in that world,¹³⁸ L_{ϕ} holds in any possible world where, whenever $D_1, D_2, ..., D_n$ are all instantiated, *E* is instantiated too. It follows that if properties $D_1, D_2, ..., D_n$ are instantiated when *e* occurs as an effect of process Π in one of the closest possible worlds, then L_{ϕ} holds in that possible world and, consequently, in that possible world, either $d_1, d_2, ..., d_n$ are all direct causes of *e* or some other events have their causal roles on *e*. In conclusion, if *w* is one of the closest possible worlds at which the antecedent of counterfactual (α_1) is true then:

(β_1) it is at least one of properties D_1 , D_2 ,... D_n that is not instantiated in *w* when *e* occurs as an effect of process Π .

Therefore, counterfactual (α_1) has the same truth value as the counterfactual obtained from (α_1) by conjugating its antecedent with the clause expressed by (β_1). Similar considerations may be made about (α_2).

Thus, we are now able to formulate a necessary condition on overdetermination. An effect *e* is overdetermined by two or more of its causes, c_1 , c_2 , ..., only if **(CNC)**:

¹³⁸ Carroll (1994) challenges the thesis that the laws supervene on the distribution of facts, but I find his attempt to beg the question.

(1) there are two processes, Φ and Π , each of them being causally sufficient for *e*, such that some of c_1 , c_2 , c_3 , ... feature among the originating causes of the processes;

(2) the following counterfactuals are non-vacuously true:

(γ_1) if Π 's nodes and their causal ancestors were the only causes of *e*, at least one of the properties of *e*'s direct causes from [$\Phi - \Pi$] being uninstantiated on the occasion of *e*'s occurrence, then *e* would occur in virtue of L_{Π} s's activation; (γ_2) if Φ 's nodes and their causal ancestors were the only causes of *e*, at least one of the properties of *e*'s direct causes from [$\Pi - \Phi$] being uninstantiated on the occasion of *e*'s occurrence, then *e* would occur in virtue of L_{ϕ} s's activation.

Even though I do not present here a fully-worked out argument, I believe that (CNC) must be supplemented with one more restriction:

(3) Φ and Π do not stand in a part-whole relationship.¹³⁹

The crucial reason for imposing (3) is that we do not want to pass (CNC) cases like the following: c_1 is causally sufficient for c_2 , and c_2 is causally sufficient for e. See § 6.1, especially the case illustrated in Figure 6.1_1.

One more thing before closing this section. In the above condition, there is no specification to the effect of excluding a case of overdetermination where the two causal processes, Φ and Π , share one of their parts. In § 5.4.3, it has been shown that, if Φ and Π have something in common besides *e*, the largest part shared by Φ and Π , [$\Phi \cap \Pi$], is a causal process whose single ending effect is *e*, such that [$\Phi \cap$

 $^{^{139}}$ For a discussion on what is for a causal process to be a part of another causal process, see § 5.4.2.

 Π is either a parallel part of Φ (i.e., the case where the originating causes of $[\Phi \cap \Pi]$ are exclusively among those of ϕ – see Figure 6.5a), or a parallel part of Π (i.e., the case where the originating causes of $[\Phi \cap \Pi]$ are exclusively among those of Π), or a parallel part of both Φ and Π (i.e., the case where the originating causes of $[\Phi \cap \Pi]$ are exactly the originating causes shared by ϕ and Π – see Figure 6.5b).



Fig. 6.5a

Fig. 6.5b

Let us consider the latter case, namely, the one where $[\Phi \cap \Pi]$ is a parallel part of both Φ and Π . It follows that the originating causes of $[\Phi \cap \Pi]$ are exactly the originating causes shared by Φ and Π . Let us say that c_i and c_i are the common originating causes. (See Figure 6.6.)

In such a case, by requiring that Π 's nodes and their causal ancestors are the only causes of e, the antecedent of the counterfactual (γ_1) requires, among others, that ϕ 's nodes cease to be cause of e, being excepted from the requirement those nodes of $\boldsymbol{\Phi}$ shared with $\boldsymbol{\Pi}$ – namely, the nodes belonging to $[\boldsymbol{\Phi} \cap \boldsymbol{\Pi}]$.



6.5 Unnecessary causes, overdetermining causes

It may be noted that there is a prevalent belief that causes are not unnecessary for their effects: no matter to what account of causation one is committed, *normally*, causes should be necessary for their effects. But what should we understand by 'cause unnecessary for its effect'? Simply put, in the absence of the cause, the effect would still occur. That is, a cause, *c*, is unnecessary for *e*, one of its effects, if and only if the following counterfactual is non-vacuously true: even if *c* had been absent, *e* would have been brought about.

Let it be two counterfactual propositions with the same antecedent. If there is a world where the antecedent is true (i.e. if the counterfactuals are not vacuously true), then the two propositions are contradictory. Thus, 'if c had been absent, then e would
have been absent too' has the same logical value with the negation of 'if c had been absent, then e would have been brought about'. It follows that counterfactual 'if c had been absent, then e would have been absent too' says that c is a necessary cause of e – i.e. a causally necessary condition of e. Note that 'causally necessary' involves 'counterfactually necessary'. To say that c is a necessary cause for e is not is not to say that it is impossible to have e without c, because there is no guarantee that there is no world where e is produced in the absence of c.

Then, given that *c* is a cause of *e*, according to the counterfactual account of causation, if there is also a relation of causal dependency between *c* and *e*, cause *c* is necessary for its effect, *e*. If by 'direct cause' it is meant a cause on which the effect is causally dependent, it follows that, the counterfactual approach renders any direct cause as not being unnecessary for its effect.¹⁴⁰

On the one hand, generally speaking, there is the prevalent view that a case where a cause is unnecessary for its effect is an aberration, an abnormality; hence such a case, if any, should be tolerated (accepted) at most as an exception. On the other hand, it is a very popular idea that an event, *e*, is overdetermined by some direct

¹⁴⁰ As I have remarked before, in Lewis's view, any cause is linked to anyone of its effects through a chain of causal dependences. Of course, then, if *e* is causally dependent on *c*, *c* is a cause of *e* and, by definition, had *c* been absent, *e* would have been absent too. Further, it follows that if *e* is causally dependent on *c*, *c* is a necessary cause of *e*.

Maybe, this is why Lewis prefers not to discuss "symmetrical cases of overdetermination, in which two overdetermining factors have equal claim to count as causes", admitting that he lacks "firm naive opinions about them" (1986 [1973]: 171, note 12). By 'symmetrical cases of overdetermination', Lewis seems to mean those cases where, an event *e* has two causes, c_1 and c_2 , such that, for each cause, there is a causal chain of dependencies linking the respective cause to the effect. Within each of these causal chains, there is a direct cause of e, namely, c_1^* and, respectively, c_2^* . But, as we have noted above, *e* is causally dependent on each of these direct causes and, implicitly, each of these direct causes is necessary for *e*. For instance, had c_1^* been absent, *e* should have been absent too. But, considering a case where the presence or the absence of c_2^* is indifferent to the absence of c_1^* , 'had c_1^* been absent, *e* would have been absent too' could be false, because, even in the absence of c_1^* , c_2^* might cause *e* to occur, rendering c_1^* as an unnecessary cause for *e*.

causes, let *c* be one of them, only if the following counterfactual is non-vacuously true: 'had *c* been absent, all other direct causes of *e* remaining unchanged, *e* would still have occurred'. Thus, in this acceptance, it is tempting to assimilate a cause overdetermining some effect with a cause unnecessary for the effect. Accordingly, given that the existence of unnecessary cause is seen as an aberration, it would be clear why the idea of systematic and widespread overdetermination is unacceptable: one cannot rationally embrace a systematic and widespread aberration.

Yet, considering c and e such that c is a direct cause of e, is it always the case that the closest worlds where 'c is absent' is true are the same as the closest worlds where 'c is absent, all other direct causes of e remaining unchanged' is true? Unfortunately, I do not have strong reasons for rejecting the cases where there is a certain kind of relation holding between c and some other direct cause, c^* , of e such that it is impossible that c is absent and c^* occurs. Therefore, given an effect e, I do not feel free to take its unnecessary causes as being overdetermining causes. I do not have enough reasons for "jumping" from unnecessary causation to overdetermination.

Moreover, being committed to a nomic account of causation, I believe that both causal sufficiency and causal necessity should be understood in the light of the sufficiency and the necessity revealed by the nomic relations. A relation of nomic necessity holding between the instances of two properties, *A* and *B*, amounts to a necessitation of *B*-instances by *A*-instances. But, of course, this is not a reason for saying that *A*-instances are necessary for *B*-instances. The assertion of some nomic relation holding between *A*-instances and *B*-instances is accompanied by neutrality

regarding whether or not *A*-instances are necessary for *B*-instances. It only implies that any *A*-instance is sufficient for some *B*-instance, at least in the actual world and any other possible world empowered by the same laws that actually sustain the necessitation of *B*-instances by *A*-instances. (Any *A*-instance necessitates some *B*-instance in virtue of some laws triggered any time an *A*-instance is present.)

According to the nomic account of causation, a causal relation is a *part* of some relation of nomic necessity. Also, a causal relation where a cause is sufficient for its effect is a *whole* relation of nomic necessity. As I have shown in the previous section, for *e* to be overdetermined, *e* has to have at least two causes such that each of them is independently sufficient for the occurrence of *e*. Recall that independent causal sufficiency requires a preservation of the causal role under the circumstance when the cause or, better said, the causal process in question were abstracted from the rest of causal processes converging to the same effect. Thus, the right thing to say is that 'cause overdetermining *e*' is rather synonym to 'cause sufficient for *e*, independently of other direct causes of *e*, if, independently of *c**, *c* is causally sufficient for *e*, then *c** is causally unnecessary for *e*,¹⁴¹ but the reciprocal is false. Therefore, it is not enough that *e* has unnecessary causes for concluding that *e* is overdetermined.

¹⁴¹ Formally put, the implication can be write as follows:

⁽*c* occurs without $c^* \square \rightarrow c$ has the same sufficient causal role in bringing about e) \rightarrow (c^* is absent $\square \rightarrow e$ is still brought about)

6.6 Looking around for a case of overdetermination

The previous sections involved us in a good deal of talk about overdetermination, but no particular case of overdetermination has come under examination. This section will review some classical (and, also, notorious) examples of claimed cases of overdetermination. Besides these, by the same pattern, some other new examples will be added. This will lead us to the general kind of structure usually employed when one tries to provide an example of overdetermination. The emphasize of this structure is intended to serve in building on an argument for the idea that the most favored examples mentioned in the literature on overdetermination are highly debatable, many of them illustrating only cases of pseudo-overdetermination or partial overdetermination.

One of the most famous or, rather, cited example of overdetermination is the following: a man is shot dead by two assassins or by two members of a firing squad whose bullets hit him at the same time in some vital points such that each of the two gunshots is causally sufficient for bringing about the death; hence, at least *prima facie*, the victim's death is overdetermined.¹⁴² Of course, there could be imagined a lot of versions of the given example such as the case when a man falls over two sharp pales, being deadly wounded in each of the two places of penetration.

Another notorious example of overdetermination is that of a building catching fire because of two distinct and independently sufficient causes. For instance, the causes could be a short circuit in the faulty wiring and a bolt of lightning that hits the building

¹⁴² See, for instance, Kim (1989), Braddon-Mitchell and Jackson (1996), and Bennett (2003).

at the same instant.¹⁴³ It inspires me to propose another example in the same vein: in the some electric device, during the time when it is plugged in, two short circuits happens at the same time in two different points, each of them being sufficient for ending the electric activity of the device.

As a matter of fact, Kim allows many other physical events to be overdetermined. For example, he says that both the loudness and the suddenness of a thunderclap are causally sufficient for his jumping.¹⁴⁴ Branddon-Mitchell and Jackson (1996) are also generous enough with overdertermination cases. For instance:

Suppose that Jane and Mary are both persuading Bill of the merits of socialism. He is persuaded, but we know that Jane and Mary are both such good arguers that either of them alone could have done the job. We say that Bill's being persuaded was overdetermined. Again, the death of some one who has major heart failure while falling from the top of a skyscraper will typically be overdetermined. (p. 10)

In the light of the above examples, one might come to believe that the world around us and, especially, the universe of human deeds are full of overdetermined events. In the face of such a deployment of examples, one may think of overdetermination that, even though not systematic and widespread, it is not such an unusual phenomenon as it may seem at the first sight.

¹⁴³ The example spelled in this form belongs to Kim (1993a [1989]: 252).

¹⁴⁴ See Kim (1993a).

6.7 The logic of the threshold function

Prima facie, any of above cases seems to pass any reasonable test one may require for deciding whether the effect is overdetermined. Let us focus on the case of the man killed by two bullets, each hitting the victim deadly. The case looks as if it passes the most popular necessary test on overdetermination, namely, (TNC).¹⁴⁵ The test requires the non-vacuous truth of the following two counterfactuals:

- If b_1 penetrated the victim's brain, b_2 missing the heart, the victim would still have died;

- If b_2 penetrated the victim's heart, b_1 missing the brain, the victim would still have died.

Also, the case passes, at least prima facie, the test I have advanced, namely, (CNC).

In a certain approximation, the test requires the non-vacuous truth of the following two counterfactuals:

- If b_1 penetrated the victim's brain, b_2 missing the heart, b_1 's penetrating the brain would still have brought about victim's dying in the same way as it actually did;

¹⁴⁵ See § 6.1, where it has been discussed the traditional necessary condition on overdetermination. By all means, Kim seems to be one of those philosophers committed to (TNC). For example, in Kim (1993a [1989]), one may find at lest two passages supporting the above claim:

C and C^{*} are distinct and each a sufficient cause of E. We may think of them as belonging to two distinct and independent causal chains. This then is a case of *causal overdetermination*: E would have occurred even if either C or C^{*} had not occurred, or had not caused it; the other would have been sufficient to bring it about. (p. 252)

The other relevant quotation is this:

Consider the two questions: (1) Would *E* have occurred if *C* had not occurred? and (2) Would *E* have occurred if C^* had not occurred? If the answer is a "yes" to both questions, this is a classic case of overdetermination ... (p. 253)

- If b_2 penetrated the victim's heart, b_1 missing the brain, b_2 's penetrating the heart would still have brought about victim's dying in the same way as it actually did.

It seems to me that all cases presented in the previous section are held to be examples of overdetermination in virtue of the assumption that they all follow a certain *logic*. Arguably, the logic is analogical with the logic of activation of a devise presented in Figure 6.7. The device takes two input signals, x_1 and x_2 , makes the sum, *a*, and maps it into an output signal, *y*, according to the following threshold (or step) function:

$$y = \begin{cases} s & \text{if } a \ge \theta \\ 0 & \text{else} \end{cases}$$



Fig. 6.7

The analogy requires that x_1 and x_2 are each higher than threshold θ . Thus, either of them could activate alone an output signal equal with *s* (that is, "the step"). The capacity to activate alone the output signal is a capacity had by each input signal.

Such a capacity is analogical with the power had by each cause in the above cases of causation: each cause has the power of bringing about, by itself, the effect. Further, in the cases of causation, what operates according to a threshold function is a causal law for whose triggeration is necessary and sufficient the exemplification of a certain property. Moreover, no matter how many exemplifications of that property occur in the time, the triggeration of the law happens only once, generating the same result as if there were just a single exemplification of the triggering property.

Let us return to the example of the man killed by two bullets, b_1 and b_2 , assuming that b_1 damages the heart and b_2 explodes the brain. As I have said, it is tempting to say that *the victim's dying* is overdetermined by b_1 's penetrating the brain and b_2 's penetrating the heart. It is pretty obvious that both organs – the brain and the heart – are vital for the man's life: some serious damage of any of them would cause the death of any human being.

