

Praise for the French-language edition of *Causalité et lois de la nature*:

“... it is a pleasure to read Kistler’s book and ... its argument is very well developed. It is a remarkable example of the standards of clarity and precision that are achieved in today’s analytical philosophy of science.”

Michael Esfeld, University of Lausanne: review in *Dialectica*

“... a wonderfully rich book by Max Kistler ... not only a skilled writer in the history of philosophy; he also makes important and novel contributions both to the theory of causation and to the philosophy of laws of nature. ... This book is already packed with good arguments. My impression is that *Causalité et lois de la nature* is important reading for all philosophers with an interest in laws of nature and causation. I will certainly be on the watch both for the translation of it and for forthcoming work by Max Kistler.”

Johannes Persson, Lund University, review in *Mind*



# Causation and Laws of Nature

Causation is important. It is, as Hume said, the cement of the universe, and lies at the heart of our conceptual structure. Causation is one of the most fundamental tools we have for organizing our apprehension of the external world and ourselves. But philosophers' disagreement about the correct interpretation of causation is as limitless as their agreement about its importance. The history of attempts to elucidate the nature of this concept and to situate it with respect to other fundamental concepts is almost as long as the history of philosophy itself.

In this first English translation of *Causalité et lois de la nature* Max Kistler seeks to reconstruct a unified concept of causation that is general enough to adequately deal with both elementary physical processes and the macroscopic level of phenomena we encounter in everyday life. It will be of great interest to philosophers of science and metaphysics, and also to students and scholars of philosophy of mind where concepts of causation and law play a prominent role.

**Max Kistler** is Maître de conférences, in the Department of Philosophy at the Université Paris X-Nanterre, France and member of Institut Jean Nicod, Paris.

## **Routledge Studies in Contemporary Philosophy**

### **1 Email and Ethics**

Style and ethical relations in computer-mediated communication

*Emma Rooksby*

### **2 Causation and Laws of Nature**

*Max Kistler*

# Causation and Laws of Nature

**Max Kistler**



First published in French:  
Causalité et lois de la nature  
© Librairie Philosophique J. Vrin, Paris, 1999  
<http://www.vrin.com>

English-language translation first published 2006  
by Routledge

Published 2017 by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN  
711 Third Avenue, New York, NY 10017, USA

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

English language translation © Routledge, 2006

Typeset in Times by  
Taylor & Francis Books

The Open Access version of this book, available at [www.tandfebooks.com](http://www.tandfebooks.com),  
has been made available under a Creative Commons Attribution-Non  
Commercial-No Derivatives 4.0 license.

*British Library Cataloguing in Publication Data*  
A catalogue record for this book is available from the British Library

*Library of Congress Cataloging in Publication Data*  
A catalog record for this book has been requested

ISBN 13: 978-0-415-39859-6 (hbk)

# Contents

<i>Acknowledgments</i>	viii
Introduction	1
1 What is a causal relation?	9
2 Laws of nature and universal generalizations	75
3 Applicability conditions and the concept of “strict law”	99
4 Consequences	122
5 The nomological theory of causation and causal responsibility	139
6 Efficacious properties and the instantiation of laws	165
7 Causal responsibility and its applications	188
Conclusion	213
<i>Notes</i>	224
<i>References</i>	251
<i>Index</i>	261

# Acknowledgments

This book is a modified version of *Causalité et lois de la nature*, published in 1999, which itself grew out of my PhD dissertation from 1995, at EHESS (Ecole des Hautes Etudes en Sciences Sociales). I express my warmest thanks to my supervisor Joëlle Proust, whose generosity never failed, with both encouragement and sharp criticism. Many other people at the Institut Jean Nicod in Paris (which was then part of CREA, Centre de recherches en épistémologie appliquée) have helped me overcome difficulties, doubts and hesitations. I thank the French Ministry of Research and the Ecole Polytechnique for their financial help during the preparation of this work.

I am very grateful to Joan Cullen who has read the present English version and helped me improve it in style and clarity. The errors remain of course mine.



# Introduction

It is not very controversial to say that causation is one of the most fundamental conceptual tools we have for organizing our apprehension of the external world and ourselves. But philosophers' disagreement about the correct interpretation of causation is as vast as their accord as to its importance. The history of attempts to elucidate the nature of this concept and to situate it with respect to other fundamental concepts is almost as long as the history of philosophy.

Our endeavour can only succeed if we begin by setting ourselves a well-defined aim. The question we shall try to answer is the following: can we reconstruct a concept of causation that claims both objectivity and compatibility with a world view grounded in contemporary science? My strategy will be to answer this question in the affirmative by effecting a reconstruction of the concept of causation, articulating its presuppositions and testing its fecundity in the analysis of causal explanation, as it exists both in the sciences and in commonsense.

In spite of the enormous efforts given over to the analysis of causation, there is a persistent variance as to what may be concluded. Abandoning the hope of finding a satisfying analysis of the concept, many authors whose work is not directly focused on the concept of causation use it both extensively and with little rigour. For lack of a consensus on the "correct" interpretation of causation, several incompatible conceptions exist alongside of commonsense intuitions. Among these, three are particularly influential:

- 1) Often causation is identified with determination. This identification stems presumably from the assimilation of causation with the traditional Principle of Causality. If the Principle of Causality states that "every event has a cause", that "every determinate phenomenon has a determinate cause", or that "the same causes always have the same effects", it is apparently easy to slip from this Principle to the thesis that it is equivalent to say that the event  $e$  has a cause and to say that the fact that  $e$  has happened has been determined by the circumstances, i.e. by the set of events preceding  $e$  (or, for that matter, by the set of events following  $e$ ). According to this conception, causation is assimilated to the natural course of events, taken in a deterministic sense. To say that everything has a cause is tantamount to

## 2 Introduction

saying that nothing happens without a reason, sufficient to determine its nature. To say that all phenomena are subject to the constraint of causation means, according to this conception, not only that nothing happens by chance, but that there are no irreducibly indeterministic processes. On the contrary, on the supposition that such processes exist, e.g. in radioactive decomposition, this conception implies that the fact that the atomic nucleus  $n$  decomposes at instant  $t$  is not only not determined by laws of nature but furthermore that it has no cause.

2) The dominant doctrine in contemporary philosophy of science – a position that can safely be qualified as “classical” – differs from the position I have just sketched by acknowledging the existence of indeterministic processes. According to the deductive-nomological model, to explain (or to predict) a phenomenon is to show that there is a description of the phenomenon which is such that it can be logically deduced from a nomological statement, together with a description of the initial conditions specifying the circumstances in which the phenomenon occurred (or is expected to occur). According to the classical position, to be cause or effect is nothing other than to be subject to a law of nature. However, this conception avoids confusing causation and determination by acknowledging, within the domain of phenomena covered by laws, and therefore within the domain of causation, the existence of phenomena that are subject only to probabilistic but not deterministic laws.

An important variant of the classical doctrine consists in interpreting a phenomenon’s falling under a law as resulting from a subjective act, grounded on laws which are themselves created, not discovered, by the human mind. In this framework, the causal relation is reduced to a relation whose existence is due to an act of judgment by a subject. Without scientists trying to put order into the multitude and diversity of phenomena, it would make no sense to speak of laws of nature or of causation. The strategy of reducing causation to falling under a law – in an instrumentalist interpretation of the latter concept – has led certain empiricist philosophers of science to an anti-realist theory of causation. According to Hume, causal judgments are grounded in the psychological mechanism that transforms observations of regularities into habits and expectations. For those contemporary philosophers of science who work in the tradition of logical empiricism – whose position we have above qualified as “classical” and of which Carl Hempel is an outstanding representative – causal judgments are justified by their integration within the framework of the set of scientific judgments. This is a significant difference between Hume and Hempel: by being identified with falling under a law used in explanatory arguments, causation has retrieved its place in the scientific conception of the world.

3) The classical conception is opposed to a tradition which goes back more directly to Hume and which has in this century been defended by Carnap and Russell. Until today, many find it attractive for its simplicity. Causation is conceived of as a concept belonging exclusively to common-

sense and one that cannot possibly be replaced by any scientific concept: instead of being *reducible* to other scientifically definable notions – what the other two conceptions try to show – nothing, within a scientific world view, corresponds to the commonsense notion of causation. Therefore it should be eliminated from the scientific world view. An explanation is scientific only to the extent to which it is *not* causal.

When Bertrand Russell presented this thesis in 1912, he took it for granted that mathematical physics was the paradigmatic science and that the emergence of a unified science, which would follow the explanatory model of mathematical physics, was only a question of time. When all sciences have achieved maturity, scientific explanations will take the form of deductions of measurable quantities from functional laws and quantitative descriptions of phenomena. Today we no longer feel it is legitimate to suppose that in the near future biology and psychology will follow the example of mathematical physics. Still, it was on the basis of this conviction that Russell had hoped to ban the concept of causation from the whole of science.

This book elaborates a realist conception of causation that is opposed to all three conceptions outlined above. This realist conception vindicates the obstinate presence of the concept of causation in explanation in all the disciplines of applied science. For one can easily observe that not only in psychology and biology, but also in solid state physics and geology, many explanations are causal. This is the case each time one switches from the theoretical study of a model to the explanation of the real phenomena the model is supposed to describe. In this way I hope to show that causation can be considered as a scientifically respectable concept. This would allow us to reject Russell's exaggerated parallel drawn between the concept of causation and the English monarchy, judging that both had been allowed to survive only because they had wrongly been thought to do no harm (cf. Russell 1912, p. 173).

My conception of causation is based on a very simple hypothesis. It says that two events are related as cause and effect if and only if a determined amount of a conserved quantity is transferred from one event to another. It goes without saying that I have not developed this conception from scratch. The underlying intuition comes directly from commonsense. Only the elaboration of details will show whether the hypothesis is fruitful.

Attempting to analyse the concept of causation directly may well raise suspicion in some quarters. The long and complex history of philosophical work on this notion is well known. Contemporary philosophers of science who favour the third position sketched above think it is vain to place any hope in the theoretical fecundity of the concept of causation. Let me quote the harsh judgment Ernest Nagel has made on this topic:

“It would be an ungrateful and pointless task to canvass even partially the variety of senses that have been attached to the word ‘cause’ – varying from the ancient legal associations of the word, through the popular conception of

#### 4 Introduction

causes as efficient agents, to the more sophisticated modern notions of cause as invariable functional dependence. The fact that the term has this wide spectrum of uses immediately rules out the possibility that there is just one correct and privileged explanation for it” (Nagel 1961, p. 73).

I think it is possible to resist this judgment, although the challenge is intimidating. As Nagel rightly notes, the concept of causation has, throughout the history of Western philosophy, accumulated such a wealth of meaning that it may seem futile to try to ground modern naturalism on this concept. Nevertheless, it is still possible to construct a theory of causation which justifies the commonsense idea that causes are efficacious while also being compatible with contemporary science. I do agree, however, with Nagel’s implicit critique that the identification of every cause with an agent reveals naive anthropomorphism and I join Nagel in rejecting contemporary anthropomorphic conceptions of causation according to which human actions are paradigmatic causes, while nature outside mankind contains causal relations only in a derivative sense (see Chapter 1). Similarly, the analysis of the concept of law proposed in Chapter 2 should, if it is correct, allow us to avoid the objection to have loaded the word “cause” with “ancient legal associations”, by making the concept of causation dependent on the concept of law. We are not obliged to deny that scientific law is a conceptual descendant of human, or divine, law (see Ruby 1986; van Fraassen 1989, Introduction; Lloyd 1995), in order to claim that our reconstruction of the concept of law is free of such associations.

An important risk, run in particular by the classical position, is to see causation in *all* scientific explanations. With an eye on its first appearance in history, I suggest calling this the “Aristotelian threat”: to the extent that all explanations are considered as causal, the specificity of the concept of causation diminishes until, in the end, the expression “causal explanation” becomes a pleonasm. A concept is useful only to the extent that it expresses a relevant distinction: it is instructive to call an explanation *causal* only to the extent that there are also *non-causal* explanations. It is only to the extent that there are nomic regularities which are nevertheless not causal that the statement that a given relation is causal has any content. Otherwise the concept of causation would be redundant and we would be well advised to abandon it.

I do not intend to claim that Aristotle himself is a victim of this “Aristotelian threat”. For Aristotle, as for the pre-Socratic philosophers whom he criticizes in his philosophy of nature, the question whether nature should be explained in causal terms does not even arise. For them, the concept of causation rather provides the framework within which the debate on the best philosophical analysis of natural phenomena and their explanation takes place.

In modern times, the concept of causation no longer plays the role of a general framework for all scientific explanations. What Aristotle considered as explanations in terms of formal causes are today often considered as paradigms of non-causal explanations. This is the case of mathematical

explanations in terms of relations between numbers (Aristotle, *Physics* II, 3, 194b28; trans. 1970, p. 28), and of logical explanations in terms of the logical relation between premises and conclusion (Aristotle, *Physics* II, 3, 195a16; trans. 1970, p. 29). However, might Aristotle's causal methodology not still be valid for the explanation of *natural* phenomena? With respect to nature it might still look plausible today, that explaining is pointing out causes. Should we therefore let ourselves be guided by his principle?

We think we understand a thing simpliciter (and not in the sophistic fashion accidentally) whenever we think we are aware both that the explanation [i.e. cause: the Greek word "aitia" can be translated both as "cause" and as "explanation"] because of which the object is its explanation, and that it is not possible for this to be otherwise. It is clear, then, that to understand is something of this sort.

(Aristotle, *Posterior Analytics* 71b10, trans. 1984, p. 115)

My answer is that we should not accept Aristotle's principle: it is possible to provide a scientific explanation, in the full sense of the word, of a natural phenomenon without designating its cause or even one of its causes. The concept of causation that will be worked out in this book is much more restrictive than Aristotle's. On the basis of this new analysis, one can replace the Aristotelian methodological imperative by a "naturalistic" imperative that constitutes the background of all research on natural phenomena: for a phenomenon to be natural is to be capable of a scientific analysis. However, this analysis can more generally be nomological, i.e. grounded on laws of nature; it need not be causal.

I shall proceed as follows. I begin by elaborating a concept of causation that is both very general with respect to its domain of application and very restrictive due to its physical character (Chapter 1). This analysis is intended to do justice to causation as it reigns in the physical world. The expression "physical world" is meant to include all of nature, particularly living beings, and among them, those capable of representing their environment. This naturalist project aims at reconciling requirements that might seem incompatible: giving the word "cause" a sense that is both narrowly physical, yet broad enough to make room for the possibility of explaining causally, natural phenomena that differ radically in type and complexity. It may turn out that the framework constituted by this concept of causality is simply not rich enough to even question whether or not the interactions of perception and behaviour living beings entertain with their environment are of the same type as physical causal interactions. Even this seemingly adverse result would be important to the extent that this tentative would at least indicate which aspects of these phenomena resist a causal analysis.

Our analysis of the physical concept of causation quite naturally shows us how to proceed, as several other fundamental concepts used in this analysis must be clarified. We begin with the concept of a law of nature: In Chapter

## 6 Introduction

1, physical causation is analysed as a relation of transference between events, but the concept of transference presupposes the concept of a law of conservation, for only what is conserved can be transferred. Chapter 2 will develop an account of laws according to which they are universal relations between property instances. We shall reject the idea of a dichotomy of two types of laws differing with respect to their logical structure: the laws of fundamental physics and the laws of the so-called special sciences. True, laws of the latter type do not apply to all natural objects but only to a restricted subset of them. But the analysis of the application relation of a law to concrete situations (Chapter 3) will show that the difference with respect to the size of domain of application does not entail any difference with respect to logical structure: this structure is the same for physical laws and for laws of special sciences.

This step will turn out to be crucial for the construction of a theory up to the task we have set ourselves: construe the concept of causation so that it is general enough to be applied both to interactions between elementary particles and to an animal's or a human's perception of its environment. The complex properties of a perceived scene have a causal impact on perceiving subjects in the same sense in which the properties of elementary particles have a causal impact on those particles' interactions. Now, the dichotomy between the world of microphysics and the world of the macroscopic phenomena of our experience is often justified by the idea of a nomological difference: it is often taken for granted that only simple physical phenomena obey strict laws without exceptions, whereas more complex macroscopic phenomena are irreducibly irregular or hazardous. This is why the laws governing their evolution and interactions are taken to be non-strict or valid only "*ceteris paribus*".

I shall, on the contrary, argue that there is no reason to consider all physical laws as strict, in the sense of allowing no exceptions, nor to attribute to the laws of the special sciences the property of holding only "*ceteris paribus*", i.e. under the proviso that, between one application of the law and the next, "all things (non explicitly specified by the law) remain the same". I shall challenge this dichotomy, on the basis of an analysis of the relation between a law and its applications to particular situations. This will lead us to a new interpretation both of the origin of exceptions and of what are traditionally taken to be "*ceteris paribus* laws". We shall in particular find that the hedging or *ceteris paribus* clauses are not part of the laws themselves, but rather characterize the conditions under which a given law applies to a concrete situation. Most physical laws only apply if certain conditions are satisfied. Such physical laws are not exempt from exceptions any more than biological or psychological laws. This justifies abandoning the very idea of a "*ceteris paribus* law", to the extent that this notion gets its meaning from its opposition to the notion of a strict law that applies, without exceptions, always and everywhere. This analysis of laws has important implications, some of which will be developed in Chapter 4.