It seems that *the penetration of* b_1 is an event that leads to the exemplification of the final property of *hitting a vital organ*. The exemplification of the property triggers some biological law responsible for the event of *ceasing progressively of all vital functions in the victim's body*, that is, the event of *the victim's dying*. The same goes for *the penetration of* b_2 , and for the double penetration. *Prima facie*, the whole thing amounts to saying that no matter which of the two bullets hits the victim, each of the two causes would be independently sufficient in bringing about the same effect, namely, *the victim's dying*. The sameness of the effect is the very assumption that I shall tray to challenge in the next section.

Thus, seemingly, if one is looking for a case of overdetermination, it would be enough to find some causal law (or a set of causal laws) capable of playing the role of a threshold function. Let us consider, for example, the laws according to which, in any electric circuit, any short circuit leads to the melting of some calibrated fusible within the circuit.¹⁴⁶ Thus, no matter how many short circuits occur at the same time in an electric circuit, its calibrated fusible goes melt. The melting of the fusible is allegedly the same event, one whose occurrence would be overdetermined in the case of two simultaneous short circuits. Again, the next section is intended as a contestation of the both claims.

6.8 Pseudo-overdetermination, partial overdetermination

In this section, I shall try to emphasize some considerations that, I hold, might disturb the serenity of the cases under scrutiny (that is, the cases enunciated in § 6.6), and make one to be skeptical about the pretence that they illustrates genuine examples of overdetermination.

Let us consider again the example of the man killed by two bullets. At a closer examination, it must be noted that it does not meet a certain elementary requirement held by any of the two considered necessary tests of overdetermination: *the effect should have been the same, had there been just one cause.* In particular, the victim should have had the same death, had there been just one bullet hitting the victim.

¹⁴⁶ Actually, the law in charge is Joule's first law, also known as the Joule effect. It is a physical law expressing the relation between the heat generated by the current flowing through a conductor: $Q = f^2 x R X t$, where Q is the heat generated by flowing of a constant current *I* through a conductor of electrical resistance *R*, for a time *t*. Admitting that a short circuit is nothing else than an occurrence of excessive current flow, a calibrated fusible is a conductor whose melting point is reached in virtue of the Joule effect when the current flowing though the conductor reaches values of short circuit.

Thus, basically my claim is that, by changing the number of hitting bullets, *the victim's dying* would not have been the same.

As it probably will become apparent in what follows, despite the impression of sameness across the worlds, the effects are different. The case of the man killed by two bullets is nothing else than an example of pseudo-overdetermination. For the sake of working with an enunciation friendly for any of the tests discussed in the previous sections, I propose the following definition:

An event *e* is pseudo-overdetermined by two of its causes, c^* and c^{**} , if the counterfactual effect obtained by any of the causes acting by itself is different from *e*, even though there is a strong temptation to hold the contrary.

As one might note, there are two key things featuring into the idea of pseudooverdetermination. Firstly, it is the theory of events to which one is committed that enables one to decide in the matter of transworld identity of events. Secondly, in virtue of the same theory, one is called to establish whose responsibility is for there being the wrong impression that two events are identical across the worlds, even when they are not.

As I have argued in Chapter 4, I favor the view that takes any event to be a temporal structure of states through which a substance goes during some period. Accordingly, any event consists of a substance (i.e. an entity persisting through changes through time), and a range of properties substance exemplifies successively; the form of the structure is given by a certain *timing* of property exemplification during the period of

considered change. That is to say that an event is a sequencing of states through which a substance goes through in accordance with a timing.

Thus, I believe that any event is characterized by $\langle s, P_i, P_f, (t_i, t_f), p \rangle$, in which *s* is a substance (or a system), P_i is the initial property *s* looses, P_f is the final property *s* acquires by loosing P_i , (t_i, t_f) is the period needed for change to be accomplished, and *p* is the profile of change – *p* consists of the intermediary properties *s* acquires and looses in going from its initial state to its final state, plus the respective moments or periods when *s* has each intermediary property.

Many times, an event involves more than one substance. Rather, it is a system of substances that undergoes changes in its global states. A global state of the system is the exemplification by the system of a property supervenient on some properties exemplified respectively by the constituents of the system. By 'system of substances', it is meant here nothing more than some substances capable of having a property.

Then, an event happening to the system is a transition from an initial global state to a final one, the transition being supervenient on the transitions had by the constituents of the system. In short, an event happening to the system is supervenient on the events happening respectively to each constituent substance.

For example, *the victim's dying* is an event happening to the body of the alive victim; it is nothing else than a successive loss of the vital functions ending up with a dead body. Reasonably, one may very well understand by 'the body of an alive person' a

system made of various cells that, in virtue of their properties, sustain various vital functions.

Given that any event is seen as being a change evolving, in a certain manner, in time, from an initial state to a final state of a substance or system, the profile of the transition is crucial for the identity of events across the possible worlds. Naturally, the profile of change is given by all states the transition goes through and by the timing the states succeed one by one. It follows that, given two events happening in two different possible worlds, in order to be different, it is enough for the two events either to have different constituent substances, or to have different profiles of change.

Now, let us come back to the example of the man killed by two bullets, b_1 and b_2 , b_1 damaging his heart and b_2 exploding his brain. As I have mentioned already, in my view, there are good reasons for asserting that had only b_1 hit its target, the victim would have had a different death. In other words, my claim is that *the victim's dying* would have been different if there had been just one bullet penetration instead of two.

Firstly, different penetrations in the same body modify it differently, and leave it differently wounded. Thus, *the cell system* that undergoes further a progressive diminishing of its vital functions to the death of victim is not identical across the worlds.

Secondly, different deadly wounds induce different profiles of dying. Thus, there are differences in *the succession* of loosing vital functions, in *the timings* of diminishing

CEU eTD Collection

vitality, and in *the durations* of transiting from *being deadly wounded, but alive* to *being dead*.

Plausibly, what gives us the wrong temptation of assimilating the counterfactual deaths with the actual one is our focusing on *the death of the victim understood as a person, not as a body of that person.*¹⁴⁷ Of course, different penetrations in the person's body do not modify the person. Personal identity seems not to be affected by the body's being wounded. So, the person that undergoes the death is identical across the worlds. More than that, in this acceptance, in the temporal structure of the event features only two states: the state when the person is alive, and the state when the person is dead.

Therefore, if it is considered the event happening to the person as such, *the victim's dying* seems to be the same event across the possible worlds. However, on second thought, one may realize that, probably, the death of victim would have come later than it actually came, had only b_1 hit its target. Thus, admitting that the timing of an event is relevant for its identity, one is justified in insisting that the counterfactual death having as only cause b_1 's penetration is different from the actual death.

Finally, it is true that in this case there is a threshold function, but it does not map causes to effects. It is not events that are causally connected in virtue of some biological law connecting instances of *hitting a vital organ* with instances with

¹⁴⁷ Probably, the occurrence of *the death of the body of a person* involves the occurrence of *the death of that person*.

instances of *being dead*. Thus, the threshold function maps the final states of the event-causes to the final state of the event-effect.

Braddon-Mitchell and Jackson (1996) almost make a similar point, except that they challenge that, in the counterfactual situation evoked by the antecedents of the counterfactuals, the *complete effect* would have been the same. I understand that Braddon-Mitchell and Jackson do not take to be aberrant the overdetermination of an effect by some causes, but rather a case of full overdetermination of *everything* that happens as the effect of those causes. Thus, they argue that, in the case of the man killed by two bullets, among what is determined by the two causes, just some – like the death of the victim –, not all effects are overdetermined. The authors speak about things as the paths to the vital organ, the magnitude of the entry holes, and the exit wounds that all would have been different had there been just one bullet hitting the target. Thus, according to Braddon-Mitchell and Jackson, this is just a case of "partial overdetermination":

[T]ypically the case¹⁴⁸ will not be one of full overdetermination, but rather of partial overdetermination. Perhaps each bullet by itself would kill [the victim], but together they do the job more quickly. Then the death is overdetermined, but the timing of the death is not; or perhaps the time of the death would be just the same, but the damage of the surrounding organs would have different, in which case the death is overdetermined, but the damage to the surrounding organs is not; and so on and so forth.

The point is that full overdetermination of everything that happens by distinct factors is very hard to come by. The very fact that there are two bullets means that some things – the paths to the organ, the magnitude of the entry hole or holes, the exit wound or whatever – would have been different had there been just one bullet. (p. 11)

¹⁴⁸ In fact, authors speak about the case when the two bullets enter the victim's body at different points but each ends up in his heart, whereas I consider the case when each ends up in the heart and, respectively, in the brain of the victim. However, the worry arose by authors is even more relevant for the heart-brain case.

The argument seems to be that, even though *e* is an effect to whose occurrence participate both c^* and c^{**} , each of them being a sufficient cause of *e*, if c^* without c^{**} causes e^* , and c^{**} without c^* causes e^{**} , this is not a case of full overdetermination. (See Figure 6.8). Let us consider the following counterfactuals:

- c* occurs without c** □→ c* has the same sufficient causal role in bringing about e and e*
- c^{**} occurs without c^{*} □→ c^{**} has the same sufficient causal role in bringing about e and e^{**}

If these counterfactuals are each non-vacuously true, it amounts to saying that e is overdetermined of e. But this is not to say that we deal with a case of full overdetermination, because neither e^* , nor e^{**} is overdetermined.



Fig. 6.8

6.9 An argument for rejecting any case of overdetermination

In the light of the above considerations, one has to admit that the putative examples of overdetermination have to face considerable objections. In fact, I strongly suspect that there is no overdetermination. Not even one! However, in order to ground this suspicion, it is not enough to challenge the proposed examples of overdetermination, not mentioning that the challenge relies upon a certain criterion of identity coming from a not yet fully-worked out theory of events.

Therefore, in what follows, I shall advance briefly an argument intended to challenge the very concept of overdetermination. Let us consider the following case of causation: e is an effect to whose occurrence participate both c^* and c^{**} , each of them being a sufficient cause of e, and each of them belonging to a separate causal chain. (See Figure 6.9.) *Prima facie*, this seems to be a case of overdetermination by the book.



So, for the sake of the argument, let us assume that *e* is overdetermined by c^* and c^{**} . Further, it is perfectly conceivable a world where c^* and c^{**} have the same cause. (See Figure 6.10.) But, the argument proceeds, if *e* is overdetermined in the former case, it should be overdetermined by the same causes in the latter case. At

this point, the troublesome observation is that our world is full of cases similar with that illustrated in Figure 6.10. For one thing, many times, in the natural world, an event, c, causally sufficient for some effect, e, is composed of events, each of them being causally sufficient for the same effect; the part-events, c^* and c^{**} , are both caused by the same cause that brought about the whole-event, c. Thus, the initial assumption that e is overdetermined leads us to the unacceptable conclusion that the cases of overdetermination are not exceptional, but rather abundant. Therefore, we have to reject the initial assumption, namely, that the case of causation from Figure 6.9 is an example of overdetermination. Consequently, claim that there is overdetermination becomes quite implausible.

Let us consider again, but in a changed version, the notorious example of the man killed by two assassins: bullets, b_1 and b_2 , hit in the same point, in the victim's head, say. Of course, one would not meet too much resistance if one takes this to be a case of overdetermination, i.e. *the victim's dying* is overdetermined by b_1 's *deadly penetrating in the victim's brain* and b_2 's *deadly penetrating in the victim's brain*. Now, I assume that I do not meet too much resistance either if I claim that, given the above case, it is perfectly possible a case when the same man is killed by a single assassin shooting one single bullet in the victim's *dying*, which very probably is transworld identical with the death from the former case, would be caused by *b's penetrating in the victim's brain*. On the assumption that *b's penetrating* is an event composed of b_1 's *penetrating* and b_2 's *penetrating*, ¹⁴⁹ it follows that if *the victim's dying* is

¹⁴⁹ An alternative version would be one where b splits in two pieces, b_1 and b_2 , when it starts to penetrate the victim's head.

overdetermined in the former case, then it is overdetermined in the latter as well. But let us take the things at their real value: in the latter case, the man is killed by a bullet that hits his head. It would be too much to say that the death is overdetermined.

6.10 Rejecting widespread and systematic overdetermination

In §2.3, I have mentioned that the inconsistency of the CEP assertions is claimed to stem from their allegedly entailing a *highly implausible situation*: each and every bit of neurological changes whose causes are mental is overdermined. Further, in § 2.5, characterizing such a situation as being *widespread and systematic overdetermination*,¹⁵⁰ Almost everybody would agree that any account of mental causation is more preferable to one requiring a commitment to a widespread and systematic overdetermination.

Although the above view is almost unanimous and, in § 2.5, I have already provides some reason for it, I believe that, at this point, a more detailed examination could be advance. Thus, I shall devote this section to the question of what should motivate people to reject a systematic and widespread overdetermination.

Let us begin by considering what Kim calls 'the principle of explanatory exclusion' (briefly, EE).¹⁵¹ Going through Kim's essays on mental causation, one meets various

¹⁵⁰ Recall that phrase 'systematic overdetermination' refers a situation whose characteristic is that, whenever some kind of causal relations hold, a case of overdetermination happens. By 'widespread and systematic overdetermination' it is meant a systematic overdetermination whose causal relations systematically involving overdetermination occur with a considerable frequency and in a large number of cases.

¹⁵¹ See, for instance, Kim (1993a [1989]: 239). Kim's terminology varies through his writings. Sometimes he chooses instead 'causal exclusion' – see Kim (1993d: 360) –, while in other occasions he prefers speaking of 'causal-explanatory exclusion' – see Kim (1993a [1991], 1993c: 281).

expressions of EE, all of them conveying the same idea that, given an event, it

cannot have two complete and independent explanations. For example, in Kim

(1993a [1989]):

No event can be given more than one complete and independent explanation. (p. 239)

Also:

Two or more complete and independent explanations of the same event or phenomenon cannot coexist. (p. 250)

Also:

No one may accept both explanations unless one has an appropriate account of how they are related each other. (*Op. cit.*, p. 257)

As far as I can see, there are three Kimian arguments for EE. The first one plays on

the card of the role any explanation is supposed to play, namely that of getting us out

of a state of "epistemic predicament":

When we look for an explanation of an event, we are typically in a state of puzzlement, a kind of epistemic predicament. A successful explanation will get us out of this state. ...[T]oo many explanations will put us right back into a similar epistemic predicament, which can be relieved only when we have an explanation of how the explanations are related to one another that epistemic predicament is relieved. (*Op. cit.*, p. 254)

But, the argument goes, the relief of any epistemic predicament is achieved through

understanding, and explanation is able to provide understanding through the

unification and simplification of the assumptions.

Unity is achieved through the promotion of interconnections among items of knowledge, and simplicity is enhanced when these interconnections are seen or interpreted as dependency relations. For the main role of dependency relations in a system is that they help reduce the number of required independent assumptions or primitives. (*Ibidem*)

Thus, considering an event that has two explanations, the only way to put us in a

state of epistemic comfort regarding the occurrence of the event is to find out how the

assumptions of one explanation are dependent on the assumptions of the other

explanation such that "the two explanations taken together are committed to no more independent assumptions about the world than is the second explanation taken alone." (*Op. cit*, p. 255)

If the first argument for the Kimian principle of explanatory exclusion call our attention on the fact that the existence of two complete and independent explanations of the same event would keep us in a sort of epistemic predicament, the second Kimian argument emphasizes that those two explanations would endanger the coherence of our knowledge network:

Accepting the two sets of premises as constituting explanations of the same event (or any one thing), each complete in itself and independent of the other, may induce a sort of incoherence into our belief system. (*Op. cit.*, p. 258)

Finally, an interesting argument is also the following. EE is a prolongation of Ockam's Razor. Event though EE relies somehow on Ockham's razor, EE takes this abstract aspiration to simplicity further and makes it a concrete requirement. Ockham's razor tells us that *we should avoid having excess baggage* by multiplying entities without need, whereas EE goes deeper, telling us, in the particular matter of explaining, *what would be the excess baggage*:

The exclusion principle does the difficult work: it says that for any event more than one complete explanation is excess baggage. (...) We can indeed think of explanatory exclusion as a special case falling under the general simplicity requirement: it is a specific rule concerning one important way in which simplicity is to be gained in explanatory matters, and it explains why this form of simplicity is to be desired. That is, the explanatory exclusion principle provides a rationale for the application of Ockham's Razor to multiple explanations of a single explanatorum. (*Op. cit*, p. 260)

It should be noted that none of the above arguments invokes a particular nature of explanation. However, it seems to me that, in order to draw from EE a justified attitude toward overdetermination, it should be considered EE from the particular perspective of *causal explanations*.

Actually, for many philosophers, to give an explanation comes to give a causal explanation. For instance, the main thesis in Lewis (1986a) is that "to explain an event is to provide some information about its causal history" (p. 217).

But, if one wants to translate EE in terms of causal explanations, one should make explicit the meaning of some phrases like 'complete explanation' and 'independent explanations'. Lewis, for instance, is not of much help here, because he equates a complete explanation of an event with its "whole explanation":

Among the true propositions about the causal history of an event, one is maximal in strength. It is the whole truth on the subject – the biggest chunk of explanatory information that is free of error. It is, of course, very unlikely that so much explanatory information ever could be known, or conveyed to anyone in some tremendous act of explaining! (*Op. cit.*, pp. 218-9)

Nor Kim does provide a straight answer, but he seems to let us understand that a complete explanation of the event is one that mentions a causally sufficient condition abstracted from the causal history of the event.

Now, if an event, *e*, has two causal explanations, E^* and E^{**} , what would be for them to be independent one from another? Arguably, their independence comes to the independence of their truth-conditions. Thus, we may say that E^* is an explanation of *e* independently from E^{**} iff E^* is true about *e* independently from E^{**} in the sense that the truth of E^* does not lead to the truth of E^{**} .¹⁵²

¹⁵² Another way of saying the same thing is: E^* is an explanation of *e* independently from E^{**} iff E^* would still explain *e* even if E^{**} failed to explain *e*.

For simplicity, let us consider that each of E^* and E^{**} mentions only one sufficient cause of *e*. Let them be c^* , and c^{**} , respectively. E^* loses its independence from E^{**} in explaining *e* when c^{**} is a cause of c^* . The suggestion is that E^* and E^{**} are independent one from another only when there is no causal chain going through c^{**} , c^* , and *e*. (That is, only when each cause arrives to the effect through a different causal chain.)

Given that principle EE excludes the coexistence of E^* and E^{**} as two complete and independent causal explanations of e, it follows that there is not such a case that c^* and c^{**} is each a sufficient cause of e and no causal relation holds between c^* and c^{**} . In conclusion, EE seems to oppose to the existence of an event that, in accordance with any of the considered tests, is eligible for being considered overdetermined.

However, it is puzzling to note that, against what EE appears to rule out, Kim still admits, in a number of cases, the existence of some events having two unrelated sufficient causes. Actually, Kim is ready to give credit to the idea that overdeterminative cases are an exception to EE, exempting them from the requirement of explanatory exclusion.¹⁵³ But this stance leaves us in a greater predicament than any epistemic predicament that would justify our commitment to EE. If EE excludes overdetermination, how could one be at the same time both committed to EE and indulgent with the existence of some overdeterminative cases?

¹⁵³ See, for instance, Kim (1993a [1989]: 253).

It might be a way out if we consider Kim's speculation that even though E^* and E^{**} mention each a sufficient cause of *e*, none of them is a complete explanation of *e*:

[I]t is not implausible to think that failing to mention either of the overdetermining causes gives a misleading and incomplete picture of what happened, and that both causes should figure in any *complete* explanation of the event. If this is right, the present case is not one in which there are two complete and independent explanations. (*Op. cit.*, p. 252)

Yet, it appears to be an *ad-hoc* consideration, changing the meaning of 'complete explanation' whenever it is needed in accordance with the author's needs. Moreover, it opens a never ending looking for "a truly complete explanation". Probably, in this line, the only complete explanation of an event would be the untouchable Lewisian "whole explanation".

Now, if Kim, as many others, admits the existence of overdeterminative cases, what seems to bother him is a picture of the world massively populated with overdeterminative cases. Thus, Kim uses EE in order to argue that, even though there are overdeterminative cases, they are not abound, but rather quite rare and exceptional.

I believe that what motivates people to reject massive overdetermination in a given world or, at least, in our world is an assumption of rationality and intelligibility: *The world is such that its nature (its constituents and its constitution) is according to "our" principles of rationality; else it would be unintelligible for us.* Nobody would admit a metaphysical theory that delivers us an irrational picture of the world. By 'our principles of rationality', it should not be understood some human-mind-dependent principles, but rather some objective principles governing both our thinking and the constitution of the world. The commitment to the existence of *widespread and systematic overdetermination* would require one to admit a picture of a world devoid of elegance, grace, balance, shapeliness, and simplicity. Note that the talk is not primarily made here from an aesthetic point of view. Rather, it is the violation of some basic requirements of rationality that entitles us to blame the picture of a world massively overdetermined. As it has been often noted, if such a world were the result of someone's design, one would be justified in describing the design as being deficient. For one thing, the designer would wrongly oversize the whole thing by making the world to spend more resources than it needs for its becoming.

But, if we feel entitled to reject the idea of massive overdetermination as the expression of a deficient design, why don't we apply the same reason when it comes to whether there are any (even though exceptional) overdeterminative cases? A design having fewer deficiencies is still a bad design.

If one holds the irrationality of having overdeterminative cases all over the places, one cannot ground his position by saying: 'It is not overdeterminative cases, but rather their abundance that is not alright.' Instead, a strong reason for discarding a certain metaphysical theory is that it pictures a world where there is one and only one category of events that are always overdetermined, whereas the rest of events are only exceptionally overdetermined.

Turning now to CEP, it builds its credibility on the claim that the commitment to the CEP assertions has the unwanted consequence that any neural event that has

CEU eTD Collection

mental causes would be overdetermined. Even though one accepted that sometimes events are overdetermined, one would not rationally afford to buy the idea that all mental causation is done by overspending causal resources.

Chapter 7: Beating the CEP Argument

This final chapter is thought as an answer to the argument from causal exclusion problem against interactionism (briefly, the CEP argument) expanded in Chapter 2. The whole train of thought is designed to follow a zigzag route. The reason is that, as I have pointed out in § 2.4 and § 2.7, the CEP argument may receive more than one version, in accordance to what reading is taken for the notion of physical closure.¹⁵⁴ That is why the progression through the chapter will be made by taking alternatively one step forth on the basis of a certain reading of physical closure, and one step back for reconsidering the CEP argument in the light of another reading. Thus, progressively, from the weakest to the strongest, each possible reading one may choose for the notion of physical closure will be tried in the fire of analysis ('What are its credentials?'), the corresponding version of the CEP argument remaining to be answered each time.

The approach that I shall take here is methodologically similar with that displayed by *E. J. Lowe in his 'Causal Closure Principles and Emergentism' (2000a). Lowe argues that "there are various forms of naturalistic dualism, of an emergentist character, which are perfectly consistent with the strongest physical causal closure principles [i.e., in my usage, the strongest readings of the notion of physical closure] that can plausibly be advocated."¹⁵⁵ (pp. 572-3) Some of these forms of naturalistic dualism*

¹⁵⁴ Regarding the large variety of formulations of the notion of physical closure to be found in literature, Lowe rightly remarks that "one might have hoped for more exactitude and agreement amongst physicalist when it comes to the formulation of a principle so central to their position." (2000a: 574) ¹⁵⁵ Lowe use 'causal closure' to refer to what I take to be physical closure. Instead, I reserve 'causal closure' for referring to a certain reading of the notion of physical closure. See § 7.4.

are on the same track as the compatibilist answers (scenarios) that I shall propose in the following. However, our approaches differ over what reasons would entitle one to decide whether a compatibilist situation (scenario) has been found.

7.1 General strategy

Let us recall, in short, the CEP argument (see § 2.5). The main theses of interactionism imply two of the CEP assertions. Also, apparently, the conjunction of the CEP assertions implies further that in any case of mental causation the physical effect is overdetermined. Given that widespread and systematic overdetermination is unacceptable (see § 6.10), the CEP assertions cannot all be true. There are some reasons for holding the truth of the third CEP assertion. Therefore, the CEP assertions implied by interactionism cannot all be true. It follows that interactionism is false.

According to § 2.1, the interactionist theses are:

(D) Dualist thesis: Mental properties are not token-identical with physical properties;

(C) Causal thesis: Mental phenomena and physical phenomena can causally affect one another.

According to § 2.3, the CEP assertion (i.e. the assertions of the causal exclusion problem) are the following:

(NI) Mind-body causal non-identity: Mental causes are not identical with physical causes;

(MC) Mental causation: The mental is able to cause physical occurrences;

(PC) Physical closure,

where (PC) involves at least that, for any physical occurrence, there was a physical cause or some set of physical causes sufficient for bringing about the respective occurrence.

It may be helpful to set out the foregoing argument in a rather more formal fashion,

each numbered line constituting a step in the proof.

- 1. (D) & (C) implies (NI) & (MC)
- 2. (NI) & (MC) & (PC) implies the overdetermination of any physical effect of mental causation
- 3. Overdetermination is not a systematic and widespread phenomenon in nature.

Therefore,

- 4. (NI) & (MC) & (PC) is false.
- 5. (PC) is true.

Therefore,

6. (NI) & (MC) is false.

Therefore,

7. (D) & (C) is false; that is, interactionism is false.

It seems fair to say that the only possible flaw of the above argument is the logical implication stated in line (2). Moreover, I believe that line (2) really is a fatal flaw in the argument, and the natural way of proving that is to set up at least a compatibilist scenario, as I shall call it. By definition, in such a scenario, if it were actual, conjunction (NI) & (MC) & (PC) would not involve the overdetermination of the

physical effect. Without a guarantee for the truth of line (2) there is no guarantee for the soundness of the CEP argument; hence it cannot play the role of a conclusive argument against interactionism.

It is very important to note, though, that, by finding a compatibilist scenario, it is not enough for proving beyond reasonable doubt that (2) is false. The scenario would do this job if either:

- there were substantial reasons to believe that the scenario is according to reality, that is, there are some cases of mental causation where the relationship among the protagonists of causation really follow the pattern of the suggested scenario;

or:

- (2) were a "strict implication"¹⁵⁶, in other words, overdetermination necessarily follows from (NI) & (MC) & (PC).

Neither is the case. But the scenario does cast doubt on the truth of (2) and this is all is needed for my purposes. To show that the scenario is only fictional a separate argument is needed. There is no way of saying whether or not it is according to reality just by looking at it.

On the other hand, in reply, a defender of the CEP argument may require an argument for the *possibility* of the respective scenario, that is, a rationale for holding that there is at least a possible world where the scenario reflects the reality. This is

¹⁵⁶ Clarence Irving Lewis is the one who introduced the notion of strict implication. Strict implication was defined as ~ 0 ($p \square \sim q$), in which 0 means "is possible" or "is not self-contradictory." See, for example, his 'The Calculus of Strict Implication" in Mind 23.2 (April 1913): 240-247. Also, it is said that p strictly implies q if and only if it is necessary that p implies q.

not an easy job! However, as we will see, I am going to work it out by giving some examples that shed a light of plausibility on the respective compatibilist scenario.

For setting up a compatibilist scenario in a framework where conjunction (NI) & (MC) & (PC) is true, I shall follow a strategy already hinted in § 2.6. The answering strategy is based on finding a necessary condition on overdetermination, (NC), and some relation, R, such that, if R held among some of the causes of the physical effect, (NC) would be violated.

In Chapter 6 I have argued that an effect, e, is overdetermined by two or more of its causes, c_1 , c_2 , ..., only if **(CNC)**:

- (1) there are two processes, ϕ and Π , each of them being causally sufficient for *e*, such that some of c_1 , c_2 , c_3 , ... feature among the originating causes of the processes;
- (2) the following counterfactuals are non-vacuously true:

(γ_1) if Π 's nodes and their causal ancestors were the only causes of *e*, at least one of the properties of *e*'s direct causes from [$\Phi - \Pi$] being uninstantiated on the occasion of *e*'s occurrence, then *e* would occur in virtue of L_{Π} s's activation; (γ_2) if Φ 's nodes and their causal ancestors were the only causes of *e*, at least one of the properties of *e*'s direct causes from [$\Pi - \Phi$] being uninstantiated on the occasion of *e*'s occurrence, then *e* would occur in virtue of L_{ϕ} s's activation.

(3) Φ and Π do not stand in a part-whole relationship.

I shall take into account two possible working frameworks for mental causation. To imagine a compatibilist scenario that would fit in any of them, all is needed is to

CEU eTD Collection

conceive a relation, R, holding among some causes of e such that condition (CNC) is not met. Thus, the claim that I intend to defend in this chapter is the following:

(S) Even though conjunction (NI) & (MC) & (PC) is true, there may be set up at least one compatibilist scenario where R is composed of causal relations going from the causally sufficient condition of physical nature to the causally sufficient condition containing mental causes.

If (S) is right, it sheds a shadow on the CEP argument, because (S) raises the question whether or not line (2) is true. For one thing, if (S) is right, and R is actually exemplified, the logical implication from line (2) is false, and the CEP argument is unsound. In short, according to (S), the possibility of R makes the CEP argument by itself to be inconclusive. That is to say that, once (S) is proved, the CEP argument appear to in a serious need of a guarantee that R does not hold in the actual world.

7.2 The notion of physical closure in a weak reading (PC_w)

The physical closure appears to be a largely shared metaphysical image. Let us consider this view in the enunciation of Corbi and Prades (2000):

Causation cannot be a matter of magic; therefore, it should be taken for granted that, for any causal chain that affects the physical world, there must be a physical process that can explain the effect. (...) [T]here is a complete causal explanation... that does not include any mental predicates. (pp. 5-6)

In the above quotation, leaving aside its being introduced by a fallacy,¹⁵⁷ the assertion of the physical closure is not primarily a metaphysical one, but rather it reveals itself as claiming *the causal explanatory closure of the physical domain*. Generally speaking, an ontological domain is closed in the causal explanatory sense if and only if, given any change in the domain, it has a complete causal explanation in terms referring only to entities belonging to the domain. Thus, we get the principle of physical closure in what I claim to be its weak reading:¹⁵⁸

(**PC**_w) For any physical effect there is a complete causal explanation made only in physical terms.¹⁵⁹

¹⁵⁷ The above argument is an example of a false dilemma, a fallacy known also as "false choice". Two mutually exclusive alternatives are wrongly held to be the only possible options: *Given a physical effect, either it has been occurred by miracle, or its occurrence is explicable through a physical explanation.* Then, the falsity of the first alternative being obvious, it is concluded in a fallacious manner that *any physical effect should have a physical explanation.*

True, any physical effect should have an explanation; it is not rationally acceptable that it has been brought by miracle. But the requirement of an explanation does not imply by itself the requirement of a physical explanation. The implication were true only if any other kind of explanation would be an appeal to magic. But it is not settled that an explanation including non-physical terms (i.e. an explanation not made exclusively in physical terms) is not an explanation alright.

¹⁵⁸ As we go further in this section, it will become evident that it is the same reading that I have used for (PC) in § 2.7 in order to state the CEP argument in what I have called there "its most liberal version".

¹⁵⁹ It may be noted that the above enunciation is not exactly the same as that advanced by Corbi and Prades. They take a physical causal explanation to one that "does not include any mental predicates", whereas, in the above enunciation, by 'physical causal explanation', it is meant a causal explanation made only in physical terms.