The earlier chapters in the book give us the conceptual instruments required to deal with a second aspect of causation, the notion of *causal responsibility*, grounded both on the concept of causation in terms of transference between events and on the concept of law. This is a richer and more powerful tool for the analysis of complex causal relations than the simpler causation between events. It allows us both to satisfy the requirement that every natural phenomenon must in principle yield to a causal explanation, and to take into account the fact that not all laws of nature are microphysical laws. Indeed, objects interact causally not only in virtue of their microphysical properties but in virtue of all their nomological properties. When I heat water in a closed bottle, the fact that the pressure reaches a certain point can be causally responsible for the fact that the bottle explodes. But pressure is an essentially macroscopic property: it makes no sense to attribute pressure to microphysical objects. The relation of causal responsibility brings into play all the nomological properties of a given event.

Just like the expression “law of nature”, the expression “causal responsibility” is borrowed from human law. Our reconstruction of the latter concept must be just as independent of the legal associations the expression brings to mind as the analysis of the former. The structure of the relation of causal responsibility is more complex than that of the causal relation between events. Its terms are not events themselves but rather *facts*. The fact that the cause–event has certain properties is causally responsible for the fact the effect–event possesses certain properties. Chapter 5 analyses the logical structure of the relation of causal responsibility, and the way it is grounded on the causal relation between events and on the instantiation of a law of nature in a particular situation. The account of causation in terms of transference and causal responsibility is compared to the traditional “nomological theory of causation” which reduces causation to the instantiation of a law, or to explanation in terms of such a law. We clarify the relation between causation linking events and causal responsibility linking facts by studying the patterns of inference between statements of both types, and address some important objections against the idea that both facts and events can be terms, or “relata”, of causation.

Chapter 6 offers an analysis of the nature of the properties apt to be efficacious in causation. It develops the idea that the intrinsic natural properties of events are the sources of their efficacy, where a property is “intrinsic” if it is entirely determined at the space–time location occupied by the event possessing it. A property is “natural” if it participates in a law of nature. It is also argued that the distinction between causal and non-causal is better construed as applying not to laws, but rather to the particular situations to which laws apply. We criticize one important proposal to understand causation in terms of a particular species of causal laws: Tooley’s version of the thesis that causes, by instantiating causal laws, raise the probability of their effects.

## 8 *Introduction*

Finally, Chapter 7 explores the fecundity of the concept of causal responsibility in clarifying some important problems encountered by the analysis of causal phenomena and statements reporting them. Among other things, we shall enquire whether there is something true in the statement that the cause of Socrates' death is the fact that he drank hemlock *at dawn*, although the emphasis lies on the wrong expression, whether it can be true both that an ice cube in a glass cools the water around it and that the water warms the ice cube up, whether someone's not falling can be a cause of her not dying, whether switching out the light causes the light's being out, and whether a massage can be a cause of death even if that massage is in fact beneficial and slows down the patient's dying.

It is hoped that our reconstruction of the concept of causation defends it successfully against Russell's verdict that no rational core can be extracted from this concept, purifying it from the confusion and anthropomorphist connotations it possesses in commonsense. We shall see that the concept of causation can be given a clear sense, which justifies its broad use both in commonsense and scientific explanations. Causal explanation is an important specific kind of scientific explanation. But there are also non-causal scientific explanations. Therefore, causation and explanation remain distinct though related concepts.



# 1 What is a causal relation?

The concept of causation plays a central role in many incompatible philosophical theories of nature. The aim of this chapter is to reconstruct the concept of causation without, however, pretending to do justice to the wealth of aspects the concept has accumulated due to its paramount importance.

I propose to tackle the concept of causation by taking up David Hume's analysis. This choice is promising although it may at first seem paradoxical. Hume's analysis ends up throwing into doubt the idea that causal relations are *objective*. My aim is to resist this conclusion which eliminates causation as an objective relation, and we approach this through a dialogue with Hume. Of the three necessary conditions posited by Hume and taken to be together sufficient for something to be perceived as a cause, to wit, contiguity, precedence and regular repetition, we shall begin with the last, namely regular repetition. It is indeed this condition that leads Hume to conceive of the causal relation as having a subjective rather than an objective ground. We shall examine two alternative conceptions to Hume's, which I call respectively singularist and anthropocentric. Both try to avoid Hume's conclusion although they stick to the main line of his analysis. Having found them unsatisfactory, we shall propose our own analysis of the objective causal relation: causation is basically grounded on the transference of an amount of a conserved quantity.

My proposal for the reduction of causation to a physical relation can be expressed by the following equivalence:

(S) Two events  $c$  and  $e$  are related as cause and effect if and only if there is at least one physical quantity  $P$ , subject to a conservation law, exemplified in  $c$  and  $e$ , of which a determinate quantity is transferred between  $c$  and  $e$ .

Condition (S) is intended to be necessary and sufficient for the existence of a causal relation between events  $c$  and  $e$ . It offers a reduction of the concept of causal relation albeit an a posteriori one. The thesis is that the causal relation is reducible in point of fact to the relation of transference of a conserved quantity. Still, this is no purely conceptual identity: it could be

refuted by the empirical discovery of causal relations without any transference of conserved quantities.

We shall bring out the most important consequences of this proposition by comparing it to Hume's analysis with respect to the following questions: Within a theory of causation that claims to be compatible with contemporary science, is it still correct to require that the cause be contiguous with the effect or should we rather admit the possibility of causal action at a distance? Similarly, is it still correct to require that the effect follow the cause in time or should we rather admit that "backwards" causation is a possibility that reasoning alone cannot and should not exclude? Furthermore, does our analysis allow us to retain Hume's argument against the possibility of simultaneous causation? In order to answer these questions we shall have to analyse more precisely the nature of the *terms* of the causal relation. To this end, we shall conceive of an event as the content of a space-time zone (in other words, as of what fills a space-time zone). This immediately raises the question whether we can agree with Hume when he considers that causes and effects are *objects*. Next, we shall answer an objection raising a doubt about the persistence of the transferred quantities through time. Our reply will lead to a clarification of the dependency of the concept of causation on that of a law of nature. At the same time, the fact that the concept of causation introduced here depends only on the laws of *conservation*, and the fact that our analysis *does not reduce* the causal relation to the instance of a law, show that our analysis does not belong to the category of nomological theories of causation, which constitute the most important contemporary paradigm of analyses of causation.

However, we shall see later that the analysis given by (S) is not sufficient. The concept of physical causation which is the object of the reduction statement (S) is fundamental to be sure, but it is not the only aspect of causation.

In the context of causal explanation it is necessary to mention certain *properties* of cause and effect events. To be able to do that we shall introduce the concept of *causal responsibility*. It brings into play *the fact* that the cause event  $c$  possesses a property  $F$  that is efficacious in bringing about the effect, and *the fact* that the effect  $e$  possesses a property  $G$  by virtue of the influence of the cause. However, the relation of causal responsibility depends on the relation of causation between events  $c$  and  $e$ , which will be worked out in this chapter and which is reducible according to (S). In this sense, causation between events, which we shall analyse presently, is most fundamental. Causal responsibility depends furthermore on a nomological relation between a property  $F$  of the cause  $c$  and a property  $G$  of the effect  $e$ . Causal responsibility does not directly relate the events  $c$  and  $e$ , but rather certain facts involving these events. I shall defend the thesis that causal statements expressing the concept of causal responsibility have the following logical structure:<sup>1</sup>

(O)  $C_F(Fc, Ge)$ .

(O) says that the fact that  $c$  is  $F$  (or the fact that  $c$  possesses property  $F$ , in short “ $Fc$ ”) is causally responsible for the fact that  $e$  is  $G$  (in short, “ $Ge$ ”). I shall show that (O) is equivalent to:

$$(O') C_E(c, e) \wedge Fc \wedge Ge \wedge L(F, G).$$

(O') says first that the events  $c$  and  $e$  are causally linked according to (S) (in short, “ $C_E(c, e)$ ”), second, that the events  $c$  and  $e$  possess respectively the properties  $F$  and  $G$  (in short,  $Fc$  and  $Ge$ ) and third, that  $c$ 's possession of  $F$  brings about, by virtue of a law of nature linking  $F$  to  $G$  (in short,  $L(F, G)$ ),  $e$ 's possession of  $G$ .

In most causal statements playing an explanatory role the concept of causation that is immediately relevant is the concept expressed by (O). However, as (O') shows, (O) presupposes (S), which is therefore the core of the causal relation.

### The analysis of the concept of causation by Hume

What makes one object cause another? David Hume has raised a powerful challenge to the idea that causation has an objective grounding. Any analysis of this concept that aims at giving an objective value to causal statements must answer to Hume's objections.

The first lesson Hume teaches us about causation is that it is not a quality or property of objects<sup>2</sup> but rather a relation (Hume 1955, p. 75): this is shown by the fact that everything that exists can be either cause or effect;<sup>3</sup> to judge whether something exists, there is no other means than to check its capacity of entering into causal relations with other things. No particular property can therefore be characteristic of all causes or of all effects. On the contrary, what makes a cause of something is a certain *relation* in which it stands to another thing, its effect. The relevant question is: what is the nature of this relation?

Hume's research on the concept of causation is focussed on the *idea* of causation as it exists in human understanding. One of the fundamental theses of Hume's radical empiricism is that we can achieve objective knowledge neither of the external world nor of ourselves. Thinking that the objects of perception possess an independent identity allowing them to subsist through time and to exist without us perceiving them is an illusion (Hume 1955, p. 16, p. 209–218). The same is true of the identity of the self (Hume 1955, p. 251–263).

From Hume's point of view it would therefore be vain to try to find an *objective* grounding of the relation between cause and effect. When he says that the object of his analysis is the idea of causation (Hume 1955, p. 74), the word “idea” has a psychological sense. For Hume, every idea has its origin in a perception; indeed Hume says literally that ideas *are* a kind of perception, distinguished from impressions only by the fact that they are

less strong and lively. Ideas are “faint images” (Hume 1955, p. 1) of sense impressions, and these impressions give them all the content they have.

I do not share Hume’s radical empiricist starting point. The most important difference between the framework adopted here and Hume’s is that I admit the existence of all those objects and properties which are described by scientific theories, whereas Hume restricts the supposition of existence to perceptions only (Hume 1955, p. 66–68, p. 210). Nevertheless, Hume’s analysis of the idea of causation constitutes a useful starting point for our own research on causation, precisely to the extent to which it contains a direct<sup>4</sup> and radical attack against the thesis that causal judgments have objective content.

Hume’s starts by pointing out two aspects of the causal relation, namely the *contiguity* of the cause and the effect and the *precedence in time* of the cause with respect to the effect (Hume 1955, p. 75–76). He takes the first property of contiguity between cause and effect as obvious and requiring no defence; however, he tries to provide a proof of the second property of precedence of the cause with respect to the effect. His reasoning is this. Let us admit by hypothesis, says Hume, that a cause can be strictly simultaneous with its effect. From this hypothesis we can infer that a cause that produces its effect only after some delay, is not really the (only) cause of that effect, for the fact that there is a delay during which the cause is already present, but not the effect, shows that the cause was not sufficient for the effect. Necessarily there is something else than the supposed cause which comes into play after the delay, and which brings about the effect at that moment. To admit the possibility that the effect comes about at a moment separated by a given delay, from the impact of the cause, without any new cause acting at that moment, would, according to Hume, be tantamount to rejecting the universal validity of the principle of sufficient reason. This principle is however “an establish’d maxim both in natural and moral philosophy” (Hume 1955, p. 76): it is unacceptable to suppose that the cause exists for a certain amount of time without producing its effect, but that it produces the effect at some later moment without any further intervention from outside.<sup>5</sup>

Contiguity and succession in time are not the only and, above all, not the most important ingredients making up the notion of causation. These properties are necessary but not sufficient because “an object may be contiguous and prior to another, without being consider’d as its cause” (Hume 1955, p. 77). And “contiguity and succession are not sufficient to make us pronounce any two objects to be cause and effect” (Hume 1955, p. 87). If, for example, a leaf falls on the back of a cat watching out for a mouse, and if the cat leaps precisely when the leaf touches it, it is nevertheless conceivable that the leaf has no causal impact on triggering the cat’s behaviour. We can very well imagine that the cause that makes the cat leap is its perceiving a movement at the spot where it expects the mouse to hide. In this situation, the event of the leaf’s touching the cat is contiguous and imme-

diately preceding the event of the cat's leap without thereby being its cause.<sup>6</sup>

What then, according to Hume, grounds the difference between causal relations and other sequences of contiguous events? Hume's main thesis about causation appears in the two definitions of cause that appear at the end of the chapter of the *Treatise* particularly dedicated to causation: according to Hume, what makes the sequence of two contiguous events causal is the fact that it is an instance of a type of sequence whose repetition has regularly been observed.

According to the first of these definitions – we shall come back to the second in a moment – a cause is “an object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedency and contiguity to those objects, that resemble the latter” (Hume 1955, p. 170).

Hume's position according to which causation depends on regularity can hardly be accepted by anyone who considers causation to be an objective relation. It has often been noted that we make causal judgments independently of whether they bear on the first instance of the relevant type of causal relation or not, whether it is the first instance in an absolute sense or in our personal experience, or of whether it is capable of recurring.<sup>7</sup> Hume himself admits that “ ’tis certain, that not only in philosophy, but even in common life, we may attain the knowledge of a particular cause merely by one experiment” (Hume 1955, p. 104). Indeed, the idea that the existence of one particular causal relation should depend on the existence of *other* causal relations that resemble it, seems unacceptable from an ontological point of view. It is even grammatically difficult to express this position: in the first definition of cause we have quoted, the definiendum is “a cause” in the singular, whereas the definiens is a general statement bearing on sets of similar objects. The reason for this difficulty of expression is that the definition is put in the material mode of discourse.

The distinction introduced by Carnap<sup>8</sup> between formal and material mode of discourse will prove useful for characterizing the difference between the two Humean definitions of cause. Carnap really introduces a threefold distinction. The statements of everyday discourse usually bear directly on objects, events or facts of the external world. They attribute a particular property to a particular object or set of objects. Carnap gives the following examples: “5 is a prime number”; “Babylon was a big town”; “Lions are mammals” (Carnap 1934, p. 212; trans. p. 286). He calls such sentences “object-sentences” (Carnap 1934, p. 210; trans. p. 284).

However, in theoretical and particularly in philosophical discourse, it can be useful to attribute to objects properties that Carnap qualifies as syntactical or formal. If it is not apparent from the form of the statement that it expresses the attribution of a syntactical predicate rather than a predicate expressing an objective property, the predicate and the statement are called “quasi-syntactical”: he calls such statements “pseudo-object-sentences” or

“quasi-syntactical sentences”; they are characteristic of what Carnap calls the *material mode* of expression (*inhaltliche Redeweise*). This mode is opposed to the *formal mode* of expression (*formale Redeweise*), which Carnap takes to be the correct form of expression for these statements. Statements in the formal mode of expression bear on linguistic forms. If the statement in the material mode attributes a property which is in fact syntactical to a non-linguistic object, the corresponding statement in the formal mode attributes a syntactical property to a linguistic expression designating the object. The formal mode consists in attributing linguistic properties to linguistic objects. Carnap gives the following example that allows us to compare the three modes of discourse: “5 is a prime number” is an object-sentence. “Five is not a thing, but a number” is a statement in the material mode. Its translation into the formal mode is: “‘Five’ is not a thing-word, but a number-word” (Carnap 1934, p. 212; trans. p. 286).

Hume’s first definition of a cause is put in the material mode of discourse, in this sense: it speaks about particular objects (or events) in the external world. This implies, in particular, that it does not bear on *representations* of these objects; in Hume’s terminology, on the *ideas* we form of these objects. However, the predicate Hume attributes to them and in virtue of which these objects are qualified as causes, is a quasi-syntactical predicate in the sense of Carnap’s distinction: according to the theory of resemblance put forward in the beginning of the Treatise, the similarity between different events is a relation which can strictly speaking only be attributed to *ideas* of events, but not to those events themselves. The first definition of the Treatise defines a cause as an object independently of the subject forming an idea of it. However, the main result of Hume’s analysis of causation is precisely that the existence of objective causal relations is an illusion. “The supposition of an efficacy in any of the known qualities of matter is entirely without foundation” (Hume 1955, p. 158).

In this respect the second definition is more satisfactory, to the extent that it explicitly mentions the dependence of the causal relation on the association of ideas that grows by habit in the mind of any subject who possesses the idea of causation. In Carnap’s terminology, one can say that the second definition results from translating the first definition, which was expressed in the material mode of discourse, into the formal mode. Here is the second definition. A cause is “an object precedent and contiguous to another, and so united with it in the imagination, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other” (Hume 1955, p. 172). Here the conflict between the aim of defining the concept of a particular causal relation, and the appearance in the definiens of a regularity which emerges from a large number of sequences, is resolved by a psychological interpretation: according to Hume, causation relates only the ideas of objects the mind forms during perception but not those objects themselves. However, the associations of ideas present to the mind are not subject to the constraints

of contiguity and succession. The ideas associated with the idea of a given object depend on all experiences the subject has had of similar objects. Indeed, ideas are easily associated according to their mutual resemblance (Hume 1955, p. 11). In this way, it becomes possible in the framework of Hume's conception of causation, that the conditions of applicability of the concept are different from the content of that very concept. A subject is justified in applying the concept of causation to a given sequence of events if it has had repeated experiences of similar sequences of events. The conditions of regular repetition restrict the *application* of the concept of causation, but when someone judges that *a* is a cause of *b*, this condition of applicability is completely external to the events *a* and *b*, and to the *particular relation* in which they stand to each other. The particular relation between two things to which the concept of causation can be applied must satisfy only two conditions: contiguity and temporal succession. But these objective conditions are not sufficient. A further purely subjective condition of applicability must be satisfied, which consists in a predisposition the mind has acquired from past experiences of similar successions. This predisposition is responsible for the fact that when one observes a new item of a familiar sequence in which an event of the first type is followed by an event of a second type, the idea of the second will be stronger and more lively than if the sequence had been observed for the first time. In this way, the habit to see always the same sequence produces in each instance of a new observation of a familiar sequence the impression of a "necessary connexion" (Hume 1955, p. 77) between the first and second element of the sequence.