I hold that their enunciation has the following difficulty. Arguably, Corbi and Prades assume a mental-physical distinction based on the theory approach, according to which physical terms, unlike mental ones, are used in physical theories, i.e. those scientific theories advanced in such domains as

But, the question is: in what sense should the required physical causal explanation be complete? Firstly, a complete causal explanation here should not be taken for "the whole explanation", how David Lewis calls it. The latter is supposed to provide ideally all true propositions about the causal history of the physical effect.¹⁶⁰ Thus, if this were the kind of physical explanation required by (PC_w) for any physical effect, it would mean that it is assumed that its causal history never goes out of the physical domain. However, by using such an assumption in an argument for the conclusion that the mental phenomena, if not included in the physical domain, cannot causally affect any physical phenomena, one begs the question against interactionism.¹⁶¹

Secondly, a complete causal explanation is one that assures an understanding of why an effect happens. Such an understanding is obtained as soon as a sufficient reason is found out. But a condition causally sufficient for an effect stands as a reason enough for the occurrence of that effect.¹⁶²

microphysics, physical chemistry, neurophysiology, medicine and the like. However, as I have shown in § 1.3, there is a pretty convincing argument for dismissing the theory approach: it is not impossible that the future physical theories, especially micro-physics, will posit entities of a mental kind. ¹⁶⁰ See Lewis (1986a: 218-9).

¹⁶¹ Consider following quotation from Corbi and Prades (2000):

[&]quot;Perhaps we do not know the entire physical story, but we assume that it must exist. (...) The dominant view assumes that, if the causal story [the scientists] told us were not a complete one, a completion of it would require additional statements in which no mental term could properly appear." (p. 6)

It may be read in two ways. Prima facie at least, it may be understood that Corbi and Predes ascribe to the dominant view the assumption that, even though the scientists do not know *the entire causal story* of a physical effect, that story is in principle free of mental causes. That kind of principle cannot be used in an argument against interactionism on pain of begging the question.

Yet, there is one more possible reading here: the assumption is that in principle there is a *complete* physical story. I shall discuss this "mild" reading of the physical closure (PC_M) in § 7.4

¹⁶² If any sufficient cause seems to play the role of a sufficient reason, the other way round does not go: no all sufficient reasons are sufficient causes.

In conclusion, a *complete causal explanation* of an effect is one that provides at least a condition causally sufficient for the effect. It should not be mistaken for a *complete metaphysical account* of the effect, that is supposed to make known all the direct interactions in which the explanandum is involved.¹⁶³ Again, if the principle of physical closure understood as (PC_W) requires for any physical effect a complete metaphysical account solely appealing to an underlying physical process, (PC_W) implies falsity of interactionism. Then, the assumption of that principle in the CEP argument is begging the question against interactionism.

In conclusion, the weak sense of the notion of physical closure is:

 (PC_w) For any physical effect, there is a causally sufficient condition in the physical domain.

For the time being, let us approach the CEP argument on the assumption that (PC) is taken in the weak sense.

¹⁶³ Question: if there is a genuine distinction between a complete causal explanation and the complete metaphysical account of an event, then, relative to the understanding provided by the former, what is it more that the latter brings to understanding? One may claim that what you understand from a complete physical causal explanation about *what is going on with a neuron* – an explanation that ignores the neuron's association with the consciousness – is the same as what you understand about *the behavior of the neuron* from a complete metaphysical account, even though it takes into account the neuron's association with the consciousness. Regarding what is going on with the neuron, the understanding is the same, but the metaphysical account provide also a picture of *what is going on in the "environment" where the neuron exhibit its behavior*.

7.3 A simplified framework for designing compatibilist scenarios

In this section, in accordance with the truth of conjunction (NI) & (MC) & (PC), I shall sketch the first working framework from a series of three altogether. Along the way, the designing of compatibilist scenarios will be undertaken in each framework. However, progressively and in parallel, the stronger the reading of (PC) becomes, the more restrictive for accommodating a compatibilist scenario the framework becomes.

If (MC) is true, it is all right to consider a physical event, e, and a mental event, m, such that m is a direct cause of e. For simplicity, let us assume that m is the only mental cause of e. For this time, let us consider that m is a sufficient cause of e. Even it is not the most plausible scenario, it is still in accordance with (MC) and has on its side the crucial advantage of simplicity. Let us assume further that the causal sufficiency of m is grounded on a fundamental law, L_{ϕ} , activated by the exemplification of property M, and necessitating the exemplification of property E, where M and E are the final properties exemplified on the occasion of m, and e, respectively.

Further, by obeying to (PC_W), within the physical domain, there must be *p*, a causally sufficient condition on *e*. For simplicity, let us assume that *p* is constituted of only one event. Let us call it exactly *p*. Then, if (NI) is true, *e* has two distinct sufficient causes, namely *m* and *p*, such that the exemplification of *M* is not realized by the exemplification of *P*, the final property reached on the occasion of event *p*. (See Figure 7.1)



Now, the relevant question is: Do causes *m* and *p* overdetermine *e*, no matter what? I shall try to defend the claim that, even though *m* and *p* are each causally sufficient on *e*, the overdetermination of *e* is not unavoidable. Let us begin by noting that, if there are two processes, Φ and Π , whose originating causes are among *m*, and *p*, Φ and Π being each causally sufficient for *e* such that Φ and Π do not stand in a part-whole relationship, then the only chance for (CNC) to be met is that their originating causes are *m* and, respectively, *p*.

In the given framework, let us assume that *R* is exactly a causal relation from *p* to *m* such that *p* is a sufficient cause of *m*. Therefore, process Φ cannot be but <m; -;*e*>. Instead, process Π can be assigned with more than one linking network: Π is <p; ?; *e*>.

Then, particularizing (CNC), the physical effect, e, is overdetermined by m and p only if:
(CNC₁) <m; -;e> and <p; ?; e> do not stand in a part-whole relationship, and

(CNC₂) the following counterfactual is non-vacuously true:

(γ) *e* would still occur in virtue of the same laws that actually support process <p; ?; *e*>, even if the nodes of <p; ?; *e*> and their causal ancestors were the only causes of *e*, property *M* being uninstantiated in the case of *m* not being a node of <p; ?; *e*>.¹⁶⁴

Thus, within the present framework, under the assumption that p is causally sufficient for m, the claim that I advocate come to say:

(S) Even though conjunction (NI) & (MC) & (PC_w) is true, it may be set up a scenario in which e is not overdetermined.

7.3.1 A "candid" scenario

Within the working framework, given (CNC₁), it should not be surprising that the first compatibilist scenario one comes up is what we may call a "candid" scenario. In that hypothetical situation, the laws supporting process Π are L_{Π} and L_{ϕ} , the former being triggered by the exemplification of *P*, and the latter being triggered by the exemplification of *M*. (See Figure 7.2_2.)

¹⁶⁴ Given the present framework, (γ) is exactly a particularization of (γ_1). (γ_2) is uninteresting, given that, by construction, it is trivially true.



It follows that process Π is $\langle p; m; e \rangle$. Then, $\langle m; -; e \rangle$, i.e. process Φ , is a serial part of Π , which is a violation of (CNC₁). (See Figure 7.2_1) Therefore, *e* is not overdetermined by *m* and *p* in a hypothetical situation described by the present scenario. This supports the truth of (S), the claim intended to undermine the CEP argument.

The conclusion is that, by assuming the existence of a mental cause of a physical effect, even though the mental is distinct from the physical, the physical world remains explanatorily closed in the weak sense, and the effect is not overdetermined, provided the mental cause is produced itself by a physical source.

Mentioning a version of physical closure very close to (PC_W) , namely that every physical event which has a cause has a sufficient cause, Lowe advance a solution

CEU eTD Collection

identical with a "candid" scenario.¹⁶⁵ He appreciates that this particular version of physical closure can be integrated by an interactionist doctrine of emergentism kind:

More generally, an interactionist dualist who espouses some version of emergentism can happily endorse principle (1F) [i.e. (PC_W)]. Such a dualist may consistently maintain that the universe has evolved (without 'outside' influence of supernatural powers) from a condition in which only physical events existed to one in which both physical and mental events exist – the latter conceived as non-physical – while also espousing causal determinism for all events. On such a view, if we trace back the causal history of each mental event, we eventually reach a time at which all of its causal antecedents were wholly physical events, because on this view it is the physical which has ultimately brought the non-physical real of the mental into being. (p. 576)

7.4 The notion of physical closure in a mild reading (PC_M)

It is time to move on to a stronger reading of the notion of physical closure. One might take the following stance: 'Given a physical effect, it is reasonable to assume that it has a causally sufficient condition in the physical domain. But, the closure of the physical requires more than that. Besides the sufficient causal condition and its effect, there should be also some *causal linkage* that is physical too. A complete explanation is one that, besides a condition causally sufficient for the effect to be explained, provides what makes the causal connection. If the explanation is physical, the connection should also be physical.' More or less, this view is illustrated in Corbi and Prades (2000):

Prima facie, it seems that there is a chain of physical causation that is enough to produce the physical effect and that does not require the instantiation of any mental property. (...) [E]veryone may be strongly inclined to concede that a physical mechanism must underlie any change that the world actually undergoes. (p. 6)

In the above reading, the notion of physical closure means that the physical domain is self-sufficient in providing a causal process sufficient for any of its changes. In this

¹⁶⁵ See Lowe (2000: 575-6).

view, the closure of the physical domain is met only if no intermediaries of other nature than physical are needed to mediate the causation of any physical change.

Since it is stronger than (PC_W), but not the strongest one, I shall call this requirement of homogeneity for the causal story of the physical the *mild reading* of the assertion of physical closure:

 (PC_M) For any physical event, there is a physical causally sufficient condition leading to the effect through a physical causal process.

It is important to keep in mind the difference between the two meanings of the notion of closure used respectively in (PC_w) and (PC_M). In short, for any *x*, a change in a domain, *D*, the former closure requirement speaks about the existence of *y*, a sufficient cause for *x*, whereas the latter speaks additionally about the existence of a causal route from *y* to *x* that does not go out of *D*. (See Figure 7.3.) Moreover, it is important not to take what (PC_M) implies, namely, that there is a causal route that does not need to go out of *D*, for what something stronger than (PC_M) implies, namely that there is no causal route going out of *D*, that links *y* with *x*¹⁶⁶.

 $^{^{166}}$ In § 7.6, it is discussed (PC_S), a stronger requirement than (PC_M).



Now, it is clear that, in the scenario presented in § 7.3.1, conjunction (NI) & (MC) & (PC_M) is not true. For one thing, even though its originating causes make up a condition causally sufficient for *e*, process Π is going out of the physical domain before reaching *e*.

However, truth be told, I find (PC_M) a quite excessive claim to be made. Given that a commitment to (PC_M) blocks a natural scenario like that just presented in § 7.3.1, the following questions arise from a justified frustration. What reasons do we have to believe that (PC_M) is true? How does it back up the view that, besides a causally sufficient condition for any physical event, the physical domain provides always some physical causal process leading to such an event? In the rest of this section, I am

going to expand on what "aromatizes" (PC_M), and makes it so appealing for some of us, by trying to show that, without its apparent argumentative aromas, it is nothing else than a gratuitous constraint in the problem of mental causation.

It is the nature of the current scientific theories that is usually invoked in defending (PC_M) . Thus, the standard answer would go like this: 'The causal explanations given by various theories in domains as microphysics, physical chemistry, neurophysiology, medicine and the like do not involve any event other than physical ones. So, if we do not want to give up of all these sciences, it is reasonable to assume that the causal processes to which scientific explanations appeal are exclusively physical. But then, if all around us there are only (scientific) domains in which causal processes are physically homogeneous, how reasonable would it be to believe that, in bringing about our neural or behavior events, the causal processes originated in some neural causes are sometimes mediated by non-physical events, as dualists claim mental events to be?'

According to the above argument, (PC_M) follows from two premises. In one of the premises it is assumed the mild closure of various sectors of reality. Those sectors are delineated by scientific theories explaining physical phenomena. In the other premise it is made a commitment to the homogeneity of causation, which is, of course, the triumphalistic materialization of the natural, but not always justified, reflex of inductive (and extrapolative) generalization. For convenience, let us focus on the following expressions of the mentioned assumptions:

 (PC_{MS}) : Causation among the physical events within the scope of any scientific theory of matter does not involve the mediation of any event other than physical ones.

Causal Homogeneity: Causation among physical events is the same in the whole physical domain.

Let us consider further a possible line of argumentation for (PC_{MS}): 'Our scientific theories of matter provide causal explanations using only material predicates. Therefore, to say that, in some cases, the causal processes referred by those explanations contain non-physical events is to say that many of our causal explanation given by science are misleading. But it hardly acceptable that scientific research is held captive to such regrettable error.'

It seems to me that the above route of argumentation is fallacious. Firstly, it is true that, in order to causally explain observable physical phenomena, our current scientific theories do employ only physical events. But, there is no reason for believing that the future causal explanation will be always like that. It is gratuitous to assume that no future theories will truly explain certain observable physical events by appeal to non-physical causes as mediators.¹⁶⁷

¹⁶⁷ Recall that we are not working under the hypothesis that any entity posited by a theory of matter is automatically physical as well. I have held that a physical property is any property of the same kind as those to be mentioned in an account of a paradigmatic physical entity. Note that the account in question consists in revealing of the constitutional structure of the physical entity. The constitutional account is different from a causal explanation of a physical event, that, in contrast, has nothing to do with what constitute the event, rather mentioning what is responsible for bringing about the respective constitution.

Secondly, the considered argument makes use of an unsecured assumption, namely that the *causal explanatory closure* – which any science aims at achieve – implies the *causal closure* of the ontological sector posited by the respective scientific theory.¹⁶⁸ Yet, even if the causal *discourse* of a true scientific theory is non-physical term free, its causal explanations involving only physical entities, this is not a warranty that the underlying causal *reality* is not richer than what is needed for meeting the causal explanatory exigencies.

There are at least two requirements for any scientific causal explanation. On the one hand, it should mention as explanatory causes only events for which there could *in principle* be found empirical evidences. Call it the evidential exigency. On the other hand, in its explanans, any explanation should deploy selected data about relevant reality in a comprehensible manner. Call it the intelligibility exigency.

The two exigencies make up a filter active at the "entrance" in the ontology posited by any scientific theory such that, in some cases, the reality is richer then the ontology of the scientific discourse. Besides the clear cases, where the force of evidences requires the introduction in some explanation of certain events or where the multitude of the causal factors requires a selection, there are also many other cases where a middle way between the two exigencies must be found. For example, there are some cases where, even though the existing empirical evidences are weak for positing a certain event, the intelligibility exigency exerts a considerable pressure to make a certain ontological commitment. In short, an entity might be postulated just because

¹⁶⁸ The claim that the causal explanatory closure should not be taken for the causal closure is just a particular application of a more general claim: the explanatory closure should not be mistaken for the nomic closure. I say more about this in § 7.6.

its presumed causal role makes more sense in explaining a certain observable phenomena. Sometimes, instead, against the intelligibility exigency, even though the introduction of a causal factor in an explanation makes fully sense, in the light of evidence, or, better said, given the lack of evidence, the respective positing appears to *ad hoc*.

Without doubt, there are certain limitations regarding our way of comprehending the world and our capacity of collecting evidence about what there is. That is why the world is for us only what is seen through an *epistemic window*. Thus, despite their effective causal role in bringing about the explanandum, our epistemic window makes some entities to be left out from the causal explanations given by any scientific theory.¹⁶⁹

We are bound to explain in a *human* manner. As far as I can see, the limitations that shape our epistemic window to the world are mainly determined by our cognitive resources and capacities. They are crucial for what sort of realities we can conceive and for what kind of patterns are meaningful for us. It could be the case that a certain causal pattern has a degree of complexity for whose understanding one needs to process an amount of data at certain speed that demands a cognitive system more powerful than anyone available to a human being.

¹⁶⁹ Of course, scientific theories might be constrained to leave out some entities not only because of the epistemic window. Sometimes, within the epistemic window, it might intervene further constraints of other nature. One of them is given by Ockam's razor: *Do not load the ontology of your theory with entities whose postulation does not raise the explanatory power of the theory.* It is a concretization of the desiderata of simplicity and elegance that any scientist struggles to achieve by looking for more powerful explanations at a smaller ontological cost.

But we should not take those constraints to be more than a practical guidance in issuing scientific theories; we should not take them for a guidance on how the world really is. After all, what does it ground the believe that our world is arranged and works in a rational manner?

Also, it should be mentioned that our way of understanding and conceiving is rooted in the empirical concepts we got due to our perceptive apparatus. The categories in which we group things encountered in our experiences prove not only our natural tendency of saving cognitive resources, but also the level of fine-tuning at which we are dealing with the empirical reality. Moreover, the way our empirical concepts are tailored biases the meaning of abstract and theoretical concepts.¹⁷⁰

The claim that I try to defend here is that a scientific theory could provide a true picture of the world, it not being also a complete metaphysical picture. It is like in the case of a scene taken from a bird's eye view. Event though one gets the real thing, some details are missed. Metaphorically, you can say that, many times, a scientific picture, unlike a metaphysical one, is like a photography taken at a weaker resolution or like a black-and-white photography. This does not make it untrue. The causal explanation given in a scientific picture can be true, although it provides only a piece of the causal linkage.

The entities and the explanatory picture proposed by any of the current scientific theory are framed by the casement of the epistemic window. But they do not exhaust the whole space to be discovered or disclosed. It is very likely that there are still (many) more experiments to be conducted and more scientific theories to be issued to the effect of revealing the full view to be seen through the epistemic window. Thus, we should make a distinction between *contingently* hidden reality and *in principle*

¹⁷⁰ For instance, in the domain of micro-physics, we can make sense of particles only as objects similar with the compact medium-seized objects we meet in our daily experience. Even the concept of 1' is biased by my perceptual apparatus. Since seeing is the way I get my main information about the world, I tend to localize myself more in the zone of my head. It should be no doubt that if I had my eyes somewhere on my chest or at the level of my knees or if I had my ears in my heels, I would have a very different sense of myself than I have now.

hidden reality. We may fail to posit some entities in our theories due to the state of art, whereas there are entities reluctant to any disclosure, whose causal interactions (if any) with the observable reality could not be proved empirically.

Let us consider some cases, where some of the causally effective entities are left out by the current scientific explanations. (See Figure 7.4.)



 a hidden event is a common cause; the explanation is false (2) a hidden event is an additional cause (3) a hidden event is an intermediary cause; the explanation is true



In the case illustrated in Fig. 7.4_1, the currently visible phenomena of the causal linkage are the effects, *c* and *e*, of a hidden common cause, *h*. Of course, a theory that causally explains *e*-like events by appeal to *c*-like events is false. Note that the *h*-like events could be revealed just in case there is an experiment such that, if it were

conducted, some *c*-like event would happen in the absence of any *e*-like event. The possibility of such an experiment requires the possibility of a causal route leading to some *c*-like event without passing through any *h*-like event.

In the case illustrated in Fig. 7.4_2, the currently visible phenomena of the causal linkage are c, one of the causes, and its effect, e, the other cause, h, being hidden. Of course, if c is just a partial cause of e, a theory that explains e-like events by appeal to c-like events is false. Again, the h-like events could be revealed just in case there is an experiment such that, if it were conducted, some c-like event would happen in the absence of any h-like event, leading also to the absence of any e-like event. Instead, if c is a sufficient cause of e, the theory is right, although it does not reveal the overdetermination of e. In this case, the revealing of h-like event.

It is important to note that, as long as the *h*-like events are not directly visible, one cannot infer the exact causal linkage in any of the above cases. An experiment in which some *c*-like event happens without any *e*-like event does only show that, in order to explain *e*-like events, there must be something more in the causal linkage. But such an experiment is not enough for revealing whether *h*-like events are either common or additional causes.

Finally, in the case illustrated in Fig. 7.4_3, the currently visible phenomena of the causal linkage are c, a sufficient cause, and its effect, e. In-between, as an intermediary hidden cause, there is h. Then, a theory that explains e-like events by appeal to c-like events is right. The revealing of h-like events depends on the

CEU eTD Collection

occurrence of an occasion when an *e*-like event happens in the absence of any *c*-like event. The occurrence of such an occasion requires a causal route leading to some *h*-like event without passing through any *c*-like event.

Importantly, it might be the case that none of the possible experiments is able to reveal the existence of *h*-like events. In that situation, we are dealing with events that are *in principle* hidden, even though they play a certain causal role in obtaining the observable events to be explained. The epistemic window cuts off just a part of the complete causal linkage, just some causes taking credit for all empirical evidences.

Quine defends the idea that our empirical evidences might support to the same degree two different (incompatible) scientific theories. It works for the present discussion. If some of the constituents of a causal linkage remain hidden, then, in accordance with the available evidence, there could be issued more than one causal explanation for the same physical effect.

In conclusion, the above considerations appear to highlight the debility of the considered argument for (PC_{MS}). Firstly, it has been shown that, even though our current scientific theories of matter provide causal explanations using only material predicates, there is not a reason enough for inferring the nature of the causal processes leading to the explanandum. It may be the case that, by extending the current causal explanatory picture in the borders of our epistemic window, some contingently hidden factors of the causal linkage that are revealed proved to have a non-physical nature. Also, speaking about *in principle* hidden causes, it may be the case that our epistemic window does not include some parts of the causal linkage linkage has a processe that our epistemic window does not include some parts of the causal linkage linkage linkage linkage causal linkage lin

whose nature is non-physical. Secondly, it has been shown that the existence of some hidden parts of the causal process responsible for the explanandum does not necessarily render the current scientific explanation false. Thus, we have a number of good reasons for crediting both (PC_{MS}) and (PC_{M}) with some caution and even some skepticism.

7.5 A "splitting" scenario

Even though, given the argumentation deployed in the previous section, I am not a great fan of (PC_M), in what follows I shall propose a compatibilist scenario in agreement with this reading of the physical closure. Thus, let us add to the framework set up in § 7.3 the requirement that *p* and *e* are connected through a causal process whose constituents are exclusively physical.

Let us consider the following "splitting" scenario, as we may characterize it. Process Π occurs in virtue of some basic law, L_{Π} , activated by the exemplification of *P*, and necessitating the exemplification of *M* and *E*. (See Figure 7.5_2.)



It follows that process Π is $\langle p; -; e \rangle$. (See Figure 7.5_1) Further, the counterfactual from (CNC₂) becomes:

(γ) If *p* and its causal ancestors were the only causes of *e*, property *M* being uninstantiated, then *e* would occur in virtue of the activation of law L_{Π} .

Let us consider a possible world, *w*, where the antecedent of (γ) is true. Thus, in *w* there is an occasion on which *P* is exemplified without *M*. It follows that law L_{Π} does not hold in *w*. *p* causes *e* on that occasion, but this is done in virtue of a law different from L_{Π} . Thus, for any possible world *w*, either the antecedent of (γ) is true, but *p* does not causes *e* in virtue of law L_{Π} , or the antecedent is false. Consequently, counterfactual (γ) is either false or vacuously true; hence (CNC₂) is not met. Therefore, *e* is not overdetermined by *m* and *p* in the scenario under discussion. This comes to support (S), the claim intended to undermine the CEP argument, even though (PC) is read as (PC_M).

Lowe does not exactly take into account (PC_M), but, not very far from it, he considers another mild reading of (PC) inspired by Kim (1993a): "Any physical event which has a cause at time *t* has a physical cause at *t*" (p. 280). The interesting thing is that, given this reading of physical closure, Lowe's compatibilist solution makes use of a certain idea I employ myself in "splitting" scenarios that proves to be crucial for making the CEP assertions compatible. "Simultaneous causation", as Lowe calls it, is a case of causation in which two causal routes starting from the same causes and ending to the same effect. Lowe applies it as follows:

P [a physical event] has a sufficient cause at time t_1 , namely, the set of physical events $\{P_{11}, P_{12}\}$. But *P* also has a cause at t_1 , the non-physical event *M* [which is caused by P_{12}]. *P* is not causally overdetermined by $\{P_{11}, P_{12}\}$ and *M*, because it is not the case that in the absence of either one of them *P* would still have occurred. (p. 577)



Thus, Lowe takes the situation depicted in Fig. 7.6^{171} to be one in which event *P* is not overdetermined just because he takes for granted the counterfactual account of

¹⁷¹ As one may note, even though I use capital letters for denoting properties or instances of properties, for the sake of presenting Lowe's argument, I make an exception here.

causation: "since *P* is caused by both P_{11} and P_{12} , if $\{P_{11}, P_{12}\}$ had been absent, *P* would not have occurred." (p. 577).

However, in chapter 3, I have mentioned some arguments against such an account of causation, favoring instead the nomic approach. I see no reason for claiming that in the counterfactual absence of a sufficient cause, $\{P_{11}, P_{12}\}$, its actual effect, *P*, would not occur. For one thing, even though $\{P_{11}, P_{12}\}$ had been absent, another sufficient cause might replace it in bringing about *P*. In the second place, it appears that, according to Lowe, *P* is not overdetermined because condition (TNC) is not met. Yet, in chapter 6 I have advanced another necessary condition on overdetermination that arguably is able to avoid the dilemmas involved by the application of (TNC). (See § 6.1.) Finally, it is not very clear why Lowe holds that " $\{P_{11}, P_{12}\}$ would *not* have sufficed to cause *P* if P_{12} had not caused *M*' (p. 577). Probably, it is assumed that in the closest worlds even though $\{P_{11}, P_{12}\}$ occurs, the fact that it does not bring about *M* is a (sufficient) mark for its changing its causal powers. However, a change in its causal powers should not be understood as a loose of all its causal powers to the effect that $\{P_{11}, P_{12}\}$ looses including the power of determining *P*.

7.5.1 Is it plausible a triangular linkage of laws?

The aim of this section is to plead for plausibility of the causal linkage of the compatibilist scenario presented in § 7.5. (See Figure 7.5_2, and also Figure 7.7_1.) I shall begin by answering some possible objections. In the end, in the vein of arguing for empirical plausibility, some examples will be conceived for showing that the nomic linkage under discussion may be met as support of some cases of causation.

Lowe rightly remarks that the physicalist cannot simply deny the possibility of simultaneous causation.¹⁷² Such a denying, since it is a "further substantive claim", is in need of an argument. It cannot be taken for granted as an unquestionable truth. On the other hand, the fact that the compatibilist conceives such a thing like simultaneous causation is not enough for grounding its possibility. A question remains: how plausible is it to speak of such a "weird" causal linkage? In particular, admitting the nomic account of causation, how plausible is it to operate with a triangular linkage of laws that would support a case of simultaneous causation?

Let us begin by considering some possible objections to the nomic triangle. In Figure 7.7_1, law L_{Π} (which is the same as L_{Π} from the nomic linkage of the above scenario) is activated by the exemplification of *P*, and necessitates the exemplification of *M* and *E*.



¹⁷² See Lowe (2000a: 577).

Let us consider the following objection. Since the process Π , leading from *p* to *e*, is independent of the physical process Φ , leading from *m* to *e*, one should split not only processes but *laws* as well. In fact, we have here three laws being instantiated: one applying to the physical process from *p* to *e* (say L_{Π}^*); one applying to the bringing about of *m* by *p* (say L_{Π}^{**}); and one applying to the mental–physical process from *m* to *e* (i.e. L_{Φ}). See Fig. 7.7. In other words, L_{Π} is a derivative law obtained from L_{Π}^* and L_{Π}^{**} . Thus, in counterfactual situations in which we have *p* without *m*, L_{Π}^* may still apply, and we may have *e* occur according to the same physical law. And this is enough for overdetermination.

To this it might be replied that even though processes Π and Φ are each individuable, it does not follow that they are independent. Moreover, even though Π is an individuable process, it does not follow that the law grounding it should be individuated inside the causal process. As I have noted already, any causal relation is part of a relation of nomic necessity. Thus, even though a process occurs in virtue of some law, it does not follow that the causal relations making up the process are identical with the law. In particular, even though Π consists of the causal relation from *p* to *e*, it does not follow that there is a law, call it L_{Π}^* , to be described as "if *P* then *E*".

One might ask: What would prevent us from saying that, in fact, law L_{Π} is a derivative law obtained from L_{Π}^* and L_{Π}^{**} , where L_{Π}^* is triggered by the activation of *P*, and necessitates just the exemplification of *E*, and L_{Π}^{**} is triggered by the activation of *P*, and necessitates just the exemplification of *M*?

The direct answer is the following: neither L_{Π}^* , nor L_{Π}^{**} is a law, and the reason for this is simply because none of them is individuable. The idea is that L_{Π}^* cannot be activated without L_{Π}^{**} being activated as well and vice versa. Also, L_{Π}^* cannot be activated without L_{Π} being activated. Thus, L_{Π}^* does not have a law status.

However, one may insist that if it is a law that any instance of *P* is attended (in a way) by instances of *M* and *E*, then it is also a law that any instance of *P* is attended by an instance of *E*. For instance, if it is a law that *Plumbers are tall and handsome*, isn't it also a law that *Plumbers are handsome*?

It seems to me that the dispute comes down to the problem of what it is for an empirical necessity to be a law. From my point of view, a law is the whole recurrent necessity acting at a certain point. The whole necessity is the whole pattern instantiated by necessity. If it's a law that *Plumbers are tall and handsome*, it is true that *Plumbers are handsome*, but it is only an aspect of the pattern of necessity. Consider another law: *Red piles are tasteful and deadly*. That *red piles are tasteful* would be only a half truth.

Let us consider that is a law that *spheres are yellow*. We know that the components of yellow colour are red colour and green colour. So, it is true that by necessity the red colour is present in the colour of any sphere, but in virtue of the law any sphere is yellow.

As I have mentioned, it could be the case that the entities that exemplify property *M*, for example, are *in principle* hidden, that is, they are not inside the epistemic window.

CEU eTD Collection

In such a case, in as much as the nomical necessitation involved by L_{Π} is accounted by any scientific theory, the law posited is inevitably L_{Π}^* . However, even though this is wrong, since L_{Π}^* is not truly a law, the explanations mentioning L_{Π}^* as a law are scientific explanations alright. In other words, even though from a metaphysical point of view a mistake is made, the scientific explanations use correctly the fact that *P*instances necessitate *E*-instances.

Another question one might ask about the nomic linkage of the "spitting" scenario regards law L_{ϕ} . By admitting the existence of L_{Π} , once *P* is exemplified, *M* and *E* are both exemplified as well. Then, what reason would it be for claiming that on the occasion of the exemplification of *P*, the exemplification of *M* still necessitates the exemplification of *E*? In other words, once there is law L_{Π} , what reason would it be for holding the existence of law L_{ϕ} ? The only reason would be given by the existence of a law, L_{Ψ} , according to which *M*-instances are necessitated in the absence of any *P*-instance. (See Figure 7.8.) If there is law L_{Ψ} , *M* can be exemplified independently of any *P*-instance, and thus law L_{ϕ} can be activated without L_{Π} being activated.



An interesting claim made by Lowe (2000) is that the reading one may want to give to physical closure should not be so strong that empirical support for it cannot plausible mustered. By this standard, Lowe appreciates that a reading like (PC_M) is not unacceptable, since, admitting that any naturalistic form of emergentism is plausible, "it is empirically plausible to suppose that the universe contained only phyiscal events at an earlier epoch and that laterday mental events are the ultimate effects of causal chains traceable back to events occurring at that earlier epoch" (p. 578). By contrary, Lowe holds, a reading requiring for any physical effect the existence of a physical cause at any time is too strong. For one thing, Lowe argues, when we consider that temporal ordering is dense, such that between any two times there is another time, the idea of a corresponding density of causes appears to be empirically implausible.