In this way Hume arrives at the conclusion that the crucial element distinguishing the concept of causation from that of succession of contiguous events is of subjective origin. It is the idea of *necessary connection*. In other words, among all sequences of contiguous events, those are judged causal which are accompanied by the idea that the second event follows the first necessarily. However, there is nothing in any particular causal relation that could provide the objective content of the idea of necessary connection. The judgment that there exists a necessary connection between events *a* and *b* does not depend on the particular relation between *a* and *b*, but on the previous experiences *of their observer*. Hume situates the idea of the concept of necessary connection in the habit formed by the mind after having encountered a large number of similar successions (cf. Hume 1955, p. 164).

Hume's theory, according to which the crucial element that makes a succession of events causal has a subjective origin, has an aspect that has become an obstacle for the research on the nature of causation, even outside the strictly empiricist framework. Hume takes a whole set of concepts associated with causation to share more or less the same meaning, because all originate from the experience of resemblance between experiences of succession: "I begin with observing that the terms of *efficacy*, *agency*, *power*, *force*, *energy*, *necessity*, *connexion*, and *productive quality*, are all nearly

synonymous; and therefore 'tis an absurdity to employ any of them in defining the rest" (Hume 1955, p. 157). It is precisely by denying that these terms are synonymous that we can escape the conclusion that it is impossible to give the concept of causation an objective content. In this way, it becomes possible to accept Hume's thesis that the notion of necessity has no objective content, and nevertheless to reject the idea that the repetition of a causal sequence is essential for the causal nature of each of its particular instances. The possibility to construct an objective concept of causation depends on the possibility to escape a circularity that Hume judged inevitable. Science, in particular physics, has developed accounts of a certain number of the above-listed commonsense concepts that Hume considers as more or less synonymous. Thanks to their scientific elaboration, the concepts of force and of energy have for example been given a precise meaning that both distinguishes them from one another and from the other concepts on Hume's list. Now that the concept of energy possesses a precise scientific meaning which is independent of the concept of causation, it becomes possible to use it in the reduction of the concept of causation without running the risk of getting caught in a circle. We shall later try to fulfil this promise and to show that it is possible to avoid, in the determination of the concept of causation, both the condition of regular repetition and the concept of necessity.

Several strategies have been followed with the aim of constructing a realist account of causation. Such a theory can subscribe to the commonsense conception that Hume expresses only to unmask it as illusory:

What! The efficacy of causes lie in the determination of the mind! As if causes did not operate entirely independent of the mind, and would not continue their operation, even tho' there was no mind existent to contemplate them, or reason concerning them. Thought may well depend on causes for its operation, but not causes on thought. This is to reverse the order of nature, and make that secondary, which is really primary. [ . . . ] the operations of nature are independent of our thought and reasoning [ . . . ]

(Hume 1955, p. 167–168)

### **Singularist theory**

One strategy that remains faithful to Hume's empiricist spirit denies that the content of experiences that give rise to causal judgments can be reduced to the conjunction of contiguity and temporal succession. Some authors have claimed that we have a direct experience of causation that cannot be reduced to the experiences of contiguity and succession. However, the arguments in favour of this thesis do not go much beyond a simple appeal to commonsense. Here is how G.E.M. Anscombe replies to Hume's challenge to



“produce some instance, wherein the efficacy is plainly discoverable to the mind, and its operations obvious to our consciousness or sensation” (Hume 1955, p. 157–158): “Nothing easier: is cutting, is drinking, is purring not ‘efficacy?’” (Anscombe 1971, p. 93). From the fact that there is a large number of causal concepts that have objective content, such as “*scrape, push, wet, carry, eat, burn, knock over, keep off, squash, make* (e.g. noises, paper boats), *hurt*” (Anscombe 1971, p. 93), Anscombe concludes that the concept of causation itself possesses an objective content after all. She admits that this is not really an argument, to the extent that Hume tries precisely to convince us that what is specifically causal in these concepts has no equivalent in the objective relations between things, because it really results from our subjective attitude.

Anscombe observes correctly that the main obstacle to progress in the research on the objective essence of causation is Hume’s “curious belief that ‘efficacy’ means much the same thing as ‘necessary connection’” (Anscombe 1971, p. 93).<sup>9</sup> Anscombe notes that the identification of causation with necessary connection is far from obvious, especially as soon as one leaves the framework of Hume’s subjective empiricism: on condition of accepting that necessity is a phenomenon that cannot be reduced to singular facts and cannot therefore<sup>10</sup> be “discovered”, making the hypothesis that the causal relation has a necessary character means to beg the question against the idea that causation is objective.

Before Anscombe, C.J. Ducasse had already defended the idea that it is possible to attribute a meaning to singular causal judgments without requiring that their meaning be grounded in the experience of a regularity:

The supposition of recurrence is thus wholly irrelevant to the meaning of cause; that supposition is relevant only to the meaning of law. And recurrence becomes related at all to causation only when a law is considered which happens to be a generalization of facts themselves individually causal to begin with. A general proposition concerning such facts is, indeed, a causal law, but it is not causal because general. It is general, i.e. a law, only because it is about a class of resembling facts; and it is causal only because each of them already happens to be a causal fact individually and in its own right (instead of, as Hume would have it, by right of its co-membership with others in a class of pairs of successive events). The causal relation is essentially a relation between concrete individual events.

(Ducasse 1926, p. 129)<sup>11</sup>

However, Ducasse and Anscombe do not complete their well-grounded criticism of Hume’s theory by a satisfactory positive proposal.

I think that they have not found the strongest argument against Hume’s thesis that the idea of causation presupposes the idea of regularity.<sup>12</sup> This argument consists in showing that Hume has not explained in a convincing

manner why there can be true causal judgments about sequences of events that are observed for the first time. It is quite easy to find examples of such judgments. When a child observes for the first time in its life a car accident, it can make a true causal judgment about the fact that the state of motion of the cars before the collision is the cause of their deformation after the collision. The fact that she has never before observed any events of this type does not undermine her capacity to judge that the particular sequence she has observed is *causal*. In a similar way, a particle physicist who observes the collision of two new particles in a bubble chamber for the first time need not ever have observed sequences of events of the same type to be able to make a correct causal judgment. We can suppose him to be the first human to observe this causal sequence, because those particles have been created for the first time in a new accelerator. No one has ever observed the sequence of events in question: how could the formation of a habit of seeing always the same sequence ground the causal judgment bearing on this *first* sequence?

To save the thesis that every causal judgment presupposes an earlier experience of a regularity, in spite of the fact that many perfectly ordinary causal judgments bear on events that never recur, Hume offers the following explanation (Hume 1955, p. 104–105). The fact that a subject observes a large number of regular successions in general gives rise to “this principle; *that like objects, plac’d in like circumstances, will always produce like effects*” (Hume 1955, p. 105; italics in original). This is a version of the principle of causality according to which the same causes always have the same effects. Hume calls upon this principle in a slightly different form. In light of the fact we have just mentioned, namely that there are true causal judgments bearing on events that are the first occurrences of their type (or at least the first observed occurrences of their type), Hume admits that causal judgments can be grounded on the earlier observation of *similar* sequences of events. Once this principle is established, it can be applied each time a subject encounters a succession of events for the first time. However, this argument really destroys Hume’s theory of the origin of the difference between causal and non causal sequences. For if the fact that we have observed many *other* sequences that recur regularly were sufficient to judge *this new sequence* to be causal, then *all* sequences would have to appear to us as causal. Such a conclusion would be disastrous for Hume’s enterprise which is precisely to look for the specific difference that makes certain but not all sequences of events causal (cf. Hume 1955, p. 77). It turns out that the vague meaning of the word “similar” plays a crucial role here. As soon as one fixes a limit beyond which new sequences of events do not count as similar to sequences already observed, there will be clear counterexamples to Hume’s theory. There are sequences that are so radically dissimilar to everything a given individual (or mankind as a whole) has ever observed, that the theory will not be able to justify that they are judged to be causal. However, insofar as *everything* counts as similar “in a certain respect” to

something that has already been observed, the principle of causality in Hume's version is incapable in principle of establishing a dichotomy between causal and non-causal sequences of contiguous events.

Ducasse tells us that the fact that a certain causal relation occurs more than once or even regularly is inessential for the causal character of each singular relation of that type.<sup>13</sup> His proposal consists in reducing causation to a succession of changes. According to him, "*the cause of the particular change K was such particular change C as alone occurred in the immediate environment of K immediately before*" (Ducasse 1926, p. 127; italics Ducasse's). This definition places us in the situation described by Hume (Hume 1955, p. 77), before he introduces the idea of a necessary connection. If a causal relation is nothing else than a succession of two changes, how can there be successions of changes that occur for independent reasons, in other words that are not causally related?<sup>14</sup> Imagine two balls rolling over a billiard table that pass by each other so closely that they almost touch, but without any of them influencing the movement of the other at all. The situation is constructed precisely so as to make sure that there is no causal relation between the balls. However, at the moment at which the balls are at a minimum distance from each other, the movement of each of them satisfies Ducasse's conditions for being the cause of the movement of the other. The movement of ball *A* and ball *B* "occurred during a time and through a space terminating at the instant *I* at the surface *S*"<sup>15</sup> and "no change other than [...] [the movement of the other ball] occurred during the time and through the space" (Ducasse 1926, p. 127) around the encounter of the two balls. However, by hypothesis, they do not exert any causal influence on one another.<sup>16</sup>

This counterexample is enough to refute Ducasse's analysis. The movements of the two balls satisfy all the conditions set by Ducasse, but are nevertheless causally independent of each other. In point of fact, Ducasse simply tries to complete the conditions of contiguity and to succession by the condition that the cause must be *the only* change preceding the effect and contiguous to it. But being the only event present at temporal and spatial proximity has really only *heuristic* relevance: in the search for causes, it is indeed useful to try to isolate the events about which one is trying to discover whether they are causally related or not. But this strategy does not lead to a sufficient condition: Hume had already noted (Hume 1955, p. 77) that there are successions of contiguous events that are, nevertheless, not causally related. There really exist contiguous successions in which no other change occurs and which are, nevertheless, causally independent, thereby refuting Ducasse's analysis. Conversely, the fact that several causes are present at the same time and the same place does not prevent the existence of a causal relation: if two balls simultaneously hit a third ball, this does nothing to prevent the hits from being both causal. However, according to Ducasse's definition, the presence of one hit forces us to deny that the other hit is causal.

Ducasse's analysis makes the following fundamental error: it takes the heuristic method described by Mill (1843, Chapter 8) as a guide to the

search for causes, as a *definition* of causation itself. It confuses an indicator that is heuristically reliable in normal circumstances with the essence of causation. Mill's "method of differences", which is an excellent method for *discovering* causes, is nevertheless not infallible, as is shown by the example of the billiard balls that *almost* touch without influencing each other. But the mistake is above all conceptual. Ducasse's analysis takes the observable indicator for what it indicates: if a given event is the only observable change nearby, this is a good indicator that it is the cause, but it is not what makes it the case that it is the cause.

This mistake gives rise to another which Ducasse's analysis shares with a number of other theories of causation: it is the requirement that causation relates only *changes*.<sup>17</sup> However, the judgment that the stability of a bridge is the cause of the survival of the people passing over it in spite of the earthquake brings the same concept of causation into play as the judgment that the lack of stability of another bridge causes its crash on the same occasion. In both cases, the cause is a stable state involving no change; the effect is in one case a change and in the other a stable state. Furthermore, it is impossible to improve Ducasse's analysis by trying to accommodate it for the existence of stable causes. This cannot be done because the theory is based on the idea that the cause is the *only* change that contrasts with a generally stable situation. Stable conditions cannot be considered as causes for the simple reason that they are typically not isolated. As soon as *two* stable conditions are present, the logic of Ducasse's analysis does not allow us to consider either of them as the cause. Ducasse's theory implies (mistakenly) that every event has only one cause. However, there are always many stable conditions that are contiguous and preceding any given event.

Ducasse's analysis shares a third problematic aspect with Hume's analysis, and with all those that are based on the idea that the relation of cause and effect *consists* in a succession in time, in other words that succession is part of the very concept of causation. The thesis that the concept of causation implies that the cause precedes the effect in time, excludes the possibility of *simultaneous* and *backwards* causation for purely conceptual reasons. The reality and even the possibility of the latter is controversial, but many authors have held that *simultaneous* causal relations obviously exist.<sup>18</sup> Gasking (1955) offers the example of a causal relation between the fact<sup>19</sup> that a piece of iron reaches a temperature of 1000°C and the fact that the piece begins to turn red. It seems indeed plausible to suppose that there is no delay at all between the change of temperature and the change of colour, in other words the change of the quality of emitted radiation. Gasking draws the conclusion that analyses in terms of temporal succession are fundamentally defective: they are incapable of taking cases of simultaneous causation into account. Indeed, such analyses reduce the *asymmetry* between the cause and effect to a temporal relation of precedence. In a given causal relation, which of the related events is the cause and which the effect? Hume's (and Ducasse's) answer is grounded on the purely conceptual

argument presented above, that it is a priori impossible that an effect be simultaneous with its cause. Therefore it seems obvious to Hume that the event preceding the other must be considered as the cause.

But the example of the change of temperature bringing about a change of colour may throw doubt on Hume's conclusion. When none of two causally related events precedes the other, Hume's theory lacks a criterion on which it could base the answer to the question as to which is the cause and which the event. However, there clearly seems to be an objective direction to causation, even in cases apparently lacking temporal succession. It is the change of temperature that seems to bring about the change of colour of the light rays emitted by the piece of metal, not the other way round.<sup>20</sup>

### Anthropocentric theory

Gasking's own analysis, although still empiricist in spirit, abandons a central presupposition of Hume's theory of causation. Hume presupposes indeed that causal relations are discovered by *observation*. Gasking introduces the idea that causal judgments depend on a type of experience that is fundamentally different from passive observation: the experience of one's own *interventions* in the course of nature. "The notion of causation is essentially connected with our manipulative techniques for producing results" (Gasking 1955, p. 483). This proposal is intended as a conceptual analysis;<sup>21</sup> it says that every causal statement is synonymous with a statement in which the causal verb is replaced by an action verb. The meaning of the statement "*A* causes *B*" is according to Gasking very similar to the meaning of the following statement: "Events of the *B* sort can be produced by means of producing events of the *A* sort" (Gasking 1955, p. 483). Each time this is literally impossible, the notion of causation in play is improper or analogical: when it is said, for example, that during a certain geological era the melting of polar ice brought about a rise in the sea level, it is obvious that the related events are outside of the reach of human action. Gasking tries to save his thesis in the face of apparent counterexamples such as this, by saying that the truth of a statement of this type depends on the truth of *other* statements that express relations between events which are both similar to the events reported in the original statement but which are nevertheless within human reach. In the case of the melting of the polar ice caps, we must consider the statement that "one can melt ice in order to raise the level of water in a certain area" (Gasking 1955, p. 483). The fact that in this case we can actually find an analogue statement does not suffice to show that this is *always* possible. But most importantly, we lack an argument for the thesis that the truth conditions of one statement can be reduced to those of another statement by virtue of a mere relation of analogy. What would be the analogy for a statement about the causal responsibility of nuclear fusion on the surface of the sun for the heat reaching the earth? In cases in which a certain type of causal relation is completely

outside of the reach of human influence, Gasking's only way out is to say that a different kind of causation is in play, which constitutes with respect to the central notion, "a sophisticated extension from its more primitive and fundamental meaning" (Gasking 1955, p. 487). Once more, it seems that he owes us an independent argument for the thesis that the statement that "gravity makes unsupported bodies fall" expresses a concept of causation that differs from the concept expressed in the statement "great heat makes steel melt". The latter statement would express a causal relation in the strict sense whereas the relation expressed in the former would be causal only by analogy. Without an independent criterion of what distinguishes the two cases, Gasking's move seems simply to be an ad hoc justification of a counterintuitive result of the theory: that only causation due to human intervention constitutes real causation, whereas causal relations of types that are outside of human reach are ontologically derivative. Another major problem of the anthropocentric theory is its circularity. It presupposes indeed the notion of production which is fundamentally causal.<sup>22</sup>

Instead of taking up the anthropocentric analysis as an attempt to achieve an ontological reduction of causation, as some authors have done,<sup>23</sup> Georg Henrik von Wright has reinterpreted it in the framework of his research on the *epistemology* of causation. This allows him to criticize the pretension that one can draw ontological conclusions from such an analysis. Von Wright takes up Gasking's idea according to which "the causal relation can be said to be dependent upon the concept of (human) *action*" (von Wright 1973, p. 117; emphasis in original). However, he hastens to clarify his thesis by adding that "this dependence is *epistemological*, rather than ontological" (von Wright 1973, p. 117; emphasis in original). Once this obstacle has been removed, von Wright can immediately admit that "causation, needless to say, operates throughout nature independently of agency, also in regions of the world forever inaccessible to human interference" (von Wright 1973, p. 117).<sup>24</sup> In particular, once it has been recognized that the relation between the concept of causation and action is epistemological, the ubiquity of causation does not force us anymore to introduce different types of causality according to whether or not they are at least potentially within reach of our actions.