In the spirit of the lesson to be learned from the above considerations, I believe that the issue of the nomic triangle must be approached in the light of the empirical plausibility of there being such a nomic linkage. In what follows, I shall try to show that its empirical plausibility is credited by a number of examples. All of them appear to have in common the fact that the effect is necessitated independently by two distinct laws whose activation is made by the same cause.

Example (1) A bomb has a mechanism of detonation that can be activated both manually by simply hitting it and by radio waves, the latter having a small delay. At some distance there is a remote controller such that by pushing its "on" button it is sufficient for the radio activation of the mechanism of detonation. In the neighborhood there is also a small missile targeting the bomb's detonator such that the missile can be radio triggered by pushing the same "on" button of the remote controller. Thus, on the occasion of the controller button's being pushed, the bomb is exploded because of two sufficient causal processes initiated by the same cause. (See Figure 7.9.)



Fig. 7.9

Example (2) To be sure of the success of its mission, an assassin decides to use a gun with poisoned bullets. Thus, once the victim is hit by such a bullet, the death

happens because of two causes: on the one hand, the bullet deadly penetrate some vital organ, on the other hand, by penetrating the body, the bullet releases some deadly poison. (See Figure 7.10.)



Fig. 7.10

Example (3) Given the success of the above execution, the killer decides to suppress the next victim by using a knife having attached in its tip some deadly explosive capsule. Thus, once the victim is wounded by the killer's knife, the wound is double deadly. (See Figure 7.11.)



Example (4) A person is connected to a cardiostimulator such that the person's life depends on the ongoing functioning of the device. At some distance an atomic bomb is detonated such that its atomic radiations are deadly for the person. Moreover, the same radiations cause the damaging of the cardio-stimulator, bringing about the death through one more route. (See Figure 7.12.)



Fig. 7.12

Example (5) A relay receives as input an electric signal sufficient for its activation. The same signal also goes to an amplifier whose output signal goes further to the same relay, through a parallel derivation. (See Figure 7.13.)



Fig. 7.13

7.6 The notion of physical closure in a strong reading (PC_s)

Have we reached the end of the story? One may object that we are still far from the end. For one thing, many people, when using the CEP argument, work under another reading of the physical closure that is stronger than those two considered so far. I shall call it "the strong reading of physical closure":

(PC_s) The physical domain is nomically closed.

Before pointing out what makes some people to feel entitled to assert such a reading of the closure of the physical domain, some short preliminary clarifications are in needed. As far as I can see, it should be distinguished between two kinds of closures: the *explanatory* closure and the *nomic* closure. The former is true of any domain coming into the scope of a *complete* scientific theory. For example, once we get a complete micro-physical theory, we are entitled to consider its domain of discourse (i.e. micro-physics) as being explanatorily closed: for any event happening in the domain of micro-physics, there is an explanation provided by the theory.

Let us assume that, ultimately, any scientific theory explains the behavior of some entity as being the result of the application of certain laws to some initial conditions, where the laws are stated by the respective theory. It follows that:

For any scientific theory, its domain of discourse is *explanatorily closed* if and only if the behavior of any entity from the domain is explicable in virtue of the laws of the theory.

Instead, in the case of the second kind of closure, the idea is that the entities from the domain are not connected with any "alien" entity in virtue of any laws. It follows that:

A domain should be considered as being *nomincally closed* if and only if the laws governing the interior of the domain do not involve its exterior.¹⁷³

If we turn back to the question 'What would it be the reason for asserting the nomical closure of the physical domain?', we should begin by noting that (PC_S) is a claim usually made by someone with sympathy for physicalism. It should be noticed that physicalism is a monist doctrine emanating from materialist monism. The reason for the detaching from materialism, i.e. the doctrine holding that everything is material, was that, along with the scientific progress, various scientific theories, especially, those from the domain of micro-physics, posited all sort of entities (e.g. corpuscles, force fields, subatomic particles and the like) to which the traditional concept of matter does not apply anymore. Thus, the smart move made by the materialist monists was the commitment to physics, which, as Crane (2001a) correctly points out, enjoyed high consideration:

A physicalist is someone who gives a certain kind of authoritative role to physics. This role is partly epistemological – physics has an authority in telling us what to believe, partly ontological – physics has an authority in telling us what there is. (p. 44)

By assigning epistemological authority to physics, there is expressed the belief in the explanatory closure of the domain of physics, and also the expectancy of finding a complete theory of physics. Whereas, since the physicalist invests physics with ontological authority,¹⁷⁴ it appears that physicalism is a doctrine that takes any entity to be ontologically reducible to the very entities posited by a supposed complete

¹⁷³ If it is granted the nomical account of causation, the explanatory closure involves the causal explanatory closure (see § 7.4). The same, the nomic closure entails the causal closure. ¹⁷⁴ Or, simply, the physicalist *acknowledges* the ontological authority physics.

theory of physics. The problem here is that, as it has been discussed in Chapter 1, if the physicalist holds that the primitive constituents of our world are the entities of physics, this exposes physicalism to the risk of ceasing to be a monist doctrine. For one thing, the future may prove that the complete theory of physics introduces in its ontology entities completely unlike those posited by any of the current theories in physics.

My suspicion is that the physicalist derives the nomical closure of the physical from the following assumptions: the physical domain is the same as the domain of discourse of physics; the domain of physics is explanatorily closed; but, the explanatory closure involves the nomical closure.

I have already discussed the first assumption. Regarding the second, it has in the background the belief that a complete theory of physics is *in principle* possible. Let us take a closer look at the third assumption. Prima facie, it may seem that the domain of discourse of a complete scientific theory is not only explanatorily, but also nomically closed. If the behavior of any entity from the domain is subject of some laws of the theory, one might believe that any such entity is nomically connected exclusively with entities from the same domain. But, this is not right, because it might be the case that certain laws of the theories are derivative. For example, a law, L_D , may be derivative in virtue of the conjugation of some other laws. $L_C^{(1)}$, $L_C^{(2)}$,... For simplicity, let us consider only the first two. If each of the constituent laws, $L_C^{(1)}$ and $L_C^{(2)}$, trespasses against the domain of discourse, connecting internal entities with external ones, the domain is not nomically closed: even though the derivative law

CEU eTD Collection

supports only relations holding among internal entities, those relations are mediated by external entities. (See Figure 7.14.)



Fig. 7.14

Finally, if the physical were a nomically closed domain, it would be closed in both directions, and the following assertions would be each true:

(PC_{so}) Physical entities do not nomically necessitate any non-physical entity;

(PC_{SI}) There is no nomical necessitation oriented from the things outside of the physical domain to those inside.

As I have argued in Chapter 2, minimal interactionism is a conjunction of two theses:

(D) Mental properties are not token-identical with physical properties;

(C) Mental phenomena and physical phenomena can causally affect one another.

By granting the nomic account of causation, interactionism as the conjunction of (C) and (D) implies both the existence of some nomic relationships from the things outside of the physical domain to those inside, on the one hand, and the existence of others going the other way round, on the other hand. Therefore, by asserting either (PC_{SO}) or (PC_{SI}) , there is also asserted the falsity of interactionism. The moral is that if one uses the strong reading of (PC) in the CEP argument, which is an argument against interactionism, one simply begs the question.

7.6.1 Lowe's "divine" scenario

At some point, Lowe (2000) takes into account a certain reading of physical closure that is very close to what I presented above as being an implication of (PC_S). He calls it "causal closure of the physical domain":¹⁷⁵ "The immediate causes of all physical events are always and only other physical events." (p. 582) I have argued that if it is used in an argument against interactionism the assumption of nomic closure of the physical domain is question-begging, because it is asserts what the dualist denies. If the above is right, granting the nomic account of causation, the assumption of the causal closure of the physical domain appears to be question-begging against interactionism, too. However, Lowe, by contrast, advances a compatibilist scenario intended to work even under the causal closure hypothesis:

It could conceivably be the case that ... sometimes a non-physical mental event M causes it to be the case that certain physical events, P_1 , P_2 ,... P_n , have a certain physical effect, P. What this would involve is the causation by a mental event of a physical causal fact. So I am assuming now a distinction between event-causation and fact-causation. In a case of fact-causation, what is brought about is not an event, but a fact or state of affairs. Some facts are causal facts, such as the fact that certain physical events P_1 , P_2 ,... P_n , are causes of another physical event, P. There is no

¹⁷⁵ Recall that Lowe and many others use 'causal closure of the physical domain' to refer generically to physical closure. In contrast, I have reserved the phrase for referring to a particular reading of the notion of physical closure.

reason in principle why such a causal fact should not have a causal explanation... Now, if a mental event *M* causes it to be the case that certain physical events P_1 , P_2 ,... P_n have a certain physical effect, *P*, then, it seems clear, *M* is itself a cause of *P* – but not an immediate cause of *P*, nor an immediate cause of any of the physical events P_1 , P_2 ,... P_n . (p. 582, see Figure 7.15.)



Fig. 7.15

I have two critical remarks regarding Lowe's proposal. First, Lowe says that there is no reason why "the physical causal fact" *should not* have a causal explanation, but, on the other hand, one may say that there is no reason why "the physical causal fact" *should* have a *causal* explanation. It seems al right to say that the physical causal fact should have an explanation or, better put, a metaphysical account, but why should it be causal in nature? Take, for example, the fact that a thing has a property. Arguably, there should be an account of the given fact, but there is nothing to point out to causal account. One cannot plausible hold that a fact should have a causal explanation just because it is a causal fact. Neither, it can be held that any causal fact can receive an account in causal terms.

Second, admitting that a certain event is responsible for (or necessitates) the occurrence of some causal fact, why should one admit that the responsibility (or the necessitation) is causal in nature? It seems to me that such a responsibility is of a different nature than the responsibility had by causes involved in the causal fact. Consequently, even though the event necessitates the causal fact, and even though events are usually the protagonists of causal relations, it seems more reasonable here to assume that the necessitation is not causal in nature.

In order to show that the notion of a mental event causing a physical causal fact is intelligible, Lowe deploys a very nice possible case:

Suppose that there is a world in which the immediate causes of all physical events are always and only other physical events, a world in which every physical event has a sufficient physical cause and no physical causation is either simultaneous or backward. Such a world can have no beginning in time, because it can have no first physical event. (...) And yet we could still ask of this world why *it* should exist or be actual rather than any other. One intelligible answer would be to say that this world was actual because God had chosen it to be actual. God's choice, then, would have caused it to be the case that a world containing certain physical causal facts was actual – and this would be mental causation of physical causal facts. (p. 583)

True, nobody would be expected to claim that the above story is unintelligible. We understand the idea that the world is the way it is in virtue of God's will. However, Lowe speaks about God's choice of *causing* the world to be the way it is. But is it legitimate to take 'God caused the world' as being the same thing with 'God made the world'? For that comprehending, it would be need a very loose sense of 'cause'.

7.7 More working frameworks, more compatibilist scenarios

One may object to the working framework set up in § 7.3 that, even though it is in accordance with (MC), nobody would go so far as to really believe that the mental can modify the brain without any help from the physical. I am neutral in this matter, not seeing any substantive reason for taking (MC) one way or another. However, for the sake of the argument, I think it is fair enough to consider one more framework, in which this time the mental is just a "helping" factor. That is, the mental just makes a contribution to the obtaining of some neural effect, without being capable by itself to assure the neural modification.

Thus, let us consider that physical event *e* has mental event m^* as a partial cause. For simplicity, suppose that m^* is the only mental cause of *e*, and causation from m^* to *e* is made without any intermediary. Given that m^* is just a partial cause of *e*, there must be a physical event, p^* , such that m^* and p^* together make up *m*, a causally sufficient condition for *e*. Let us assume that the causal sufficiency is grounded on a fundamental law, L_{ϕ} , activated by the exemplification of properties M^* and P^* , and necessitating the exemplification of property *E*. Each property is exemplified on the occasion of m^* , p^* , and *e*, respectively. (See Figure 7.16.)



In accordance with (PC_W), there is also p, a causally sufficient condition on e such that p belongs to the physical domain. Next, (NI) requires that mental properties are not realized by physical properties; hence, mental events are not identical with physical events. Therefore, given that condition m contains a mental event, m^* , and condition p contains only physical events, m is not identical with p.

But this is the moment when one may realize that there could be two cases and hence two frameworks. In one case, the two conditions, m and p, do not share anything. In the other case, m and p overlap each other, sharing some physical cause. In what follows, each case (framework) will be treated individually.

7.7.1 m and p do not overlap each other

Let us set up the new working framework according to the hypothesis that m and p do not overlap each other. For simplicity, I take p to be constituted of only one event, call it p as well, its final property being P. (See Figure 7.17.)



CEU eTD Collection

We come back to the relevant question: Do the two distinct causally sufficient condition, *m* and *p*, overdetermine *e*, no matter what? Again, I shall try to prove that the overdetermination of *e* may be avoided. Let us begin by noting that, if there are two processes, Φ and Π , whose originating causes are among m^* , p^* , and *p*, Φ and Π being each causally sufficient for *e*, then the only chance for (CNC) to be met is that their originating causes are m^* and p^* for Φ , and *p* for Π .

Thus, in the given framework, let us assume that *R* is composed of two causal relations, one from *p* to *m*^{*}, the other one from *p* to *p*^{*}, such that *p* is a sufficient cause both for *m*^{*} and for *p*^{*}. Therefore, process Φ cannot be but $<m^*$, *p*^{*}; -;*e*>. Instead, process Π can be assigned with more than one linking network: Π is <p; ?; *e*>.

Then, particularizing (CNC), physical effect *e* is overdetermined by m^* , p^* , and *p* only if:

(CNC₁) <*m**, *p**; -;*e*> and <*p*; ?; *e*> do not stand in a part-whole relationship, and
(CNC₂) the following counterfactual is non-vacuously true:

(γ) *e* would still occur in virtue of the same laws that actually support process <p; ?; *e*>, even if the nodes of <*p*; ?; *e*> and their causal ancestors were the only causes of *e*, at least one of properties *M** and *P** being uninstantiated in the case of *m** and *p** not being nodes of <*p*; ?; *e*>.

Thus, within the present framework, under the assumption that p is causally sufficient both for m^* and for p^* , the claim that I advocate comes to saying:

(S) Even though conjunction (NI) & (MC) & (PC) is true, there might be at least a scenario in which *e* is not overdetermined.
7.7.2 Another "candid" scenario

As long as the notion of physical closure receives a weak understanding, another "candid" scenario is enough for supporting (S). Thus, let us consider a hypothetical situation where the laws underlying process Π are L_{Π} and L_{ϕ} , the former being triggered by the exemplification of *P*, the latter by the exemplification of *M** and *P**. (See Figure 7.18_2.)



It follows that process Π is $\langle p; m^*, p^*; e \rangle$; hence, process Φ , i.e. $\langle m^*, p^*; -; e \rangle$, is a serial part of Π . (See Figure 7.18_1) But this is a violation of (CNC₁). Therefore, *e* is not overdetermined by m^* , p^* , and *p* in a hypothetical situation described by the

scenario under discussion. This supports (S), the claim intended to undermine the CEP argument.

7.7.3 Another "splitting" scenario

However, one may insist to give a stronger reading to the notion of physical closure. By taking for granted (PC_M), which is, as I have argued, the strongest acceptable reading of (PC). Then, a compatibilist scenario must be one of the same style as the previous "splitting" scenario. Thus, let us consider a hypothetical situation where process Π occurs in virtue of some basic law, L_{Π} , activated by the exemplification of *P*, and necessitating the exemplification of *M**, *P**, and *E*. (See Figure 7.19_2.)

It follows that process Π is $\langle p; -; e \rangle$. (See Figure 7.19_1) Further, the counterfactual from (CNC₂) becomes:

(γ) If *p* and its causal ancestors were the only causes of *e*, at least one of properties *M*^{*} and *P*^{*} being uninstantiated, then *e* would occur in virtue of the activation of law L_{Π} .



Let us consider a possible world, *w*, where the antecedent of (γ) is true. Thus, in *w* there is an occasion on which *P* is exemplified without *M** or *P**. It follows that law *L*_Π does not hold in *w*. *p* causes *e* on that occasion, but this is done in virtue of a law different from *L*_Π. Thus, for any possible world *w*, either the antecedent of (γ) is true, but *p* does not causes *e* in virtue of law *L*_Π, or the antecedent is false. Consequently, counterfactual (γ) is either false or vacuously true; hence (CNC₂) is not met. Therefore, *e* is not overdetermined by *m**, *p**, and *p* in the hypothetical situation described by the present scenario. This supports the truth of (S), the claim intended to undermine the CEP argument, even though (PC) is read as (PC_M).

7.7.4 m and p overlap each other

I turn now to the alternative in which conditions *m* and *p* overlap each other. For simplicity, I take their common part to be maximal, namely, the whole physical constituent of *m*, which is exactly p^* . Moreover, p^* being a constituent of condition *p*, there must be another physical event, let it be p^{**} , such that p^* and p^{**} together make up *p*. (See Figure 7.20.)



I believe that the two distinct causally sufficient condition, m and p, do not always overdetermine e. As before, the strategy is to find at least a scenario describing a situation in which condition (CNC) is not met. In this sense, the immediate relevant

remark is that, if there are there are two processes, Φ and Π , whose originating causes are among m^* , p^* , and p^{**} , Φ and Π being each causally sufficient for e, the only chance for (CNC) to be met is that m^* and p^* are Φ 's originating causes, and p^* and p^{**} are Π 's originating causes

7.7.5 One more "candid" scenario

The following scenario is advanced as a compatibilist one in the context of granting (PC_W). Let us consider a hypothetical situation where *R* is a sufficient causal relation holding between p^{**} and *m*. In the line with all candid scenarios proposed above, let us take process Φ to be $\langle m^*, p^*, -; e \rangle$, and process Π to be $\langle p^*, p^{**}; m^*; e \rangle$. The laws supporting process Π are L_{Π} and L_{Φ} . Law L_{Π} is triggered by the exemplification of P^{**} , it necessitating the exemplification of M^* . Law L_{Φ} is triggered by the exemplification of M^* and P^* , it necessitating the exemplification of *E*. (See Figure 7.21.)

It is obvious that Φ is a serial part of Π , which is a violation of (CNC₁). Therefore, *e* is not overdetermined by m^* , p^* , and p^{**} in a hypothetical situation described by the present scenario. This supports (S), the claim intended to undermine the CEP argument.



7.7.6 One more "splitting" scenario

The following scenario is advanced as a compatibilist one in the context of granting (PC_M). Let us consider a hypothetical situation where *R* is composed of two causal relations, one from p^{**} to m^* , the other one from p^* to m^* , such that p^{**} and p^* is a causally sufficient condition for m^* . Therefore, if there are two processes, Φ and Π , whose originating causes are among m^* , p^* , and p^{**} , Φ and Π being each causally sufficient for *e*, the only chance for (CNC) to be met is that Π is $<p^{**}$, p^* , -;e>, and Φ is $<p^*$, m^* ; -;e>. L_{Π} , i.e. the law supporting process Π , is triggered by the exemplification of P^{**} and P^* , it necessitating the exemplification of M^* and *E*. L_{Φ} , the law supporting process Φ , is triggered by the exemplification of P^* and M^* , it

necessitating the exemplification of *E*. Note the novelty brought by this scenario: p^{**} is a cause of *e* in virtue of two distinct laws. In fact, there are two causal relations between p^* and *e*, in each of them p^* being just a partial cause of *e*. (See Figure 7.22.)





Then, particularizing (CNC), the physical effect, *e*, is overdetermined by m^* , p^* , and p^{**} only if:

(CNC₁) <p**, p*; -;e> and <p*, m*; -; e> do not stand in a part-whole relationship, and
 (CNC₂) the following counterfactual is non-vacuously true:

(γ) If the nodes of $<p^{**}$, p^{*} , -; e> and their causal ancestors were the only causes of e, property M^* being uninstantiated, then e would occur in virtue of the activation of law L_{Π} .

Let us consider a possible world, *w*, where the antecedent of (γ) is true. Thus, in *w* there is an occasion on which P^{**} and P^* are exemplified without M^* . It simply follows that law L_{Π} does not hold in *w*. p^{**} and p^* causes *e* on that occasion, but this is done in virtue of a law different from L_{Π} . Thus, for any possible world *w*, either the antecedent of (γ) is true, but *e* is not caused in virtue of law L_{Π} , or the antecedent is false. Consequently, counterfactual (γ) is either false or vacuously true; hence (CNC₂) is not met. Therefore, *e* is not overdetermined by m^* , p^* , and p^{**} in the hypothetical situation described by the present scenario. This supports (S), the claim intended to undermine the CEP argument, even though (PC) is read as (PC_M).

My thesis is not intended as a key fitted for unlocking the question of how psychical phenomena intermingle with bodily ones. Rather it is an invitation to a more thoughtful analysis of the too easily claimed conflict between the idea of a causally potent mind of a non-physical kind and the hypothesis of the closure of the physical domain. Such a conflict is known as the Causal Exclusion Problem (CEP). The whole thesis is nothing else than an attempt at showing that CEP is not quite such a firm ground as is often thought for building up an argument that interactionism could not be an answer to the mind-body problem.

I have shown that interactionism is a particular view of the dualist doctrine whose main thesis is that, at the level of the ultimate realities of our world, there are two basic ontological realms, mental and physical.¹⁷⁶ As I have argued, the core dualist thesis boils down to the following: mental properties are not token-identical with physical properties. The dualist thesis of non-identity makes a claim of non-identity in the strongest sense; it rejects not only any type-identity between mental and physical properties, but also any realization relation and hence any token-identity between the instances of these properties.

But what is it for an entity to be physical? What is it to be mental? In my search for a mental-physical distinction available for all mind-body theories, I have reviewed two

¹⁷⁶ Generally speaking, on the monist view everything that exists is ontologically reducible to one single kind of basic entities, whereas the dualist doctrine claims that any non-basic entity has a constitution made up either of only pure physical entities, or of only pure mental entities, or of mixed entities.

influential approaches. It has been shown that the *theory approach* is vulnerable because, in order to "locate" the basic physical entities, the approach relies on the ontology of the current micro-physics. If the future explanations given in micro-physics appeal to some of what we definitely take to be mental entities, the mental-physical distinction traced by theory approach will collapse.

The analysis of the *paradigmatic approach* has suggested a better way to start in tracing the mental-physical distinction, namely, first to settle the meaning of 'mental property' and only then to derive from it the meaning of 'physical property'. Thus, I have come to the conclusion that, by 'mental property', should be meant the kind of property instantiated by (or in) phenomenal experiences. Then, a physical property is nothing else but a property of the same sort as any property featuring among those to be mentioned in an exhaustive account of a paradigmatic physical object; such an object would be one devoid of phenomenal experiences.

Yet, it might be objected to the above distinction that it leaves no room for neutral monism, namely, the view that there are ultimately entities that are neither physical, nor mental, but account for any physical or mental reality. To this I have answered by showing that the notion of 'neutral' is highly problematic in the context of the mind-body problem, given the difficulty of conceiving something capable of constituting both phenomenal experiences (or minds or subjects) and non-sentient objects. Moreover, the content (the meaning) of 'neutral' remains mysterious, because in principle neutral monism cannot identify what makes some entity to be physical and, at the same time, what is missing from its basic constituents such that they are only neutral.

* * *

Besides the dualist thesis of mental-physical distinctness, the interactionist version of the dualist doctrine claims that mental phenomena and physical phenomena can causally affect one another. I believe that one cannot undertake a substantial discussion on the truth of interactionist theses, in general, and on how it is affected by CEP, in particular, unless one takes a definite stance on the matter of causation itself. One of the fundamental convictions that have shaped my analysis of CEP is that the issue of causation should be understood in the light of the metaphysical relations supported by the laws of nature.

I have held that there is an a priori correlation between causes and laws. The reason I have brought for this is that there are two relevant features inbuilt in our commonsensical notion of causation. First, we all seem to agree that it is a sort of *necessitation* that is involved in bringing about an effect in the sense that, given the occurrence of its causes, it is compelling for the effect to occur. Secondly, we all seem ready to acknowledge a case of causation provided that, with other occasions, like causes produce like effects. Thus, it is a plausible claim that a causal relation is an instance of a relation of *recurrent necessity*. But, in the empirical domain, such a relation of necessity is the very necessitation involved by the notion of a law, namely *empirical necessitation*.

Given the above considerations, we are naturally led to say that the features of necessity and of generality constituting our concept of causation are to be naturally explicated as deriving from the necessities and generalities involved by the

CEU eTD Collection

application of some laws of nature. In other words, the recurrent constraint exerted by similar causes on similar effects is nothing over and above the sort of necessitation to which we are committed when we acknowledge the laws of nature.

Therefore, I have based my study on a nomic account of causation whose main claim is that any causal relation is part of some relation of nomic necessity in the sense that the causal relata are among the relata of the necessitation. Thus, in this view, if c is a cause of e, then there is some relation of nomic necessity holding between two sets, C and E, such that c is a member of C and e is a member of E. That c is a cause of e boils down to the fact that they are involved in a relation of necessity holding in virtue of certain laws. Importantly, the laws underlying the necessitation of e are not triggered only because of c, but rather because of the whole bunch of individuals belonging to set C.

In the spirit of D. Lewis's analysis of causation, the nomic account of causation urges that an effect is linked by any of its causes through some chain of causal dependences. However, contrary to Lewis, a causal dependence is not taken to be a counterfactual dependence, but rather a nomic dependence. Following F.G. Clendinnen, I have held that just some of the nomic dependences, namely, those empowered by fundamental laws, should be thought of as being causal.

By definition, a primitive causal relation is a direct or intimate causal connection in the sense that there is no need of an intermediary cause. Terminologically, I have stipulated that a direct cause of an effect is such that there is a primitive causal relation standing between the cause and the effect. An argument has been deployed

CEU eTD Collection

for the conclusion that a causal relation is primitive if and only if the effect is causally dependent on the cause. One consequence is that a primitive causal relation is part of some fundamental relation of nomic necessity. The second consequence is that any causal relation is either primitive or decomposable in a chain of primitive causal relations.

* * *

I am committed to the idea that the world contains both substances and events – two ontological kinds of particulars. A belief playing a crucial role in my analysis of CEP is that there is no causation without events. The identity of the causal network involved by the study of CEP comes in part from the identity of the events featuring as causal relata.

The issue of event identity is highly dependent on what general nature events have. I have pointed out that D. Davidson makes a credible plea for including events in our ontology, resting his case on the benefits they could bring in accounting for action sentence entailments. However, as I have argued, Davidson's approach leaves unanswered the question of what events are. Their nature cannot be derived or deduced from the role they play in accounting for certain entailments.

I have suggested that events are entities having a metaphysical structure. Generally speaking, by saying that an entity, x, has a structure, it should be understood that *something* – some ontological stuff – is structured in a certain way such that x is

CEU eTD Collection

obtained. In the Aristotelian style, by 'matter', I have called what is structured, and by 'form', how the matter is structured.

Importantly, I have held that one has to resist the temptation of taking a structured entity as being nothing over and above its matter and the form that structures that matter. Such an entity is also what I have called the *structuration of the matter in that form*. That is, a structured entity is nothing besides some entities, some places in a form, and the matching between the entities and the places.

Further, I have argued that what Kim takes to be an event, namely the exemplification of a property by a substance at a moment of time, is not really an event, but rather a state of the respective substance. An event happens to a substance when and only when that substance undergoes a change in the exemplification of some of its properties during a period of time. Then, naturally, an event is a transformation in states of a substance during a period of time. The conclusion is that any event has a metaphysical structure characterized by a form that is temporal in nature, its places being occupied by states. At a higher level of detail, any event has a temporal metaphysical structure consisting of a substance (i.e. an entity persisting through changes through time) and a range of properties the substance exemplifies successively; the form of the structure is given by a certain *timing* of property exemplification during the period of the considered change.

Any change is time consuming; it involves a period of time during of which a substance usually loses properties in favor of acquiring others. Such a change in properties is made, in the case of each event, in a specific manner. In most of the

CEU eTD Collection

cases, for characterizing (individualizing) an event, it is not enough to provide the initial state and the final state; it is also necessary to mention the profile of the transition between those two states. This kind of profile is given both by the intermediary states and by the timing of transition.

Such kind of considerations have led me to the idea that the identity of an event is given by structure $\langle s, P_i, P_f, (t_i, t_f), p \rangle$, where *s* is the substance (or the system of substances) participating in the event, P_i is the initial property had by *s*, P_f is the final property *s* acquires instead of P_i , (t_i, t_f) is the period needed for change to be accomplished, and *p* is the profile of change. Profile *p* consists of the intermediary properties *s* acquires and looses in going from its initial state to its final state, plus the respective moments or periods when *s* has each intermediary property.

* * *

In the issue of how psychical phenomena intermingle with bodily ones interactionism proposes a metaphysical picture where they can causally affect one another even though mental properties are not token-identical with physical properties. Then, if the causal relata are events or states, granting that events are changes in properties, interactionism involves that there are some mental events causally relevant for some other events of the physical world in spite of the fact that mental events have a nonphysical nature.

To the above suggestion it might be objected that it is in conflict with a popular assumption of our days, namely, the hypothesis of physical closure. The idea would be that, if one is committed to physical closure, the interactionist doctrine practically asks one to admit that in any case of mental causation the physical effect is overdetermined. Thus, the trouble would be that interactionism leads to a highly implausible situation. This is in short the pretended difficulty for interactionism which CEP emphasizes.

As I have shown, what should motivate us to reject a widespread and systematic overdetermination is an assumption of rationality and intelligibility. The world is such that its nature (its constituents and its constitution) is according to "our" principles of rationality; else it would be unintelligible for us. Nobody should admit a metaphysical theory that delivers us an irrational picture of the world. By 'our principles of rationality', it should not be understood some human-mind-dependent principles, but rather some objective principles governing both our thinking and the constitution of the world.

The commitment to the existence of massive overdetermination would require one to admit the picture of a world devoid of elegance, grace, balance, shapeliness, and simplicity. As it has been often noted, if such a world were the result of someone's design, one would be justified in describing the design as being deficient. For one thing, the designer would wrongly oversize the whole thing by making the world to spend more resources than it needs for its becoming.

Even though one accepted that events are overdetermined in some cases, one would not rationally afford to buy the idea that all mental causation is done by overspending causal resources. Therefore, the moral would be that any account of mental

CEU eTD Collection

causation is more preferable to one implying that each and every bit of action we perform (or, more to the point, any neurological event involved by mental processes) is overdetermined.

* * *

But, does interactionism really imply a widespread and systematic overdetermination? Is the implication suggested by CEP really warranted? It appears that a supporter of interactionism not willing to defend his doctrine at the cost of contesting physical closure should base his comeback on finding some compatibilist scenario. In such a scenario, if it were real, the physical effect would not be overdetermined even though it would have a non-physical cause.

The key point of my thesis is the endeavor to set up such a kind of compatibilist scenario. In this sense, I have followed a strategy based on finding a *necessary condition on overdetermination* and some *relation tokening among some of the causes of the physical effect* such that, if the relation held, the necessary condition would be violated.

In all fairness, by showing that there *might* be some circumstance in which it is not the case that interactionist mental causation involves overdetermination, it is not proved that interactionism does not imply overdetermination. The falsification of the implication requires some reasons for accrediting the actuality of the considered compatibilist scenarios. However, the scenarios do cast doubt on the truth of the implication. By setting up such a scenario is enough for raising the question of

whether or not it is real. To show that the scenario is only fictional a separate argument is needed. There is no way of saying whether or not it is according to reality just by looking at it. As long as there is no convincing argument that our actual world does not accommodate the scenarios, there is no guarantee for the soundness of CEP and it cannot be used in a conclusive argument against interactionism.