In this manner von Wright is also in a position to answer the most fundamental objection to the anthropocentric approach: human actions constitute one type of causal relations among others. What reason do we have to think that the species of human actions is ontologically more fundamental than the gender, namely causal relations in general? For this is the thesis of those who hold that the concept of causation is reducible to that of action. Von Wright does not pretend that this is the case:

No proof can decide, I think, which is the more basic concept, action or causation. One way of disputing my position would be to maintain that action cannot be understood unless causation is already intelli-

ble. I shall not deny that this view too could be sustained by weighty arguments.

(von Wright 1971, p. 74)

Von Wright's main thesis does, therefore, not bear on causation itself, but rather on the necessary conditions for us to gain knowledge of it. He notes that the mere observation of regularities in nature can suggest causal hypotheses to us, but that this is logically insufficient to give us causal knowledge, in the sense of showing us that these regularities are neither coincidences nor sequences of events due to a common cause that are not causally related directly (von Wright, 1983, p. 59). On the basis of this observation, Hume, who acknowledges only passive observation as a source of knowledge, reaches the conclusion that the difference between the knowledge of a causal relation and of non-causal correlation can only have a subjective origin. For Hume, it is the habit of the observer who encounters always the same sequences of events that makes certain sequences appear "thicker" (to use Blackburn's 1990a expression) to him than others. In their case, the experience of the cause brings about the idea of the effect before the latter is perceived. This is the source of the idea that the effect had to occur necessarily, an idea which Hume holds to be equivalent to the idea that the effect is due to a causal relation.

Von Wright accepts the Humean premise that causal relations are not revealed by immediate observation (von Wright 1973, p. 114; 1983, p. 59). He also follows Hume's thesis that what makes a sequence *causal* is its *necessity*. What is new in von Wright's theory, is his interpretation according to which this necessity is objective. A causal hypothesis has consequences bearing on counterfactual situations. It does not only tell us something about the regularities occurring in the actual world, but tells us also which effects *possible* events would have, if they occurred. Now, according to von Wright, the possibility of acting is a means of realizing possible worlds. For this reason, our capacity to act gives us the unique means to get access to causal knowledge. It is only by our active intervention in the course of events that we can "realize" a world that contains the supposed cause of a causal hypothesis. The hypotheses themselves are the fruit of observation (and of the creative mind). But they can only be tested by exploring, beyond the actual world, possible worlds; and according to von Wright, this can only be achieved by human actions.

Von Wright's theory relies on the idea of the absolute freedom of action. This is the only means to justify the idea that there is not only logical necessity but also natural necessity.<sup>25</sup> The latter constitutes, according to von Wright, the essence of causation. He takes up Hume's analysis of causation as necessary succession, but gives it a new meaning: necessity is not anymore considered as a psychological phenomenon due to habit. The necessity of causation is natural, in other words, objective. It is only *knowledge* of causation that we owe to our human particularity. In this way it

becomes possible to justify the commonsense opinion criticised by Hume, according to which causation operates throughout all of nature, independently of the intervention of agents, and in particular in regions of the world that are for ever inaccessible to man. However, the privilege of epistemic access to causation remains exclusively human because “it is *established* that there is a causal connection between  $p$  and  $q$  when we have satisfied ourselves that, by manipulating the one factor, we can achieve or bring it about that the other is, or is not, there. We usually satisfy ourselves as to this by making experiments.” (von Wright 1971, p. 72, italics in original).

As convincing as the epistemological thesis<sup>26</sup> may be, it does not touch upon the ontological issue of whether the causal relation consists in a necessary succession of events. To the extent that von Wright (1973) attributes this necessity to a causal law, of the form “necessarily, whenever  $p$  is the case,  $q$  immediately follows” (von Wright 1973, p. 113), it seems more appropriate to discuss von Wright’s ontological thesis after our analysis of laws of nature. We shall defend the following theses which are incompatible with von Wright’s theory: first, that there are no good reasons to think that there are specifically causal laws (Chapter 6); and second, that these laws do not even guarantee the strict universality of their instances, and a fortiori not their necessity (Chapter 2).

However, we are already in a position to provide an argument against the ontological thesis that causation as a general phenomenon reduces to the specific causation of human action. Consider the following reasoning: apart from the epistemological (and of course practical) importance of our experimental actions, do they constitute a particular species of causal relations or not? In other words, are human actions causal in the same sense as arbitrary natural causal relations are or not? If they are not causal in the same sense then why would we be justified in believing that they give us a privileged access to *causation* rather than to that specific relation that is human action? This would be to confuse species and genus. And if human actions are causal in the same sense as natural causal relations, what reason do we have to think that action gives us access to possible worlds, so as to allow us to evaluate counterfactual conditionals?

Consider the latter possibility: actions are causal in the same sense as other natural causal relations. In this case we cannot any longer hold that only human actions have the power of transforming possibility into actuality, because we have just supposed that human actions do not differ from other causal relations by some “specific causality”. We have therefore the following choice: if we say that *all* causal relations transform possibility into actuality, isn’t this equivalent to saying that natural causality reduces to a temporal relation? The possible would be simply that which is not yet actual, but *will be* at some moment or other. The word “actual” must here be understood in the logical sense in which it contrasts with “possible” and “necessary”. This implies in particular that it must not be given the con-



notation of the present. In this sense, the actual would be construed as consisting of the set of all *past* and *present* events. Finally, the necessary would reduce to what is *always* the case. Consider the other horn of the dilemma. If causal relations that are not actions occur only in the actual world (by opposition to possible worlds), this must also be true of actions, given the hypothesis that they do not contain any different sort of causality. As soon as one abandons the idea that action is, as a causal relation, fundamentally different from natural causal relations, one is left with no reason to believe that our actions give us privileged access to possible worlds, and therefore to the evaluation of counterfactuals.

It may be useful to insist that this criticism bears only on the ontological thesis. It is not meant to throw doubt on the explanatory value of von Wright's epistemological thesis: according to the latter, the immediate experience we have of causation in our own actions constitutes our privileged access to causation. Even if, at the ontological level, we acknowledge that *all* causal processes transform possibility into actuality, our experience as agents allows us to fully understand what that means. Before we act, we represent our actions and their results as possible, to become subsequently the most immediate witnesses of their realization. *For us*, this constitutes without any doubt the type of situation where the causal transformation of the possible to the actual appears the most clearly.

To sum up, the main problem with the anthropocentric approach to causation is that it tries to explain the essence of the general phenomenon by looking at the particular case of action. However, as von Wright himself admits, action could not play the role of the explanans if it were not a particular causal concept itself. The anthropocentric approach reduces causation in general to action, which is a particular species of causation, without providing any analysis of what makes action causal in the first place.

### **Transference theory**

As soon as one drops the dogma of radical empiricism according to which the content of the concept of causation must be determined exclusively with the help of the immediate data of experience, and one accepts that the explanans may contain scientific concepts, it is possible to gain a new perspective on the concept of causation. In this way it becomes possible to make the hypothesis that a necessary and sufficient condition for the existence of a causal relation can be found by using physical vocabulary.

Several authors<sup>27</sup> have suggested that such a condition can be found by building on the idea that the causal relation consists in the *transference* of an amount of a conserved quantity. In what follows I shall defend a particular version of this thesis: I suggest that condition (S) is necessary and sufficient for two events *c* and *e* to be causally related:

(S) Two events  $c$  and  $e$  are related as cause and effect if and only if there is at least one conserved quantity  $P$ , subject to a conservation law and exemplified in  $c$  and  $e$ , a determinate amount of which is transferred between  $c$  and  $e$ .

(S) is meant to reduce a fundamental aspect of causation to the physical relation of transference, without claiming that this reduction provides an exhaustive analysis of the concept of causation. Later in the book, I shall indeed describe a different aspect of the concept of causation, which can be found in the semantic analysis of causal statements. Then we will have to spell out carefully the exact relation between that richer aspect of causation and the physical relation described here, which constitutes its foundation.

(S) expresses an a posteriori thesis. This means that it may be refuted by counterexamples. However, to show its inadequacy, it is not sufficient to show that there are possible worlds, different from the actual world, in which causation is realized in other ways than by the transference of a conserved quantity. In order to defend an a posteriori hypothesis of reduction, it is not necessary either, as Sosa and Tooley (1993, p. 3) suggest, to characterize causation by the general role it plays in all possible worlds. On the contrary, our thesis does not go beyond the claim that *in this world* causation is realized by the physical relation expressed in (S).<sup>28</sup>

If the concept of transference were considered as a *causal* concept, hypothesis (S) would be circular and even almost tautological. However, it is possible to avoid circularity by giving the term “transference” the following meaning: an amount  $A$  is said to be transferred between  $c$  and  $e$ , if and only if this very amount is present in both events. In order for that to be possible, the space–time localization of event  $c$  relative to event  $e$  must obey to very specific constraints. We shall analyse these constraints in the following sections. Once the interpretation I have just spelled out has been admitted, the most important ontological difficulty we have to overcome in order to defend (S) is to justify that it makes sense to say of a *particular* amount of a conserved quantity that it is present in two spatio-temporally separate events. This issue will be taken up in a later section.

### ***Contiguity***

Is Hume’s concept of causation compatible with the reduction suggested in (S)? We can immediately note that the idea that causation is grounded on transference allows us to preserve one of the phenomenal characteristics of causation spelled out by Hume, namely contiguity. Transference is a continuous process in the course of which the amount of conserved quantity at issue crosses the spatio-temporal distance between the cause event and the effect event at a finite speed. This means that, if cause and effect are not immediately contiguous themselves, the transference process guarantees the existence of a causal chain whose links are contiguous.

We have, however, a considerable advantage over Hume, to the extent that we consider it legitimate to rely on scientific notions and knowledge whereas Hume's radical empiricism obliges him to reason exclusively in terms of the content of immediate experience. As far as contiguity is concerned, Hume disposed of no argument to exclude the possibility of action at a distance. He treats the idea that causal influence can act only by contact as an obvious truth. Today, we can count on the scientific result according to which all fundamental forces act at a distance only by virtue of a propagation that spreads at a finite speed. This is true in particular of gravitation.<sup>29</sup> As to the mechanism of the action of the fundamental forces at a distance, physics gives us solid arguments for thinking that it is always the result of underlying causal processes that obey strictly to contiguity.<sup>30</sup> However, we have no a priori argument against the possibility that there are nevertheless, in point of fact, instantaneous actions at a distance. We have no means to decide, by purely conceptual means, the question of whether there are non-local correlations that refute the condition of contiguity.

Einstein *et al.* (1935) have shown that quantum mechanics predicts the existence of so-called entangled systems which have the following intriguing property: some of the system's quantities are such that, if they are measured simultaneously at two spatially separate points of the system, the value measured at one point depends on the value measured at the other. However, such correlations are nothing spectacular in themselves: after what has been said so far about entangled quantum systems, the correlations existing at a distance might well be the consequence of the local propagation of causal processes. To distinguish the quantum mechanical phenomenon (called the "EPR" effect, after the initials of the three authors who first described it) from more usual situations where this is indeed the case, consider an ordinary coin split into two pieces, such that "heads" is on one half and "tails" on the other.<sup>31</sup> Imagine now two persons: one puts the "heads" part and the other the "tails" part in their pockets, then they walk away from each other. Imagine this happens so that neither sees which half he has in his pocket. In these conditions, when one of them looks at the half he has in his pocket, he immediately knows which half the other person has in his pocket, whatever their distance from each other. This is a case where a correlation at a distance is created by perfectly local causal processes, beginning with the splitting of the coin and ending with the discovery of which coin each has in his pocket. It is only because, the distribution of the halves being unknown before the moment at which the first of the two people looks into his pocket, this person gains knowledge about something that is the case at a certain distance, and this happens so to speak simultaneously.<sup>32</sup> Nevertheless the state of affairs itself had been determined before and independently of its discovery by one of the two persons. Its determination is due to causal processes that propagate locally.

The case of a correlation existing within an entangled quantum system extended in space is different. The situation that shows the closest analogy

with the case of the split coin is a pair of particles exemplifying a quantity that can take only two values. An example of such a quantity is *spin*, more precisely its component in one spatial direction. Spin is a fundamental quantity quantum physics attributes to elementary particles. The spin of the most important particles that compose matter, electrons, protons and neutrons, can take the values  $+1/2$  and  $-1/2$ . If two particles of this type interact and then speed off in different directions, their states are correlated in such a way that the result of a measurement of this quantity, taken on one of the particles, determines the result of a measurement of the same quantity, taken on the other particle. The fundamental difference between the quantum case and the classical case is that in the former, the state of particle *A* is objectively determined only at the moment at which the measurement is performed on particle *B*. It is not only the information, the observer's knowledge, that reaches its object simultaneously at the distance; rather, the objective physical state of the particle is determined instantaneously by an act of measurement that takes place at a certain distance from it. To the extent that this result of quantum theory seems to imply the existence of causal action that produces its effect instantaneously at a distance, Einstein, Podolski and Rosen concluded that "the quantum-mechanical description of physical reality given by wave functions is not complete" (Einstein *et al.* 1935, p. 780). Later, several authors, in particular Bohm (1951), have worked out an alternative to what is now called "classical" quantum theory, which postulates the existence of so-called "hidden variables" that allow us to reduce in a sense the case of the quantum correlation of the EPR type to the case of the classical correlation of the type of the split coin. The value taken by the hidden variables determines the result of the measurement taken on both particles, *before* we gain knowledge of it. The objective state appears therefore, just as in the classical situation, as independent of the knowledge we have of it. However, it must be added that this conception allows us to consider the situation as equivalent to the classical case only "in a sense", because the new theory is, although deterministic, "non-local". Within a quantum system, the evolution of a part of the system that is localized at a point *X*, depends on the evolution of other parts of the system that are localized at points *Y* distant from *X*. Furthermore, Bell (1964) has shown that this non-locality is a property that every deterministic alternative to classical quantum theory must *necessarily* process, in order to reproduce certain fundamental results of quantum theory. I should add, however, that the nature of the non-locality that Bell's (1964) result imposes on every acceptable quantum theory is far from being completely understood.

In this situation, the defender of the transference theory can argue that the correlation at a distance shown by the EPR phenomenon is not an instance of causation: according to classical quantum theory, within a system occupying a single quantum state, the result of a measurement taken on one part of a system localized at *X*, determines the result of other mea-

surements, even if these measurements are taken simultaneously at parts of the system localized at  $Y$  at some distance from  $X$ . According to Bohm's (1951) alternative theory, the determination at a distance is not due to the act of measurement; rather, it is conceived as objective and independent of our intervention and knowledge. However, both conceptions leave open the possibility that this is a case of *non-causal determination*.<sup>33</sup> This seems particularly plausible in light of the fact that the phenomenon in question does not allow us to *act* at a distance, in the sense that the result of a measurement taken at a distance, say at point  $Y$ , is not under the influence of our will. To do justice to the EPR effect, it is enough to suppose that this result and the result of the measurement taken at point  $X$  are *correlated*. It is impossible to exercise a voluntary influence on the measurement taken at point  $Y$ , by interfering with the system at point  $X$ . In agreement with our thesis, nothing seems to be transferred between  $X$  and  $Y$ , so that the events that are measured at  $X$  and  $Y$  are not causally related. Measurement shows the existence of a non-local correlation, but this does not threaten our thesis that causal interactions are always local.

The scientific debate on the explanation of this phenomenal is not yet over (cf. Scheibe 1991; Albert 1994). So far, nothing prevents us from thinking that quantum mechanical correlations at a distance do not give rise to causal relations, but are a form of non-local and non-causal determination. Therefore, we can keep claiming that condition (S) is both sufficient *and* necessary.

### Succession in time

As for the issue of succession in time, the situation is quite complex. Dowe (1992a) argues that Fair's (1979) version of transference theory encounters two important problems. Dowe's first objection to transference theory is that it makes a *causal theory of time* impossible. Dowe calls Fair's (1979) version of transference theory "the temporal theory" (Dowe 1992a, p. 182) of causation, because it reduces the asymmetry between cause and effect to the asymmetry of time. Dowe considers this to be a weakness, because a theory that explains the asymmetry of causation by the spatio-temporal asymmetry of cause and effect events makes a causal theory of time impossible. According to Dowe, the fact that the theory cannot provide a framework for the explanation of the direction of time on the basis of the asymmetry of causation, "endangers the objectivity of the account" (Dowe 1992a, p. 181; cf. Dieks 1986, p. 95; Sosa and Tooley 1993, pp. 3–4).

Dowe (as well as Ehring 1986) also objects to Fair's theory that it excludes for purely conceptual reasons, the possibility of causal relations that are oriented "backwards" in time, in the sense that the effect precedes the cause. In fact, it seems preferable to consider that it is an empirical issue whether there are causal relations in which the temporal order of cause and

effect is switched. The conceptual framework of the theory should therefore not exclude that possibility a priori.

It is true that Fair's (1979) analysis is unsatisfactory on both counts. We shall come back later to its incapacity to take into account the possibility of the existence of causal processes evolving "against" the direction of time. Let us begin with the first point raised by Dowe. To explain the asymmetry of the causal relation, Fair makes use of the notion of an absolute time providing a temporal order for all events, independently of the nature of the physical processes linking them. "I am [ . . . ] quite content to take time and its ordering of events as primitive in this attempted reduction of causation to physics." (Fair 1979, p. 241) Fair also acknowledges that this makes it impossible to explain the temporal order itself by the causal order. However, the hypothesis of an absolute order of time is not the only way of explaining the asymmetry of causation, within the framework of a theory of causation in terms of transference. To overcome Dowe's objections, it is sufficient to show that it is possible to explain the asymmetry of causation within the framework of transference theory *without presupposing* that time has an intrinsic direction. This can be done by the following analysis.<sup>34</sup> We begin by considering time as one of the dimensions in which events and processes are located. Just as objects are "in" space, events and processes are "in" space-time. Then we suppose that time, as a dimension, is isotropic just as space is.<sup>35</sup>

The origin of the asymmetry of causal processes lies in the existence of a special sort of processes that are intrinsically asymmetrical and transmit this asymmetry to all other causal processes. Dowe (1992a) himself has contributed to developing this idea, originally due to Reichenbach (1956).<sup>36</sup> It is based on two premises. First, all causal relations are part of a single global causal network; second, there are particular processes possessing an intrinsic asymmetry (Figure 1.1).

The structure of the network imposes on all causal relations it contains the constraint of being "aligned", in the sense of being asymmetrical in the same direction. The whole network is asymmetrical in the same direction as some of the individual processes it contains. In Reichenbach's words, "we [ . . . ] can construct a *causal net* which, as a whole, has a *lineal*<sup>37</sup> order. This term means that, if a direction is assigned to one line [within the net], a direction is determined for each line. If we reverse the direction for the first line, all other line directions are reversed." (Reichenbach 1956, p. 36; italics in original). If the network contains at least one particular causal relation possessing an intrinsic direction, this direction can be considered as the direction of the whole network. This account presupposes the existence of particular causal relations possessing an intrinsic asymmetry. The most important processes of this sort are isolated systems whose global entropy increases. The second law of thermodynamics expresses the relevant asymmetry: the global entropy of isolated<sup>38</sup> systems can only *increase*. All spontaneous changes in the global entropy of such systems are increases.<sup>39</sup>

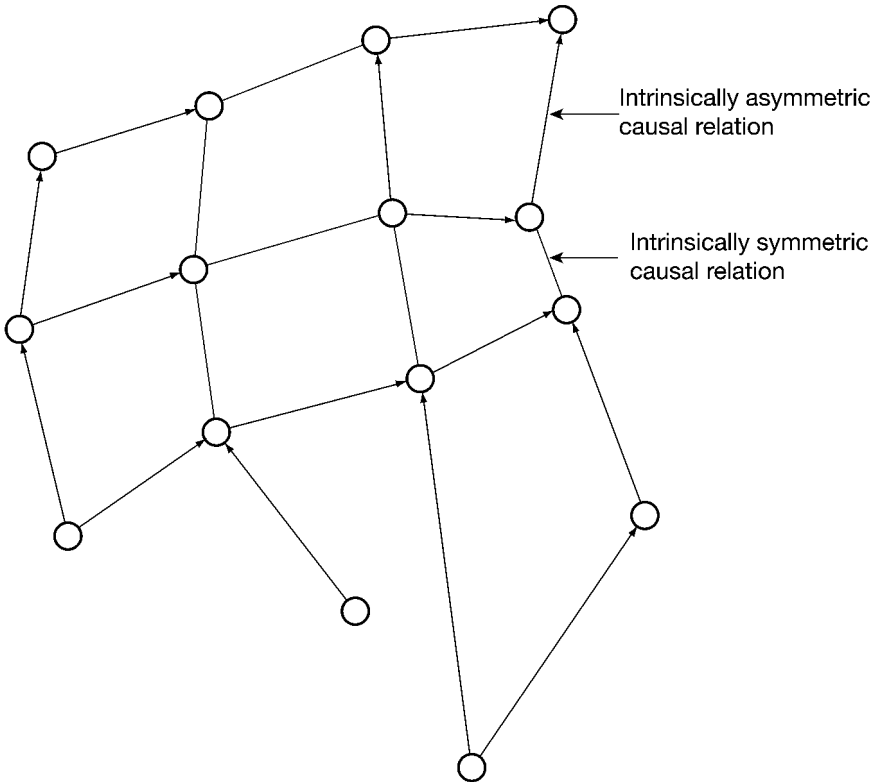


Figure 1.1. Schematic causal network  
 (after Reichenbach 1956, p. 37; with modifications)

Within the framework of the causal network, it is possible to determine which of two causally related events is the cause and which the effect, without already possessing a criterion of the direction of time according to which one occurs earlier than the other. It suffices to find somewhere in the network an intrinsically asymmetrical process and to transpose the direction defined by it on the whole temporal coordinate. Even if there were only few intrinsically asymmetrical processes, this would suffice to determine the global direction of all causal processes. It so happens that there is a large number of irreversible processes, namely all processes with spontaneously increasing entropy. In point of fact, the increase of entropy of all these processes points in the same temporal direction. Each of these processes allows us to define a permanent direction of the temporal coordinate. This direction can then be qualified as “positive”. Nothing stands in the way of considering this positive direction as “the direction of time” – nothing, provided one remembers that this is just a convenient way of speaking. One should bear in mind that time as a dimension has strictly speaking no

direction. If another empirical condition is satisfied (to which we shall come back in a moment), it is possible to consider without contradiction that all causal processes are oriented in the same direction.

Once the positive direction of the temporal coordinate has been fixed on the basis of the intrinsically asymmetrical processes, one can determine the direction of any process, even if it is reversible and intrinsically symmetrical with respect to time. In this way, causal relations which are not intrinsically asymmetrical, are nevertheless indirectly asymmetrical by virtue of the existence of *other* processes that are intrinsically asymmetrical. We can say that these relations are asymmetrical by virtue of the asymmetry of the spatio-temporal *localizations* of the cause and effect events, given that the mediation is accomplished by the determination of a positive direction of the temporal coordinate.

In a universe in which there are intrinsically asymmetrical processes, the possibility to provide an objective determination of the direction of time, implies that for a given punctual event *O*, there is an objective difference between the light cone representing the (possible) future, relative to the spatio-temporal position of *O*, and the light cone representing its past (Figure 1.2).

The spatio-temporal position of the causal processes taking their origin in an event *O* situated at the centre of both cones (the cause) must lie within one of these cones: the cone containing the future of the cause event. Processes starting at *O* can only lead to events *E* lying within the “future cone”. The asymmetry appears in the fact that *O* cannot be the effect of causal processes taking their origin at *E*. We should note however, that this conclusion depends on the premise that all intrinsically asymmetrical processes of the universe point to the same direction. We have no a priori certainty that this is indeed the case.

Therefore, before we look at the implications of the fact that our theory does not exclude the existence of asymmetrical processes pointing in opposite directions, we must insist on the following point. According to our explanation of the origin of the asymmetry of causation, this asymmetry is not *constitutive* of causation. First, this aspect gives our theory a greater explanatory force than Hume’s: for Hume, the fact that causes precede their effects is simply a constitutive aspect of the causal relation; but this means that the asymmetry has no explanation.<sup>40</sup> Second, our account makes it appear as a contingent fact that all known causal processes are oriented in the same direction, a direction which we can therefore consider as the future. It is no a priori truth that causes precede their effects; in other words it is no consequence of the concept of causation. Rather, it is a contingent fact, about which we may be mistaken.<sup>41</sup>

How can we conceive of the possibility of *symmetrical* causation? Take a causal process that is not intrinsically asymmetrical. Any process obeying the laws of classical mechanics can serve as an illustration, for example the movement of the earth around the sun. Our theory allows us to explain the fact that this is a causal process, without thereby attributing an intrinsic direction to it. Without having fixed the positive direction of the temporal



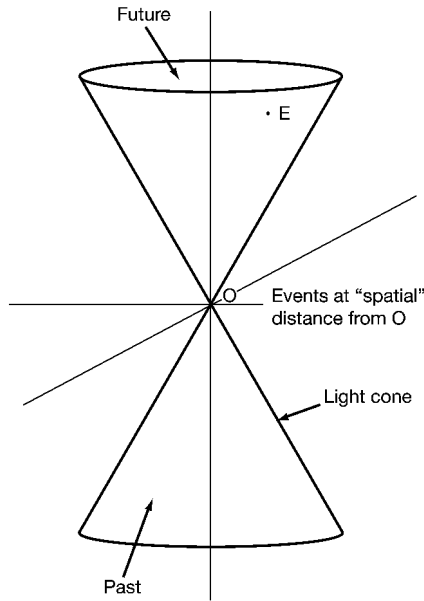


Figure 1.2. Light cones  
(after Anderson 1967, p. 154)

coordinate, in other words in abstraction of the existence of a causal network containing some asymmetrical processes, that celestial movement possesses no temporal asymmetry. However, any two spatio-temporally adjacent events contained in this symmetrical process are related by the transference of energy, mass and angular momentum. Therefore, transference theory allows us to consider temporally adjacent states of the system as causally related, without one being the cause rather than the effect. We can conceive of a possible world in which there are no intrinsically asymmetrical processes: there would not be any intrinsic difference between cause and effect either. There would be causal relations without any privileged direction. This is a conceptual possibility, in the sense of being compatible with the physical concept of causation. To be sure, it is only a possibility, for the actual world is such that it contains asymmetrical processes allowing us to define a positive direction of time: this direction can be defined as the direction in which the entropy of isolated systems increases.

We have reached the result that the asymmetry of causation is in general not an intrinsic asymmetry: it is intrinsic only for a particular class of processes; as for the other processes, they are asymmetrical by virtue of their causal connection, sometimes indirect, with processes of the particular sort which are intrinsically asymmetrical. However, this does not make the causal character of any processes indeterminate: we have said that time must be considered as a coordinate perfectly analogous to the spatial coordinates.

Then according to the criterion of transference, two events located in space and time are causally related if and only if an amount of a conserved quantity is transferred between them. It appears clear that the direction of transference is not independent of the direction imposed on the causal network and on time by the asymmetrical processes. If we imagine this direction inverted, the event from which the amount originates becomes the event at which the amount arrives, and vice versa. However, I don't think that this provides an argument against transference theory.<sup>42</sup> It only means that the asymmetry between cause and effect is no part of the concept of causation; the asymmetry of time and causation is rather a particular and contingent property of the actual world.

As for the conceptual *possibility* of causal relations whose direction is opposed to the direction of time, we can draw the following conclusions: the temporal and causal direction of an intrinsically symmetrical process is determined only indirectly on the basis of the intrinsic direction of other processes to which it is linked within the causal network. However, it is not a priori certain that it is possible to obtain in this way a unique direction for all processes. First, it remains conceptually possible that some parts of the universe are causally totally isolated from the rest: according to our theory, such an isolated part would also be *temporally* isolated. Either it contains itself asymmetric processes, so that it is possible to define a temporal direction in it. Or all processes occurring in it are symmetrical; then it would contain causal relations, but neither time nor the causal relations would be oriented in any temporal direction. There would be no objective reason to consider one of two causally related events as the cause rather than the effect. Intuitively, this consequence may seem hard to swallow. But it is hardly more than a conceptual possibility, for the following reason: according to present physical theory, there is no shield against the propagation of gravitation. If there is a region that is causally and temporally isolated from ours, it would have to be so far from us that the gravitational attraction spreading out from the mass of our region of the universe has not yet reached it.

There is a second possibility that is perhaps less fantastic. Once again, it is a possibility in the sense of being compatible with the physical concept of causation. Whether it is in fact realized is an empirical question that philosophy has no means to answer. It is conceivable that our world contains a particular set of intrinsically asymmetrical processes whose direction is *opposed* to that of the other asymmetrical processes (Figure 1.3). Such hypothetical processes evolving in the temporal direction opposed to the direction of the other asymmetrical processes, can be considered to evolve "backwards" in time.<sup>43</sup> To be able to say not only that they have their own time, the direction of which is opposed to that of normal time, but also that they evolve towards the past, we must accept the following premise. We must suppose that the number of processes evolving in the opposite direction is much smaller. This gives an objective justification to the idea that the direction shared by the large majority of asymmetrical processes can be

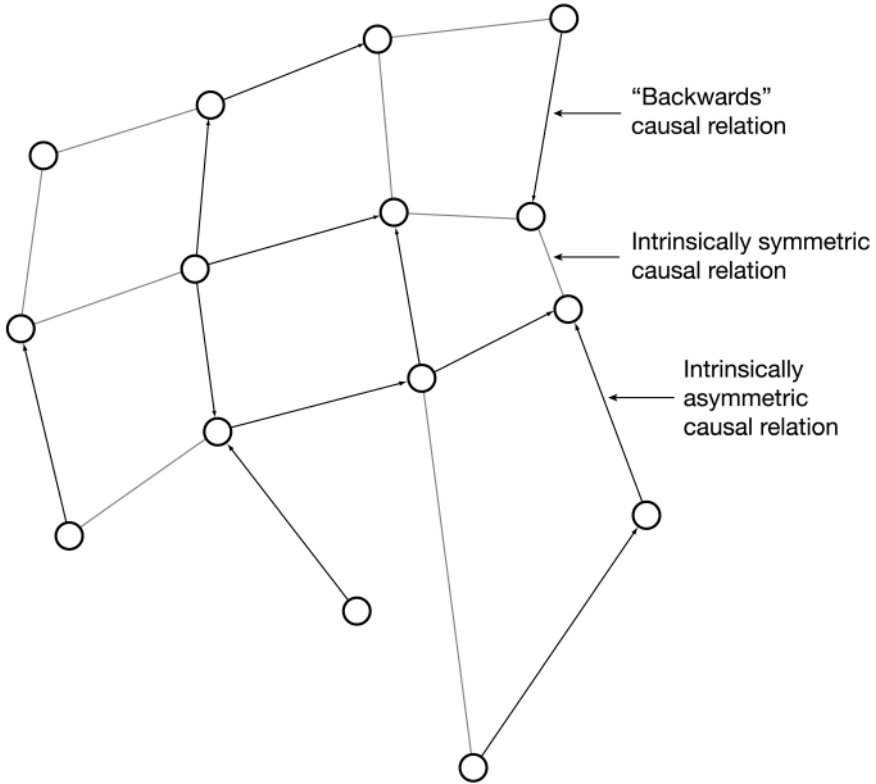


Figure 1.3. Schematic causal network with “backwards” causal relations (after Reichenbach 1956, p. 37)

considered as “the direction of time”. More precisely, we can then make the following terminological decision: we can decide to follow commonsense and say that by definition there is only one time, which is identical with “majority time”, defined by the direction of the majority of asymmetrical processes. This allows us to maintain the idea that there is a “main” direction of time. We can say that, with respect to that time, the hypothetical backwards processes evolve “towards the past”. Even if some cases of backwards causation were discovered, this would not cast any doubt upon the fact that the large majority of causal relations are oriented in the same sense, which constitutes the “future”.

The transference theory of causation proposes a reduction, albeit an a posteriori one, of the causal relation. Tooley (1990) has offered two arguments intended to show that reductionist theories of causation are generally incapable of accounting for the fact that causal relations have a well-determined direction. The foregoing reasoning gives us the means to reply to these objections.

According to the first argument, “there are possible causal worlds to which reductionist accounts of the direction of causation do not apply” (Tooley 1990a, p. 223). Tooley begins by constructing a possible world that is so simple that it does not contain any sources of asymmetry at all. He imagines a world containing only two particles, subject only to the force of gravitational attraction and being in a state of uniform rotation one around the other. For lack of any other source of temporal asymmetry, Tooley argues, the reductionist has only one possibility to justify his idea that the causal relations in this simple world have a well-determined direction. He must make use of an objective direction of time without being able to reduce that direction in turn to physical properties. It is true that this would mean to acknowledge that it is impossible to explain the origin of the fact that causal relations have a direction, in other words the fact that in every causal relation between two events the distinction between cause and effect has an objective ground.

However, we need not accept the unsatisfactory position just sketched. Indeed, we have already considered, in a slightly different form, a state of affairs quite similar to that imagined by Tooley. We have considered the possibility that our actual world contains a region that is 1) causally isolated from the region we inhabit, and 2) contains no intrinsically asymmetrical processes whatsoever. Tooley’s thought experiment proceeds by imagining a possible world which consists in its totality in such a region. Our theory implies the following analysis of this possible world: it contains causal relations, but they have no determinate direction. There are transferences of energy, but no asymmetry determines which of two events related by such a transference is the sender and which is the receiver.<sup>44</sup> We cannot, and we do not, help ourselves to the existence, in that possible world, of a determinate direction of *time*. The account sketched above in terms of intrinsically asymmetrical processes in a causal net, is meant to reduce, not time, but its direction. Tooley’s simple universe that contains only two particles is in time just as it is in space. However, the two directions of the temporal coordinate are exactly equivalent. There is neither a privileged direction of time, nor of causal relations. Tooley’s argument gives as no reason to think that causation must have a determinate direction in all possible worlds. What it shows is that there are possible worlds in which there is no physical asymmetry to ground the asymmetry of causation. However, nothing prevents us from concluding that, in such worlds, neither causation nor time has a direction.<sup>45</sup>

Tooley’s second argument aims at showing “that there are other possible causal worlds for which reductionist accounts yield wrong answers with respect to the direction of causal processes” (Tooley 1990a, p. 223). To this end, Tooley constructs a possible world in which there are two “universes” causally isolated from each other. Both are rich enough to contain irreversible processes exhibiting changes of entropy. Furthermore, the two universes are exactly identical except for the direction of their evolution. The

initial state of one is exactly identical to the final state of the other and vice versa. During their evolution, each instantaneous state of one is matched by an instantaneous state of the other, the only difference being that the direction of movement of all particles is inverted in one with respect to the other. Tooley now reasons as follows: in one of the two universes, the direction of causation corresponds, just as in the actual world, to the direction in which the entropy of isolated systems increases; however, in the other universe these two directions do not coincide, and causation is oriented in the direction opposed to that of the increase of entropy. At this point, Tooley concludes that the direction of causation cannot be reduced to the direction of physically asymmetrical processes, because the imagined possible world contains a universe in which these directions are opposed to each other.