* * *

Causes are linked in causal chains that further generate the causal history of any effect. However, in giving causal explanations, causal processes are more workable than causal histories. The identity of a causal process, as I have defined and elaborated its notion, is given by a triplet consisting in its *originating causes*, its *ending effects*, and its *linking network*, that is, the causal network that joins the events that start the process to those with which the process ends up. One constraint is that, for any couple formed by an originating cause and an ending effect, there should be a causal chain connecting them.

Another notion that I have introduced is that of 'causal role'. Its benefit is that it makes room for discrimination among the different ways the same effect is or could be directly caused. Thus, 'c is a direct cause of e' has been taken to be synonym to 'c has a certain causal role with respect to e'. Given the definition of 'direct cause', it has come out that the causal role c has had in bringing about e gains its identity from both the property obtained in c and the triggered causal laws that necessitated e. Thus, as I have shown the transworld identity of some causal role requires the tokening of the same causal property and the triggering of the same fundamental

laws of whose activation the causal property is (co-)responsible. By extension, the causal role had by a causal process in bringing about one of its ending effect is made up of all causal roles played by its nodes on the way of the causal chains leading to the ending effect.

The finding of a necessary condition on overdetermination depends very much on what one understands by 'overdetermined effect'. I have held that 'overdetermination' is a derivative notion, and hence its mastering firstly requires a conceptual analysis. Its understanding should be grounded in the grasping of other two notions: 'causal process sufficient for an effect' and 'excess in bringing about an effect'.

I have argued that a process should be taken to be causally sufficient for one of its ending effects if and only if, in the corresponding relation of nomic necessity, the originating causes are identical with the necessitating relata. Of course, some of the causes of an effect make up a causally sufficient condition for it if and only if there is some relation of nomic necessity such that the causes are identical with the necessitating relata and the effect features among the necessitated relata.

To my mind, the existence of two processes causally sufficient for an effect amounts to an excess in bringing the effect if and only if the sufficiency of each process is not dependent on the sufficiency of the other. Thus, I have argued that, considering that Φ and Π are two processes, each of them being causally sufficient for *e*, there is an excess in causing *e* if and only if neither Π 's causal role with respect to *e* depends on Φ 's causal role with respect to *e*, nor *vice versa*.

CEU eTD Collection

Thus, a necessary condition for the independence required above is the truth of the following counterfactuals:

(α_1) if Π 's nodes and their causal ancestors were the only causes of *e*, then, relative to *e*, Π 's causal role would be identical with its actual causal role.

(α_2) if $\boldsymbol{\Phi}$'s nodes and their causal ancestors were the only causes of e, then,

relative to e, Φ 's causal role would be identical with its actual causal role.

By taking *e* to be an overdetermined effect only if there is an excess in bringing *e* by two processes, Φ and Π , each of them being causally sufficient for *e* in virtue of laws L_{Π} s and, respectively, L_{Φ} s, I have shown that a necessary condition on *e*'s being overdetermined by two or more of its causes, $c_1, c_2, ...$ is the following:

(3) there are two processes, Φ and Π , each of them being causally sufficient for *e*, such that some of c_1 , c_2 , c_3 , ... feature among the originating causes of the processes;

(4) the following counterfactuals are non-vacuously true:

(γ_1) if Π 's nodes and their causal ancestors were the only causes of *e*, at least one of the properties of *e*'s direct causes from [$\Phi - \Pi$] being uninstantiated on the occasion of *e*'s occurrence, then *e* would occur in virtue of L_{Π} s's activation; (γ_2) if Φ 's nodes and their causal ancestors were the only causes of *e*, at least one of the properties of *e*'s direct causes from [$\Pi - \Phi$] being uninstantiated on the occasion of *e*'s occurrence, then *e* would occur in virtue of L_{ϕ} s's activation.

(3) ϕ and Π do not stand in a part-whole relationship.

* * *

I have taken into account three possible working frameworks for mental causation, designing in each of them scenarios in which, even if mental causation is consistent with the interactionist claims, causation does not involve overdetermination. The first considered framework allows mental cause to be sufficient for the physical effect. However, given that one may object that nobody would go so far as to really believe that the mental can modify the brain without any help from the physical, I have sketched two more frameworks suitable for the study of the cases where the mental is just a "helping" factor (i.e. a partial cause) in bringing about the physical effect.

The hypothesis of physical closure may receive more than one version. In accordance to what reading is taken for physical closure, from the weakest to the strongest, the compatibilist scenario has to meet particular requirements and pass over specific difficulties.

The *weak* reading of the hypothesis of physical closure renders the hypothesis that for any physical effect there is a complete causal explanation made only in physical terms. Instead, the *mild* reading involves some supplementary claim, namely, that there should be always a causal route from a physical cause to its physical effect such that it does not go out of the physical domain. Thus, the mild reading is that for any physical event there is a physical causally sufficient condition leading to the effect through a physical causal process.

Another reading of physical closure that has been also considered asserts that the physical domain is nomically closed. I have argued that it is too strong a hypothesis to be used in an argument against interactionism, pointing out that it should not be

CEU eTD Collection

taken for the hypothesis of explanatory closure of physics that enjoy different credentials.

In order to make a distinction, I have shown that for any scientific theory its domain of discourse is *explanatorily closed* if and only if the behavior of any entity from the domain is explicable in virtue of the laws of the theory. Instead, in the case of nomic closure, the entities from the domain are not connected with any "alien" entity in virtue of any laws. Thus, a domain should be considered as being *nomincally closed* if and only if the laws governing the interior of the domain do not involve its exterior.

By granting the nomic account of causation, interactionism involves the claim that there are some nomic relationships between the things inside the physical domain and those outside amongst which, among others, feature the mental entities. Therefore, the use of the hypothesis of the nomic closure of the physical in an argument against interactionism amounts to begging the question against interactionism. I have found these considerations to be convincing enough to lead me to discard the last reading of the hypothesis of physical closure, sticking only with the first two.

* * *

Some of the compatibilist scenarios that I have set up have in common a causal linkage that can be described as *simultaneous causation*. Explicitly, in such cases, causation is accomplished in such a way that the effect is necessitated independently

CEU eTD Collection

by two distinct laws whose activation is made by the same cause. In its simplest form, the causal linkage looks like a triangle. (See one more time Figure 7.5.)



The idea of causation in triangular linkage is not entirely new; it has been mentioned or used by philosophers like David Lewis and E.J. Lowe. How plausible is it to speak of such a "weird" causal linkage? As Lowe rightly notes, one cannot simply deny the possibility of such kind of causation.¹⁷⁷ Such a denial, since it is a "further substantive claim", is in need of an argument.

On the other hand, it should be noted that the fact that one conceives such a thing like simultaneous causation is not enough for grounding its possibility or, much more, its occurrence in our world. In this spirit, I have approached the issue in the light of

¹⁷⁷ See Lowe (2000a: 577).

empirical plausibility, by trying to credit it by a number of examples. After all, in philosophy, plausibility is all that matters.

Bibliography

- Armstrong, D. H. (1983) *What is a Law of Nature?*, Cambridge: Cambridge University Press.
- Armstrong, D. M. and Heathcote, A. (1991) 'Causes and Laws', Noûs 25, 63-73.
- Armstrong, D. M. (1999) 'The Open Door: Counterfactual Versus Singularist Theories of Causation' in H. Sankey (ed.), *Causation and Laws of Nature*.
- Armstrong, D. M. (1999a) 'Universals as Attributes', in J. Kim and E. Sosa (eds.), *Metaphysics: an anthology*, Malden Mass.: Blackwell Publishers.
- Bennett, K. (2003), 'Why the Exclusion Problem Seems Intractable, and How, Just Maybe, To Tract It', *Noûs* 37:3, pp. 471-497.
- Bigelow, J., and Pargetter, R. (1990) 'Metaphysics of Causation', *Erkenntnis* 33, pp. 89–119.
- Bigelow, J., Ellis, B., and Lierse, C. (1992) 'The World as one of a kind: natural necessity and laws of nature', *British Journal for The Philosophy of Science* 43, pp. 371-88.
- Braddon-Mitchell, D., and Jackson, F. (1996) *Philosophy of Mind and Cognition*, Blackwell.

Carroll, J. (1994) Laws of Nature, Cambridge: Cambridge University Press.

- Chalmers, A. (1999) 'Making sense of laws of physics', in H. Sankey (ed.), *Causation* and Laws of Nature.
- Chalmers, D. (1996) The Conscious Mind, New York: Oxford University Press.
- Clendinnen, F. J. (1992) 'Nomic dependence and causation', *Philosophy of Science* 59, 341-60.

- Clendinnen, F. J. (1999) 'Causal dependence and laws', in H. Sankey (ed.), Causation and Laws of Nature.
- Corbí, J. E., and Prades, J. L. (1999) *Minds, causes, and mechanisms: a case against physicalism*, Wiley-Blackwell.
- Crane, T. (2001a) *Elements of Mind: An Introduction to the Philosophy of Mind*, Oxford University Press.
- Crane, T. (2001b) 'The Significance of Emergence', in C. Gillet and B. Loewer (eds.), *Physicalism and Its Discontents*, Cambridge University Press.

Davidson, D. (1976) 'Causal Relations', Journal of Philosophy 64: 691-703.

- Davidson, D. (1980a) 'The Logical Form of Action Sentences', orig. in 1967, in his Essays on Actions and Events. Oxford: Clarendon Press, pp. 105-122.
- Davidson, D. (1980b) "The Individuation of Events," orig. in 1969, in his *Essays on Actions and Events*. Oxford: Clarendon Press, pp. 163-80.
- Descartes, R. (1642/1986) *Meditations on First Philosophy*. Translated by J. Cottingham. Cambridge: Cambridge University Press.
- Ellis, B. (1999) 'Causal Powers and Laws of Nature', in H. Sankey (ed.), *Causation and Laws of Nature*.
- Foster, J. (1989) 'A defence of dualism', in J. Smythies and J. Beloff, (eds.), *The Case for Dualism*, University of Virginia Press.

Foster, J. (1991) The Immaterial Self, Routledge.

Hare, R., and Madden, E. H. (1975) Causal Powers, Oxford: Blackwell.

Hempel, C. G. (1969) 'Reduction: Ontological and Linguistic Facets', in S.
Morgenbesser, P. Suppes, and M. White (eds.), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, New York: St Martin's Press, pp. 179-199.

Horgan, T. (1989) 'Mental Quausation', Philosophical Perspectives 3: 47-76.

Hume, D. (1739: various modern editions) *A Treatise of Human Nature*, Bk I, Pt IV, sect. VI, and Appendix.

Jackson, F. (1982) 'Epiphenomenal qualia', Philosophical Quarterly, 127-136.

- Jackson, F. (1986) 'What Mary didn't Know', *Journal of Philosophy*, 83, 5, pp. 291-295.
- Jackson, F. (1998) *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, Oxford: Clarendon.
- Kim, J. (1976) 'Events as property exemplifications'. References are to the reprint Kim (1993a), pp. 33-52.
- Kim, J. (1989) 'Mechanism, Purpose, and Explanatory Exclusion', *Philosophical Perspectives* 3: 77-108. References are to the reprint Kim (1993a), pp. 237-264.
- Kim, J. (1991) 'Dretske on How Reasons Explain Behaviour', in B. McLaughlin, Dretske and His Critics, Oxford: Basil Blackwell. References are to the reprint Kim (1993a), pp. 285-308.
- Kim, J. (1993a) *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.
- Kim, J. (1993b) 'The Non-Reductivist's Troubles with Mental Causation', in his *Supervenience and Mind*, pp. 336-357.
- Kim, J. (1993c) 'The Myth of Nonreductive Materialism', in his Kim, *Supervenience and Mind*, pp. 265-284.
- Kim, J. (1993d) 'Postscripts on mental causation', in his *Supervenience and Mind*, pp. 358-368.

- Kim, J. (1998) *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation.* Cambridge: MIT Press.
- Kripke, S. (1980) *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
- Levine, J. (2001) Purple Haze, Oxford University Press.
- Lewis, D. (1966) 'An Argument for the Identity Theory', *The Journal of Philosophy* 63: 17-25.
- Lewis, D. (1973) 'Causation', *Journal of Philosophy* 70, pp. 556-67. References are to the reprint Lewis (1986), pp. 159-172.
- Lewis, D. (1973) Counterfactuals. Oxford: Blackwell.
- Lewis, D. (1979) 'Counterfactual Dependence and Time's Arrow', *Noûs* 13: 455-76. Reprinted in his *Philosophical Papers*. *Vol. II*, New York: Oxford University Press, 32-52. References are to the reprint Lewis (1986).
- Lewis, D. (1980) 'A Subjectivist's Guide to Objective Chance', in R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, Volume II, University of California Press. References are to the reprint Lewis (1986), pp. 83-144.
- Lewis, D. (1983) 'Mad Pain and Martian Pain' and 'Postscript', in D. Lewis *Philosophical Papers, Vol.* I, Oxford: Oxford University Press.
- Lewis, D. (1983) 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 61: 343-77.
- Lewis, D. (1986) Philosophical Papers. Vol. II, New York: Oxford University Press.
- Lewis, D. (1986a) 'Causal Explanation', in his Philosophical Papers, pp. 214-240.

Lewis, D. (1986b) On the Plurality of Worlds, Oxford University Press, Oxford.

Lewis, D. (1998) 'Against Structural Universals', in S. Laurence and C. Macdonald (eds.), *Contemporary Readings in the Foundations of Metaphysics*, Blackwell Publishers, pp. 198-218.

Lombard, L. B. (1986) Events: A Metaphysical Study, Routledge & Kegan Paul.

- Lowe, E. J. (1996) Subjects of Experience, Cambridge: Cambridge University Press.
- Lowe, E. J. (2000) An Introduction to the Philosophy of Mind, Cambridge: Cambridge University Press.
- Lowe, E. J. (2000a) 'Causal Closure Principles and Emergentism', in *Philosophy* 75: 571-585.

Lowe, E. J. (2002) A Survey of Metaphysics, Oxford: Oxford University Press.

- Malcolm, N. (1968) 'The Conceivability of Mechanism', *The Philosophical Review* 77: 45-72.
- Menzies, P. (1999) 'Intrinsic versus Extrinsic Conceptions of Causation', in H. Sankey (ed.), *Causation and Laws of Nature*.
- Mills, E. (1996) 'Interaction and overdetermination', *American Philosophical Quarterly*, 33, 105-15.
- Nagel, T. (2000) 'The Psychophysical Nexus', in P. Boghossian and C. Peacocke (eds.), *New Essays on the A Priori*, Oxford University Press.

Oliver, A. (1996) 'The Metaphysics of Properties', *Mind* Vol. 105, No. 417

- Quine, W. V. (1969) "Existence and Quantification," in his *Ontological Relativity and Other Essays.* New York: Columbia University Press.
- Robinson, H. (1982) *Matter and Sense*, Cambridge Studies in Philosophy, Cambridge: Cambridge University Press.
- Robinson, H. (2003a) 'Dualism', in S. Stich and T. Warfield (eds.) *The Blackwell Guide to Philosophy of Mind*, Blackwell, Oxford, 85-101.

- Robinson, H. (2003b) 'Dualism', in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*.
- Russell, B. (1921) *The Analysis of Mind*. Reprinted: London: George Allen & Unwin, 1978.
- Russell, B. (1927) The Analysis of Matter. London: Kegan Paul.
- Sankey, H. (ed.) (1999) Causation and Laws of Nature. Kluwer Academic Publishers.
- Schiffer, S. (1987) Remnants of Meaning. Cambridge: MIT Press.
- Stoljar, D. (2001a), 'Physicalism' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.).
- Stoljar, D. (2001b) 'Two Conceptions of the Physical', *Philosophy and Phenomenological Research*, 62: 253-281.