It would seem, however, that this argument begs the question: Tooley supposes from the beginning that the directions of both time and causation are already fixed in both universes. However, on a reductionist account of the directions of causation and time, these directions, in a given world or region of the world, are precisely not determined independently of the physical processes occurring in that world or region. Either it contains asymmetrical processes, in which case these processes determine the directions both of time and of causation, or it does not contain any such processes, in which case there is no privileged direction. The question-begging premise is that the directions of time and causation are already determined independently of the properties of asymmetrical processes. As soon as we drop that premise, it is easy to provide a reductionist analysis of the possible world described by Tooley. The two universes being causally isolated, we must consider the question of the direction of causation and time independently for each of them. According to our hypothesis, each universe evolves in the direction determined by the increase of entropy of isolated systems.<sup>46</sup> The air of paradox disappears from the possible world imagined by Tooley as soon as we note that it is impossible to *compare* the directions of evolution of the two universes; by virtue of the fact that these universes are causally separated, it makes neither sense to say that the direction of causation is the same in both nor that it is inverted in one with respect to the other. To say this, as Tooley does, means to stipulate, against our hypothesis, that there is an absolute direction of time, which allows us to compare the respective directions of causation in the two universes with respect to it.

As we have seen, the transference theory allows us to make sense of the possibility of causal relations whose direction is inverted with respect to the majority of causal processes. However, it is true that it makes a “causal theory of time” impossible, because it makes impossible the reduction of the direction of time to that of causation. However, I have argued that this is not a defect of the theory, because there is a satisfactory alternative account of the relation between time and the direction of causation. Time itself is isotropic just as space; the intrinsic asymmetry of a certain number of processes

and the fact that there is a global network within which all events are causally connected to each other, make it possible to single out a positive direction in the temporal dimension. The asymmetry of causal processes has the same origin as the asymmetry of time: a certain number of them are intrinsically asymmetrical, both temporally and causally, and all others possess a derived asymmetry, by way of the network linking them to the intrinsically asymmetrical processes.

It has been objected to transference theory that it excludes the conceptual possibility of backwards causation.<sup>47</sup> However, this is in fact only true of the particular version of the theory defended by Fair (1979). True, Fair's version of transference theory excludes the possibility of causal relations in which the effect precedes the cause and of processes evolving from the future towards the past. However, this is due to the fact that Fair abandons the main idea of transference theory in order to be able to account for a particular type of commonsense intuition. In some situations, the analysis in terms of transference conflicts with commonsense judgments. If we put an ice cube into water at time  $t$ , this "causes" a cooling of the water. At time  $t+\Delta t$ , the surrounding water will be cooler than at time  $t$ , and the water contained in the ice cube will on average be warmer than at time  $t$ . In other words, part of the cube will not have  $0^\circ\text{C}$  as at time  $t$ , and will have melted. However, in this situation the direction of transference of energy is from the surrounding water to the ice cube. By cooling the surrounding water loses part of its energy to transmit it to the frozen water of the cube. A certain amount  $\Delta E$  of the energy contained in the water at time  $t$  is transferred to the ice cube, so that at time  $t + \Delta t$  this energy  $\Delta E$  has been transformed into kinetic energy of the molecules of the ice cube. There is no doubt about which of the events that take part in this interaction is the cause and which the effect because the transference is realized by an irreversible and therefore intrinsically asymmetrical process. Therefore, there is good reason to say that the state of the water surrounding the ice cube at time  $t$  is the cause, and the state of the water contained in the cube at time  $t+\Delta t$  the effect. However, Fair considers that it is preferable to justify the commonsense judgment that the ice "causes" the cooling of the water. According to him, this constitutes sufficient reason to introduce the following "emendation" into transference theory:<sup>48</sup> if there is a transference of energy between  $A$  and  $B$ , the direction of this transference does not determine the direction of causation; in other words it does not determine which of  $A$  and  $B$  is the cause and which the effect. According to Fair's version of the theory, the cause is the object  $A$  that is the source of transference, "*unless* the physical energy-momentum redescription of  $A$  is a description of an energy *gain* and  $B$ , an energy *loss*" (Fair 1979, p. 243; italics in original). This allows him to say that the ice cube whose energy increases, in other words towards which the energy is transferred, is nevertheless the cause, and that the surrounding water that loses energy is the effect. However, the price to pay for this concession to

commonsense is that the theory loses its capacity to account for the objective asymmetry of the causal relation. But there is worse: notwithstanding his intentions, Fair cannot even account for the causal judgments of commonsense. Commonsense acknowledges without contradiction the coexistence of two simultaneous causal relations between the same terms but in opposite directions: the surrounding water causes the melting of the ice cube by heating it up and the ice cube causes the cooling of the surrounding water. Commonsense judges both that *A* causes *B* and that *B* causes *A*. Fair tries to account for the possibility of such apparently contradictory causal judgments by saying that it depends on the perspective taken by some observer which of both is true and which is false. According to Fair, there is only one transference of energy and therefore only one causal relation; but its direction is objectively indeterminate and depends on the subjective perspective one takes on the relevant events. For one observer, the direction of causation is that of the transference of energy, but for another it is the opposite.<sup>49</sup>

Dieks (1981) has pushed Fair's reasoning one step further: according to Dieks, the case of the causal relation between the melting of the ice cube and the cooling of the surrounding water provides a strong reason to think "that we must not conceive of causation as an ontological category, but rather must see it as an *explanatory concept*" (Dieks 1981, p. 106). Explanation has precisely this property that it depends on the decision of the person who is constructing the explanation, whether she considers as the explanandum the fact that the ice cube melts (by warming up) or the fact that the surrounding water cools down. In this sense, the logical direction of explanation, from explanans to explanandum, is independent of any objective asymmetry between the two facts connected by the explanation. However, this is precisely one of the main reasons for *not* identifying causation with a type of explanation, and for *not* trying to reduce causation to explanation.<sup>50</sup>

### **Simultaneous causation?**

Making reference to Aronson's (1971a) version of transference theory, Ehring (1986) says that the "notion of transference will not do the job of determining causal direction in [ ... ] cases [ ... ] of [ ... ] simultaneous and backwards causation" (Ehring 1986, p. 250). In the preceding section, we have rejected this judgment for the case of backwards causation. We have seen that it is possible to reconcile transference theory with the possibility of the existence of causal relations in which the effect precedes the cause. But what about *simultaneous* causation? If there are causal relations in which none of two causally related events precedes the other, this constitutes another major obstacle for transference theory. That theory indeed implies that, independently of the *direction* of causation, cause and effect cannot be simultaneous, on the basis of two plausible

premises. First, the notion of transference implies that cause and effect must be distinct events; second, every transference between two spatially separated events takes time.

We have also seen (in the first section of this chapter) that Hume tries to establish the impossibility of simultaneous causation by a purely conceptual argument or, in his own terms, “by a kind of inference or reasoning” (Hume 1955, p. 76).<sup>51</sup> We have not yet analysed that argument. According to Hume, if there were simultaneous causation without any delay separating the cause from the effect, then non-simultaneous causation would be inconceivable. That this conclusion is unacceptable is the basis of the *reductio ad absurdum* of the hypothesis that there exist cases of simultaneous causation. Hume’s argument makes use of the principle of sufficient reason according to which all events happening to an object are determined by factors external to the object. In other words, the principle says that there are no “spontaneously” occurring events: “ ’Tis an establish’d maxim both in natural and moral philosophy, that an object, which exists for any time in its full perfection without producing another, is not its sole cause; but is assisted by some other principle, which pushes it from its state of inactivity, and makes it exert that energy, of which it was secretly possess’d” (Hume 1955, p. 76). Hume begins by asking us to suppose that there are both cases of simultaneous causation and of non-simultaneous causation. Non-simultaneous causation not being the simplest relation between cause and effect, it must be considered as resulting from the combination of a relation of simultaneous causation and a period of delay. However, the only way of conceiving of this combination leads according to Hume to a violation of the principle of sufficient reason. A cause whose effect follows it in time must contain two temporal stages: a period of latency during which it does not yet act and a moment of efficacy in which it produces its effect simultaneously. For, having acknowledged by hypothesis that there are causes that are simultaneous with their effect, the delay of the effect with respect to the cause  $x$  in the case of non-simultaneous causation itself requires a sufficient reason, over and above  $x$ . But now we face a dilemma: if such a supplementary fact exists, it is not true that  $x$  is “the cause”; if there is none, by virtue of the principle of sufficient reason there can be no effect at all. The cause, says Hume, has no sufficient reason to act precisely at that moment, therefore it is impossible that it acts.<sup>52</sup> “For if one cause were co-temporary with its effect, and this effect with *its* effect, and so on, ’tis plain there wou’d be no such thing as succession, and all objects must be co-existent.” (Hume 1955, p. 76; emphasis in original). Indeed, “since any one of them, which retards its operation for a single moment, exerts not itself at that very individual time, in which it might have operated; and therefore is no proper cause” (Hume 1955, p. 76).

Let us consider the form of the principle of sufficient reason Hume uses in this argument: no event, of whatever type, can happen at time  $t$  without something determining its occurrence at that instant. However, this specific



version of the principle is only justified in the framework of a deterministic conception of natural processes. Contemporary physical theory does not any more support such a conception. Physics acknowledges the existence of causes that have exactly the structure of what Hume declares to be inconceivable. Particles susceptible to radioactive decomposition have (taken individually) a period of latency whose length is only statistically determined, followed by an instant at which they produce, as an effect, new particles resulting from their decomposition. The step from latency to the actualization of the causal potential of such a particle is “spontaneous” in the sense of being independent of any triggering factors external to the particle. Speaking of spontaneity in this context may appear as an abuse of language. But it need not be an abuse if we give the term, in the context of Hume’s argument, the following sense: An event happens spontaneously at a given instant  $t$  to an object  $o$ , if and only if it is not constrained to occur at that moment by any factor external to  $o$ . Furthermore, the triggering factors internal to  $o$  are not deterministic, in the sense that they only determine at each moment the *probability* with which the event may happen. In this sense, the occurrence of the event at  $t$  is “indeterminate”. Nevertheless, it is preferable to avoid using the term “indeterminate”, to the extent that it may give the erroneous impression that the occurrence of such an event is not subject to the laws of nature. Nothing determines which of several possible events occurs at instant  $t$ , as far as an individual radioactive particle is concerned. In this sense it can be said of a given decomposition event that it is spontaneous although it is also subject to a probabilistic law. This is possible because a probabilistic law does not directly determine the evolution of each individual particle to which it applies, but only the average evolution of a large number of such particles. It should be clear that this is “spontaneity” in a minimal sense: contrary to human action where spontaneity takes its full sense, in the context of a radioactive particle, the term is meant to refer to the fact that it is indeterminate which of two possibilities becomes actual. In any case, the behaviour of radioactive particles constitutes a counterexample to the version of the principle of sufficient reason used by Hume: their behaviour is not completely determined by *external* causes.

The existence of irreducibly indeterminate processes prevents us from using Hume’s argument against simultaneous causation. Indeed, the validity of the principle of sufficient reason with respect to singular events on which the argument relies is inseparable from a deterministic conception of the world that has been overthrown by the discovery of irreducibly indeterminate processes. However, if we give up using the principle of sufficient reason, must we conclude as do many contemporary authors, that there are two sorts of causal interactions, one being simultaneous, the other taking a certain amount of time? We have seen that the defenders of the anthropocentric theory, such as Gasking and von Wright, argue against Hume essentially on the basis of apparent examples of simultaneous causation.

The locomotive's starting to move causes the wagons' starting to move, apparently without any temporal delay; the increase of temperature of the piece of iron seems to be the cause of its change of colour; the increase of the temperature of the gas seems to be the cause of the increase of its pressure, in case its volume is held constant.<sup>53</sup> The argument against Hume's conception of causation consists in observing that it grounds the distinction between cause and effect entirely on the temporal order of their succession. Of two causally related events, the earlier is the cause and the latter is the effect. If, however, contrary to what Hume holds, there are simultaneous causal relations, the framework of Hume's theory does not provide any justification to consider one of its terms to be the cause rather than the other. Gasking and von Wright conclude that the concept of causation contains an additional element stemming from the sphere of human action. The conceptually primordial causal relations are those where a human brings about an event. That event is the effect and the action is its cause, independently of whether they succeed each other in time or are simultaneous.

In order to resist this argument in favour of the anthropocentric conception, it is necessary to show either that there are cases of simultaneous causation whose asymmetry is nevertheless independent of human action, or that there really are no cases of simultaneous causation. The transference theory gives us the conceptual means to justify the latter judgment. In point of fact, it implies that strictly speaking, there can be no simultaneous causation. For if every causal relation consists in a transference and if every transference takes time, how could an effect be simultaneous with its cause? The only way to conceive of this possibility would be to require that the cause be not only *simultaneous* with the effect but also occupy the same *position* as the effect. It remains for us to show that simultaneous causation is only *apparent* but not real. Indeed, the theory of transference predicts that in each *apparent* case of simultaneous causation, it either turns out that the cause is not *strictly* simultaneous with the effect after all, or else that it is not a causal *relation*, in the sense that "the cause" and "the effect" constitute really only *one* event.

In each case in which an effect appears to be simultaneous with its cause, we must first of all check whether the cause also occupies the same position as the effect. Then our thesis is the following: if cause and effect are situated at some spatial distance from each other, one can always discover that they are also separated in time, though maybe only by a very short interval. If on the contrary, "the effect" occupies the same position as "the cause", and if they are strictly simultaneous, then the relation between them is not causal. In those cases where the effect is spatially separated from the cause the simultaneity is only apparent; in the other cases the relation is not causal.

The first part of this thesis can be justified by a physical argument. As for the possibility of "action at a distance", in other words situations where an

event has a simultaneous effect happening at some distance (“action” is not meant to refer to human action), we should rely on the theory of special relativity, which constitutes the conceptual framework for the physics of space and time. The theory of relativity contains the postulate that causal influence cannot spread with a speed faster than that of light. All theories of the mechanisms of the action of the different natural forces satisfy this postulate. With respect to gravitation, we have already noted the fact that its action at a distance is only apparently simultaneous.<sup>54</sup> For gravitation as for all other forces acting at a distance, it is possible to explain their efficacy on objects at some distance from the source of the force by the existence of a local *field*. The field takes its origin at a given spatio-temporal location and propagates with finite speed, in empty space with the speed of light. The gravitational field only differs from the fields transmitting the other forces by the experimental difficulty to observe its propagation directly, and in particular its speed. But this is due to contingent facts: no known process creates or destroys sufficient quantities of matter and the gravitational interaction is extremely weak compared with the other forces.<sup>55</sup>

Similarly, it might seem that when a train starts moving, the last wagon departs at the same time as the locomotive, which would mean that the locomotive causes the last wagon’s movement instantaneously at a distance. However, this simultaneity is only apparent. The force responsible for the fact that the energy deployed by the locomotive brings about the movement of the wagons is the force of the chemical bond, which guarantees the cohesion of solid objects. One can conceive of the propagation of this force with the help of a model in which the atoms and molecules constituting the train are connected with springs. When the spring linking the last molecules of the locomotive to the first molecules of the first wagon is stretched, the spring is thereby put in a state of higher energy. The movement allowing the spring to return to equilibrium, in other words to the state corresponding to a minimum of energy, brings the first molecules of the first wagon closer to the locomotive. But this will have the effect of stretching the second spring, which will react in the same way, and so forth to the last molecules of the last wagon. As a consequence and contrary to appearances, the wagons begin to move a little later than the locomotive. Even if the delay is extremely short, it is possible to observe it experimentally.

*Static* causal relations raise a more serious conceptual difficulty. I have defended the thesis that the existence of a causal relation does not require any change (see the second section in this chapter). Now in the absence of a change in the effect brought about by the cause, it is difficult to resist the intuition that the cause is simultaneous with the effect. Zangari (1992) considers the parallel evolution of two objects mutually dependent on each other without any apparent asymmetry, such as the movements of the moon and the earth, each due to the causal influence of the other body’s presence. There is no reason to claim that the earth exerts a

causal influence on the movement of the moon, which would not also be a reason to claim that the moon exerts a causal influence on the movement of the earth. From the fact that the physical description of the evolution of these two celestial bodies does not show any asymmetry in their mutual causal influence, Zangari concludes that “the ontological characterisation of the causal connection” must rely on “the interaction term itself, which possesses the desired symmetric properties” (Zangari 1992, p. 263). But it is not sufficient to show that there *exists* a symmetric description that can serve as a basis for calculating the effects of a causal interaction, to conclude that the interaction itself is fundamentally symmetric. In the case of Newtonian and electrostatic interactions, Zangari himself notes that such a reasoning is mistaken: the classical description of these forces is symmetrical and instantaneous. But in the relativistic representation – which is ontologically primordial to the extent that the classical description constitutes only an approximation to it – an asymmetry between source and effect appears, and the action of one on the other takes the form of a propagation that unfolds in time.

However, it is not inevitable to conceive of causal relations between static states and between parallel evolutions as cases of symmetrical causation. The physics of interactions is compatible with an alternative description according to which these apparently symmetrical relations are after all grounded in asymmetrical processes. Every force even if it does not bring about any movement or apparent change, acts through a sequence of interactions. The action of forces can be conceived as resulting from an exchange of “boson” particles, of a kind specific for each type of force. This allows us to draw the following conclusion: underlying every stable causal relation, there are relations of a “transmission” type between the temporal parts of the objects taking part in the causal relation,<sup>56</sup> and the events that are the terms of these transmission relations are not strictly simultaneous.

A conceptual argument is needed for the second part of our thesis that there is no causal relation if the “effect” is both simultaneous and spatially coincident with the “cause”. I shall provide such an argument by justifying, first, that causes and effects are fundamentally events rather than objects and second that events are particulars.<sup>57</sup>

## **Objects and events**

For Hume, causes and effects are “objects”. However, we must not give this term its modern interpretation, if only because Hume takes as illusory the idea that objects persist through time. In our terminology, an object is a particular whose extension in space but not in time is part of the conditions determining its identity. An object may exist during a more or less extended time interval; however, this duration is not constitutive of the identity of the object. To identify an object, it is necessary and sufficient to indicate its

*spatial* limits. The fundamental principle of the identity of objects is that two objects cannot occupy the same portion of space at the same time. However, the identity of a chair or of a photon at a given moment does not depend on whether it has existed before or whether it will exist after that moment.

The concept of *event* differs from the concept of object in this respect. The temporal dimension plays exactly the same role in determining the identity of an event as the spatial dimensions. A causal theory of persistence of objects through time becomes conceivable only by considering causes as events, i.e. entities that are essentially bounded in the temporal dimension. In their respective versions of the transference theory, Aronson (1971a) and Fair (1979) both follow Hume and consider that causes and effects are “objects”. This choice is a source of difficulties that can easily be avoided by taking events rather than objects as terms of the causal relation.

In his critical comment on Davidson’s (1967) work on the logic of action sentences, Lemmon (1967, p. 99) has proposed to identify events with space–time zones:

- (1) An event is identical to a space–time zone.

Lemmon’s proposal is not intended as an epistemic criterion allowing us to identify events, but rather as a metaphysical hypothesis about their nature. Most of the arguments that have later been put forward against the idea that the terms of the causal relation are events extensionally conceived took Lemmon’s particular formulation as their target. Lemmon suggests that events should be *identified* with the space–time zone they occupy, and this raises difficulties as soon as one tries to apply principle (1) to particular cases.

Of course, everyone is free to redefine their conceptual tools. However, unorthodox conceptual choices must bear the burden of proof if they lead to counterintuitive results. To convince us that events *are* zones instead of occupying them such an interpretation of the concept of event must be useful, in particular in allowing us to draw relevant distinctions. If one wishes to identify events with space–time zones, one cannot say of the opening of a door that it *is* an event but must rather say that it takes place *in* an event. The examples offered by Lemmon himself illustrate well the counterintuitive character of his conception, as well as the artificial difficulties it creates. Let us say that *z* is the zone identifying an event. If we suppose that a person *p* (or an object *o*) is in this zone, it participates thereby in the event that is identical with *z*. However, insofar as the zone may have considerable temporal extension and last longer than an instant, Lemmon concludes that it is possible that “a participant may weave into and out of an event” (Lemmon 1967, p. 99). Lemmon’s conception has the merit of clarifying a weakness that it shares with Davidson’s. Lemmon reaches his extensional concept of event with the aim of resolving a difficulty

concerning the *description* of events: “We suppose that numerically the same event (action) can be described in different ways. So it is natural to ask for the criteria for two events being the same – to use the usual misleading idiom” (Lemmon 1967, p. 98). I suggest that Lemmon ends up with an inadequate concept of event because he ignores a crucial distinction between two questions. One concerns reference and possible ways in which two different symbolic expressions can have the same denotation. The other question, with which we are concerned here is to which ontological category belong the terms of causal relations. Lemmon’s error consists in putting an epistemic criterion to an ontological use for which it is not adequate.

Quine has proposed a different extensional conception of events: “Each [event] comprises simply the content, however heterogeneous, of some portion of space–time, however disconnected and gerrymandered” (Quine 1960, p. 171). In this form, the space–time zone clearly plays the role of a criterion, as distinct from the event whose identity it is supposed to determine. Quine completes his criterion of spatio-temporal position with the thesis of the arbitrary character of the determination of the limits of the zone occupied by an event; as Quine himself shows in a later work,<sup>58</sup> in this respect there is no difference between the categories of events and objects. The spatial limits of objects are also determined only to a degree that depends on pragmatic circumstances. On an atomic scale, the limits of ordinary objects are spread out over a thickness whose order of magnitude corresponds to several atoms. For Quine, the difference between objects and events is only that the delimitation of objects of medium size is subject to stronger constraints than the delimitation of the space–time zone occupied by events. However, the reverse seems to hold for elementary particles: the external limits of a particle are extremely ill defined as long as it does not interact, in other words insofar as it does not participate in an event. While a photon is not interacting, it behaves as a wave that is infinitely extended in space. However, when it gets absorbed, for example by a solid surface, the event of interaction takes place at a well-defined place. In any case, the leeway one has in defining macroscopic events by way of the space–time zone they occupy, does not constitute an objection against the fundamental character of the ontological category of events, on an equal footing with the category of objects.

In a later definition Quine takes objects to be ontologically more fundamental, but to determine their identity uses the same spatio-temporal criterion as for events. “The problem of individuation of events would seem to be dissolved now by the assimilation of events to physical objects or to some sort of constructs upon physical objects. For physical objects are well individuated, being identical if and only if spatiotemporally coextensive” (Quine 1985, p. 167). Quine assimilates the categories of events and objects, reducing the difference between them to a difference of degree: whether an entity is an object or an event depends on the temporal extension of the

zone it occupies. If its zone is temporally extended it is an object, otherwise it is an event.

However, Quine's position which is motivated by the aim of ontological parsimony – not to multiply fundamental ontological categories unnecessarily – has two unwelcome consequences. It excludes the possibility of objects existing for only a very short time. As a consequence, those entities studied in particle physics that result from decomposition processes and are themselves subject to very quick decomposition are, according to Quine's conception, not really objects but rather events. Conversely, Quine's conception excludes the possibility of events lasting for a long time: beyond a certain threshold, they are really objects rather than events.

The element of Quine's conception we shall retain is the idea that the space–time zone where an event is located must not be confused with the event itself: the spatio-temporal delimitation plays the role of a criterion of the identity of events. Contrary to Quine, we shall consider every content of a well-delimited space–time zone as an event, independently of its extension in the temporal dimension. The difference between the categories of objects and events lies in the fact that the identity of an object depends only on its spatial but not on its temporal limits whereas the identity of an event depends on both.

If we were to make the hypothesis that causes and effects were objects, and admitted that a cause cannot be identical with its effect, it would be impossible to give a causal explanation of the evolution of an object that is not in interaction with other objects. Not only would a causal theory of the persistence of objects through time be excluded by definition, in other words by the choice of the conceptual framework, but it would also become impossible to consider as a causal process the successive transformations of a chemical substance undergoing a sequence of reactions. One is of course free to name, as Aronson (1971a) does, “natural changes” those changes that take place in an object without being due to interactions with other objects, and “unnatural” (Aronson 1971a, p. 421) those changes that result from interactions. However, there are two reasons to reject the idea that only “unnatural” changes are causal. First, as we have already noted, there are static causal relations in which either the cause or the effect or both undergo no change. Second, as Ehring (1986, p. 250) notes, we have no reason to exclude the possibility of a causal process that is entirely internal to a substance, such as a chemical transformation.

As soon as the terms of the causal relation are taken to be events, these problems do not arise any more. Amounts of energy or of other conserved quantities can be exchanged while the macroscopic properties of the objects that are constituted by these events do not change. When a substance undergoes a transformation by virtue of internal constraints but without being subject to external interactions, the different stages of the transformation are so many events linked by causal relations grounded on the transference of conserved quantities.

### **The spatio-temporal separation of cause and effect and the persistence of objects through time**

Let us now take up once again the argument that we have left pending, that there can be no causal relations in which cause and effect are simultaneous. The crucial thesis in this respect is that, at the fundamental ontological level at which the transference theory operates, causes and effects are events conceived of as contents of space–time zones. The events whose causal interactions we consider are very often temporal parts of objects or of fields. Our argument against the possibility of simultaneous causation depends on the defence of this thesis against the rival idea that causes and effects are *facts*. A fact is what makes a proposition true; it consists in an object’s satisfying a predicate. By contrast, according to the definition of an event as the content of a space–time zone, an event is a particular: it has an indefinite number of intrinsic properties, i.e. properties independent from the external relations between the event and other particulars.

Justifying the thesis that events and not facts are the fundamental terms of the causal relation, also argues against the existence of causal relations where cause and effect coincide both spatially and temporally. Let us take the relation between the increase of the temperature of a gas and the increase of its pressure. According to our conception these are two aspects of a unique event, which is a change in the state of the gas: the increases of temperature and pressure take place at the same time in the space occupied by the gas. However, they correspond to two different facts; and this explains why it is at first sight conceivable (although it turns out to be wrong) to consider them as related as cause and effect. When it is said that a change in temperature causes a change in pressure, does one thereby refer to a causal relation between facts? In other words, does the word “because” in the following statement indicate that it expresses a causal explanation?

(2)  $P(g,t)$  because  $T(g,t)$ ,

where  $t$  represents an instant of time,  $g$  a sample of a gas that can be considered “ideal”,  $P(g,t)$  the pressure of  $g$  at  $t$ , and  $T(g,t)$  the temperature of  $g$  at  $t$ .

The truth conditions of statement (2) seem to be perfectly symmetrical: (2) is true if and only if its converse

(3)  $T(g,t)$  because  $P(g,t)$

is true. This is because the truth of (2) and (3) is grounded on the same law, the so-called ideal gas law:

L(IG)  $PV = nRT$ ,



where  $P$  represents pressure,  $V$  volume and  $T$  temperature;  $R$  is a constant factor and  $n$  represents the quantity of gas molecules contained in the sample, measured in moles.

This law states the proportionality of the numerical values of the quantities  $P$  and  $V$  taken by the sample  $g$  at time  $t$ . The law being symmetrical with respect to  $P$  and  $V$ , L(IG) grounds the explanation of  $T(g,t)$  by  $P(g,t)$  exactly to the same extent as it grounds the explanation of  $P(g,t)$  by  $T(g,t)$ . With causality matters are very different. Let us first admit that causation is an irreflexive relation: no event or fact can be its own cause. However, the causal relation is transitive in the sense that, if  $A$  causes  $B$  and  $B$  causes  $C$ , then  $A$  is also cause of  $C$ . Therefore, by transitivity, the conjunction of (2) and (3) –  $T(g,t)$  because  $P(g,t)$  and  $P(g,t)$  because  $T(g,t)$  – implies:  $T(g,t)$  because  $T(g,t)$ . Therefore, if causation is an irreflexive relation, the expression “because” in (2) cannot express a causal relation; we may rather interpret it as representing a relation of (non-causal) explanation directly grounded on the law L(IG). This explanation is non-causal in the sense that the explanans does not designate a fact causally responsible for the effect designated by the explanandum.

The hypothesis of the irreflexivity of the causal relation is of course justified only insofar as *particular* causes are concerned. As soon as a cause is considered as a token of a general type, it becomes possible to say “*an A* causes *an A*”. This is possible insofar as the two occurrences of “*A*” designate different events of the same type. The same term can also be used to make reference to the whole class rather than to an individual belonging to the class. However, this does not contradict the fact that, insofar as particular occurrences are concerned, no particular event and no fact bearing on a particular entity can be its own cause. The statement that “*an A* causes *an A*” can be interpreted as meaning that a particular occurrence of type  $A$  causes another occurrence of the same type.<sup>59</sup>

We should admit that the statements “ $T(g,t)$  because  $P(g,t)$ ” and “ $P(g,t)$  because  $T(g,t)$ ”, even if they do not directly *express* causal relations, can nevertheless *suggest* the existence of causal relations. However, far from throwing doubt on our analysis, this fact reinforces it because those statements suggest *different* causal relations. One might express this by saying that the statements suggest the existence of causal relations between the facts  $T(g,t)$  and  $P(g,t)$ , by virtue of the rules of pragmatic relevance.<sup>60</sup> These causal suggestions are pragmatic implicatures that are not part of the semantic content of the statements. This becomes particularly clear in circumstances in which the change in gas  $g$  is brought about by a human agent. In this case, the first statement (“ $T(g,t)$  because  $P(g,t)$ ”) suggests that the change in the two properties of the gas has been brought about by an intervention that brings energy to the gas in the form of pressure, e.g. by pushing on the walls of the container. The second statement (“ $P(g,t)$  because  $T(g,t)$ ”) suggests that the gas has received energy by being heated.

With respect to the change of the state of the gas, this difference is insignificant because within the gas all energy, in whichever form it has been transmitted, takes the form of kinetic energy.

It is important to insist on the fact that this difference bearing on causation is not part of the semantic content of statements (2) and (3) but is rather suggested pragmatically. As far as the semantic content of the statements is concerned, our argument is the following. On the basis of two premises:

- (i)  $T(g,t)$  and  $P(g,t)$  are singular facts not types of facts;
  - (ii) singular facts and particular events cannot be their own cause,
- we have reached the conclusion that the statements (2) and (3) have no causal content.

We have reasoned about the change of state of a sample of gas, without making use of the particular properties of this case. We can therefore generalize our conclusions.<sup>61</sup> First, the relation between two facts constituted by a particular's possessing two properties is never causal. Cause and effect cannot coincide both in space and time. In other words, there can be no simultaneous causation where the cause coincides also spatially with the effect. Together with the argument that there is no simultaneous causation *at a distance*, this establishes the more general thesis that *there is no simultaneous causation*. Second, if our argument is correct, it gives us reason to think that facts are not the fundamental bearers of causation. True, many factual statements express causal relations. Later in the book, we shall examine them in some detail. It will result from this enquiry that the truth of a factual causal statement presupposes the existence of a causal relation between events.

Here is another argument for the thesis that events are the fundamental terms of the causal relation. Causation seems to be the only relation capable of explaining the persistence of concrete objects<sup>62</sup> through time. Carnap (1928), after Russell (1914)<sup>63</sup> and Lewin (1923), has introduced the concept of *genidentity* to designate the relation between different moments of the existence of an object.<sup>64</sup> According to Carnap, genidentity is the property of two different world points to belong to the same world line.<sup>65</sup> He explains that genidentity, i.e. the "association of various 'thing-states' with one object" (Carnap 1928, p. 218; trans. p. 252), belongs to a class of relations of "improper identity" (Carnap 1928, p. 218; trans. p. 253) that natural language puts into the same category as identity but that differ from logical identity: two temporal parts of the same object at different moments are not identical in the logical sense, because one satisfies predicates that the other does not. To say of these parts that they are nevertheless genidentical is to attribute to them an "equivalence relation" (Carnap 1928, p. 219; trans. p. 253) weaker than logical identity.

Transference theory gives a causal sense to the concept of genidentity.<sup>66</sup> It allows us to conceive of genidentity as grounded on the transmission of

amounts of conserved quantities between each temporal part of an object to the adjacent temporal parts. Here it becomes clear that the theory presupposes a kind of fundamental diachronic identity pertaining to particular amounts of conserved quantities. We would get caught in a vicious circle if we tried to explain the identity of such a particular quantity  $Q$  through time in turn as a causal process. However, this is not what I propose to do.  $Q$  remains identical with itself by virtue of the conservation law that governs the quantity of which it is an amount. The validity of the conservation law itself must be considered as primitive. On this basis we can explain the persistence of an object carrying such quantities. We can conceive of the object as an entity that is structured with respect to its existence through time, as a continuous series of events. Analogous to spatial decomposition, we can speak of these events as of “temporal parts” of the object. The object persists because each of its temporal parts transmits amounts of the conserved quantities it bears – mass-energy being the most fundamental among them as far as material objects are concerned – to the temporal parts next to it. In this manner, genidentity, the identity of an object through time, appears as derivative with respect to the more fundamental diachronic identity of the ultimate constituents of the object, i.e. amounts of conserved quantities.

The idea of explaining the persistence of objects through time in causal terms implies abandoning the more traditional concept of substance. Russell has clearly noted this fact. In his words, a “causal line” is a causal chain of events underlying the persistence of an object through time. “A causal line may always be regarded as the persistence of something – a person, a table, a photon, or whatnot” (Russell 1948, p. 477). This explanation of persistence replaces the traditional explanation in terms of substance.

A “thing” or a piece of matter is not to be regarded as a single persistent substantial entity, but as a string of events having a certain kind of causal connection with each other. This kind is what I call “quasi-permanence”. [ ... ] When “substance” is abandoned, the identity, for commonsense, of a thing or a person at different times must be explained as consisting in what may be called a “causal line.”

(Russell 1948, p. 476)

By making a distinction between two kinds of diachronic identity – one applying to particular amounts of conserved quantities and the other applying to concrete objects bearing these quantities – our theory can explain the fact that an object can persist even though all of its constituent parts are progressively substituted for other parts, something that happens regularly to the human body. We can explain the genidentity of such an object notwithstanding its “regeneration” by the strict identity of the vast majority of amounts of conserved quantities it contains from one moment to the next. The material parts of the body persist through time because

mass obeys a conservation law. However, if all material parts of a body were substituted simultaneously, there would not be any relation of transmission relating the state of the body before and after the event of substitution. In this case, our theory correctly predicts that there is no object persisting through the substitution. Rather, the initial object ceases to exist and is replaced by a different object that may qualitatively resemble it.

Another important consequence of this distinction is the fact that it gives the transference theory the means to explain why material objects can come into existence not only by composition and cease to exist not only by decomposition of their material parts. For example, an “atom” of positronium – an electron ( $e^-$ ) and a positron ( $e^+$ ) in a bounded state – is a massive object that disintegrates into non-massive objects: two  $\gamma$  quanta. Justifying this event as a transformation rather than as a case of disappearance of matter is grounded on the possibility of identifying the amounts of mass-energy and of momentum transferred from the electron and positron to the  $\gamma$  quanta, which obey conservation laws.

The crucial idea is that the persistence of an object is grounded on the existence of causal relations between its successive “temporal parts”. These temporal parts are events: they are *particular* entities in the same sense as the object of which they are parts. The concept of an object and of an event are related in this way: most events contain one or several objects as “participants” but the objects themselves consist in series of events.

On the condition of considering the terms of causal relations as events not objects, we can answer the following objection raised by Sosa and Tooley (1993): “Although the temporal parts of an electron, for example, are causally related, this relation does not involve any transference of energy or momentum *from one object to another*. So causation cannot be identified with physical relations involving the transference of energy and/or momentum” (Sosa and Tooley 1993, p. 4, my emphasis). According to our analysis, the terms of the causal relation that guarantees the persistence of an object through time are the “temporal parts” of the object. These temporal parts are events; they are particulars which possess, over and above the spatial limits that they inherit from “their” objects, temporal limits. During those periods of existence of the object in which it does not interact, the relevant causal relation is constituted by the transference of all amounts of conserved quantities the object possesses, from one moment to the next.

This explanation of the genidentity of an object is not vulnerable to the following objection raised by Ehring. He points out that, “in the case of inertial motion [ . . . ] no causal explanation is required” (Ehring 1986, p. 254),<sup>67</sup> and concludes from this that transference, understood as “a change in position of a quantity of energy/quantity of momentum” does not always “involve causation” (Ehring 1986, p. 254). But first, the fact that one judges that a process or state does not require any causal *explanation* does not prejudice the question of whether it *objectively* has a causal character. Second, it suffices to direct one’s attention on what seemed obvious at first

sight, to be able to consider it as a new explanandum: in this sense, we explain the persistence of objects through time, their genidentity, as a specific case of transference of energy and mass from one space–time location to another. What better refutation could there be of the thesis that something doesn't require any explanation, than to explain it?

### The transmission of amounts of conserved quantities

Let us imagine two similar situations each containing two events that are contiguous in space and follow each other immediately in time. The events in these situations are in the same position one with respect to the other, both in space and time. Imagine that in one situation these events are causally related but not in the other. The question which a theory of causation tries to answer is this: what distinguishes the relations between the two events in these two situations? Transference theory gives a clear answer: what makes the relation between the events causal in one situation is the fact that an amount of a conserved quantity such as energy or momentum is transferred between these events, whereas the absence of such a transference is the reason that the relation is not causal in the other.

However, this answer is satisfying only if an important difficulty can be overcome. How can we justify the judgment that there really is a transference rather than the mere *appearance* of transference? To answer this question we must first answer another, more fundamental question. Are amounts of conserved quantities a sort of entities capable of being transferred? As Aronson (1982) has clearly observed, we can justify the thesis that the causal relation is *objectively* more than mere “contiguous succession” only if it is possible to justify the idea that *the same* amount can appear in two distinct events. “This criterion [ . . . ] makes sense of how it is that the causal relation is stronger than the mere correlation of events. In the case of *a* causing *b* to move, *b*'s momentum at time  $t_3$  is the same as *a*'s at time  $t_1$ . Denying this would lead to quite a mystery, for then we would have to say that somehow *a* lost all of its momentum at time  $t_2$  and *b* acquired the exact same amount at that very moment!” (Aronson 1982, pp. 294–295).

Several authors<sup>68</sup> have expressed doubts about whether amounts of conserved quantities are capable of being transferred. Such doubts have been motivated by both ontological and epistemic reasons: Ontologically, one may doubt the objective determination of individual amounts of conserved quantities through time; epistemically, one may doubt the possibility of recognizing the identity of such a particular amount.

Here is the ontological argument. In order to speak of the transference of an amount  $Q$  of a conserved quantity from event  $c$  to event  $e$ ,  $Q$  must have the character of a *substance*: it must remain, in a sufficiently strong sense, *the same* for the duration of the transference. The crucial step in this argument against the use of the notion of transference in the theory of causation consists in undermining the presupposition that the attribution of

diachronic identity to an individual amount of conserved quantity is well founded.

A second argument aims not at establishing that the amount  $Q$  lacks the necessary brand of diachronic identity but at establishing that it necessarily lacks a *mark* that might allow one to *recognize the same  $Q$*  in different events. This argument is epistemological rather than ontological: it says that our reduction (S) of the causal relation to a kind of transference is not adequate because it provides no means to distinguish in principle between causal processes and “pseudo-processes” in Salmon’s (1978, 1984) sense.<sup>69</sup>

Here is how Dieks (1986) spells out the first argument. He begins by observing that both classical physics and quantum physics provide reasons for denying that particular amounts of conserved quantities, such as energy or momentum, have the character of substances. It is correct that classical physics does not justify considering either energy or momentum as “a kind of substance which is transferred, while retaining its identity, from the particles and fields before the interaction to those after the interaction” (Dieks 1986, p. 88). The reason is that the only restriction imposed on interactions in classical physics is the *global* conservation of these quantities. However, this is not so in the modern theory of interactions, i.e. quantum mechanics, where the quantities in question are considered as locally conserved.<sup>70</sup>

But quantum theory provides a stronger reason to doubt the hypothesis that individual amounts of energy, momentum and other conserved quantities can, like substances, remain identical through an interaction. If it turns out to be impossible to justify the hypothesis that they remain identical through time as particular amounts, the idea that a given amount of, for example, energy could be “transferred” has been undermined. In order to be transferred, something must *persist* through time. However, according to Dieks, the theory of indistinguishable particles (also called, apparently paradoxically, “identical” particles) gives us a reason to deny that they can persist as individuals. Indeed, in a system containing several (boson) particles of the same type in interaction, it is not only impossible to *know* the evolution of individual particles, but such particles *are objectively not* distinct from each other. To suppose that they are leads to false empirical predictions. To the extent that these particles are capable of interacting, they are not only epistemically indiscernible, but really “identical”, in this sense: all particles of the same type that interact share their location. To express the same thing in a different way: when two particles that do not differ by any intrinsic property are in interaction, they do not possess any individual identity.

It is decisive for the argument against transference theory that the lack of individual identity that characterizes interacting particles of the same physical type immediately implies that *each instance* of a given property in one particle is identical to the instances of that property in all other particles. The same reasoning applies to a particle and its properties, to the extent that the question of the identity of a particle reduces to the question of the identity of the properties the particle possesses, for example the property to

be an electron or the property to possess kinetic energy  $E$ . This is so at least if one supposes that an object *consists in* a set of property instances without any substance existing beyond these properties.

During an interaction, energy, momentum and other conserved quantities can only be attributed to the whole set of identical particles, but not to any individual particles. This means that we must accept Dieks' (1986) conclusion: it makes no sense to attribute, during an interaction among identical particles, a particular amount  $Q$  of a conserved quantity to one of these particles rather than another.

Consider the following schematic interaction. Imagine that two particles  $a_1$  and  $a_2$  that are glued together lose all their momentum in an elastic collision with two other particles  $b_1$  and  $b_2$ , also glued together. In this interaction each of the  $b$  particles receives half of the total momentum given away by  $a_1$  and  $a_2$ . However, it is not only true that one cannot *know* whether  $a_1$  has transferred its momentum to  $b_1$  or to  $b_2$ , but this question has no answer (makes no sense) because the momentum lost by  $a_1$  (in opposition to the momentum lost by  $a_2$ ) lacks the identity necessary to distinguish whether *it* or the momentum of  $a_2$  has been transferred to one of the  $b$  particles. Each pair,  $a_1$ – $a_2$  and  $b_1$ – $b_2$ , forms by virtue of the interaction we have called “the glue” binding it together, a whole such that it is impossible to attribute the momentum to one rather than the other part of the whole. During the collision, this is true of the whole set of the four interacting particles: insofar as they are by hypothesis of the same type, they share their identity during the interaction, and share therefore also their momentum.

In our example of the collision of  $a_1$ – $a_2$  with  $b_1$ – $b_2$ , it makes no sense to suppose that  $a_1$  transfers *its* half of the momentum possessed by  $a_1$ – $a_2$  before the shock to  $b_1$  rather than to  $b_2$ . However, Dieks is wrong in going further: he holds that the lack of individual identity that we have just acknowledged concerns also those periods during which a particle does not undergo any interaction with particles of the same type. “If there is no interaction between the electrons, quantum theory tells us that there is no *genidentity* of the various particles.” (Dieks 1986, p. 89). The passage just quoted is really not perfectly clear because the genidentity of particles that undergo no interaction is a relation between different temporal parts of the *same* particle, and not, as Dieks suggests, between different particles.

The lack of identity of indistinguishable particles depends on the supposition that they are interacting. We may therefore conclude that a particle, to the extent that it does not interact with particles indistinguishable from it, can conserve its individual identity on the basis of the conservation of the conserved quantities it possesses, such as an amount  $Q$  of momentum. The identity of particles of the same type that interact does not warrant the conclusion that a particle lacks genidentity *even when it does not interact with other particles of the same type*.<sup>71</sup>

The results of quantum physics in this domain do not provide any reason to doubt the following statement: during the evolution of the pair  $a_1-a_2$  before the interaction, this pair of particles remains genidentical with itself because the momentum  $a_1-a_2$  contains remains the same; during the collision, this momentum is transferred to the pair  $b_1-b_2$  and then conserved during the evolution of  $b_1-b_2$ . In this sense, the pair  $b_1-b_2$  has genidentity, in other words persists through time although  $b_1$  and  $b_2$  do not possess any individual identity in virtue of which one could be considered as different from the other. By introducing a qualification relative to interaction, we can make the notion of persistence through time consistent with the impossibility of attributing an independent identity to particles in interaction. On the one hand, the particles lose their individual identity to the extent to which they interact. But on the other hand, if two particles of the same type are isolated one from another, so that they do not interact, one can attribute them an individual identity.

Dieks offers another argument, independent of the theory of indistinguishable particles, to reject the genidentity of elementary particles. According to one interpretation of quantum theory, each interaction leads to the “annihilation” of all the participants in the interaction, followed by the “creation” of “new” particles resulting from the interaction. With the exception of the properties affected by the interaction, the particles that are created may be perfectly similar to those which have been annihilated. For example, in the interaction known as the Compton effect, a photon is scattered by an electron: the electron absorbs the photon so that the photon is annihilated and a new photon is created at the moment at which it is emitted by the electron. In this sense, the physical conception of the nature of such an interaction makes the use of the concept of substance impossible, in order to describe the persistence of elementary particles. According to Dieks, this proves more: it proves that elementary particles and the amounts of conserved quantities they bear have no diachronic identity at all. However, nothing prevents us from thinking that they persist through time in a weaker sense than a substance. The theory imposes sufficiently strong constraints on the annihilation and creation of particles to justify the idea that, except in situations in which several “identical” particles are present, a given particle possesses genidentity, and the amounts of conserved quantities it bears possess a primitive kind of diachronic identity that we can compare, with Dieks (1986), to the identity of a flux. In the case of Compton scattering, the continuity of the persistence of the energy and momentum borne by the photon is guaranteed by the genidentity of the electron bearing these amounts between the annihilation and recreation of the photon.

If one distinguishes as we do, between the primitive identity through time of individual amounts of conserved quantities and the genidentity that characterizes the persistence of concrete objects, one can reject the following objection. Dowe (1995b) argues that the question how a *property* could be



genidentical leads to a dilemma. Either the property is conceived as a universal, in which case it can be possessed by different objects or by the same object at different moments, but it makes no sense to say that the property the object possesses at a later time is genidentical to the property it possesses at an earlier time. Indeed, the concept of genidentity applies exclusively to concrete objects, not to properties conceived as universals, nor to particular properties such as an individual amount of energy. Or one considers that the predicate “*a* possesses *b*” in the sentence “object *a* possesses the individual amount of energy *b*” expresses a relation, which, says Dowe, “suggests that energy is not a property at all, but a thing, a substance, separate from the object. [...] But energy cannot stand alone, possessed by nothing.” (Dowe 1995b, p. 369). The theory put forward escapes this dilemma by conceiving individual amounts of conserved quantities as something different still. An individual amount of a conserved quantity is an abstract particular, sometimes called a “trope”. Tropes of conserved quantities have a primitive diachronic identity that differs both from the genidentity of concrete objects and from the identity of universals. We might consider the relation between an object and the tropes it contains as a third meaning of the predicate “*a* possesses *b*”. However, it is less misleading to express this relation by saying: “the concrete object *a* contains the trope *b*”.

### Causal processes and pseudo-processes

Several authors<sup>72</sup> have formulated the following epistemological argument against the idea that amounts of conserved quantities can be transferred. It says that it is impossible to *identify* an individual amount  $Q$  of conserved quantity as being the same at two moments of time; in other words, if an object possesses quantity  $Q$  at time  $t$ , and if another object possesses at time  $t'$  an amount  $Q'$  of the same quantity, it is in principle impossible to *find out* whether  $Q$  is identical to  $Q'$ . In the case of kinetic energy, e.g., says Dowe, “it is not possible to distinguish energy transmission from a regular appearance of energy” (Dowe 1992a, p. 214). According to Dowe, a criterion such as our (S) does not allow us to decide whether in a given situation a certain amount  $Q$  has been transferred from event  $c$  to event  $e$ , or whether there is just a series of appearances of energy of equal amounts in  $c$  and in  $e$  but where these amounts have independent origins. The criterion does not allow us to exclude the possibility that the relation between  $c$  and  $e$  is only apparently causal. This is the case in situations where an amount  $Q$  is present in the event  $c$  and where a different amount  $Q'$  appears in the event  $e$  whose spatial and temporal proximity with  $c$  gives rise to the misleading impression that  $Q$  is identical to  $Q'$ .

This argument contains the following confusion. One can distinguish between “energy transmission” and “regular appearance of energy” only insofar as one acknowledges that individual amounts of energy can persist through time. By denying this possibility, Dowe robs himself of the conceptual

foundation of the distinction between transmission and regular appearance. Dowe objects to transference theory that it is unable to justify the distinction between causal processes and pseudo-processes. However, we shall see that it is on the contrary precisely the notion of transference that provides the foundation for this distinction, where the application of the concept of transference supposes a minimal identity through time for tropes of conserved quantities.

According to Dowe's objection, transference theory is unable to provide a criterion allowing us to distinguish causal processes from pseudo-processes, in the sense Salmon (1978, 1984) has given this term. An attentive examination of the distinction between causal processes and pseudo-processes will allow us to establish two important results. First, criterion (S) allows us to justify and explain the conceptual distinction between causal processes and pseudo-processes whereas the rejection (by, among others, Salmon himself) of the idea that amount of conserved quantities can be transferred destroys the foundation of this distinction. Second, the definition of a causal process proposed by Dowe fails because it leads to counting certain pseudo-processes among the causal processes.

Let us begin by recapitulating Salmon's distinction between causal processes and pseudo-processes. Salmon grounds this distinction on Reichenbach's criterion of "mark transmission",<sup>73</sup> which he formulates in the following way:

*Let  $P$  be a process that, in the absence of interactions with other processes, would remain uniform with respect to a characteristic  $Q$ , which it would manifest consistently over an interval that includes both of the space-time points  $A$  and  $B$  ( $A \neq B$ ). Then, a mark (consisting of a modification of  $Q$  into  $Q'$ ), which has been introduced into process  $P$  by means of a single local interaction at point  $A$ , is transmitted to point  $B$  if  $P$  manifests the modification  $Q'$  at  $B$  and at all stages of the process between  $A$  and  $B$  without additional interventions.*

(Salmon 1984, p. 148; italics in original)

With the help of two other definitions, Salmon uses the criterion of mark transmission to explicate the concept of a causal relation as a process capable of transmitting a mark. These are his definitions:

*If a process is capable of transmitting changes in structure due to marking interactions, then that process can be said to transmit its own structure.*

(Salmon 1984, p. 154; italics in original)

*A process that transmits its own structure is capable of propagating a causal influence from one space-time locale to another.*

(Salmon 1984, p. 155; italics in original)

Salmon clearly expresses his conviction to have provided an analysis of the concept of a causal link:

The propagation of causal influence by means of causal processes *constitutes*, I believe, the mysterious connection between cause and effect which Hume sought.

(Salmon 1984, p. 155; italics in original)

In this theory, a process is defined by its capacity to transmit marks. However, as soon as one adopts an ontology according to which events are more fundamental than processes, one has the means to explain this capacity instead of considering it, as Salmon (1984, p. 139ff.) does, as a primitive fact. An ontology in which the concept of event is fundamental allows us to characterize a process in the following way. A process is a series of events that satisfies a purely spatio-temporal condition: the series of events must be continuous, in the sense that there is no spatio-temporal gap between neighbouring elements of the series. Given that the events themselves are defined by their spatio-temporal extension, this criterion is univocal. If the series of events is represented in a coordinate system where one of the coordinates represents time, it must contain a continuous line that runs across the series of events, without any interruption. In the theory of relativity, such a series is called a “world line”. Its definition shows that the concept of world line is more general than the concepts of process and pseudo-process.

A world line is a causal process if and only if all events constituting it are causally related to each other.<sup>74</sup> A process is characterized by the fact that the events constituting its “temporal parts” are linked by causal relations. Hence a process is causal by definition. The concept of *pseudo-process* can then be analysed with the help of two conditions. The first is negative and opposes pseudo-processes to causal processes: the events that make up the world line of a pseudo-process must not be causally linked among them. Second, a pseudo-process must satisfy a further condition: it must have the phenomenological appearance of a causal process; in other words, it must give the wrong impression of being a causal process. Salmon introduces the notion of pseudo-process by a paradigmatic example. The crucial property of the world line of this example is that it evolves “in a highly regular fashion” (Salmon 1984, p. 142). Salmon’s explanations clearly show – although he does not say so explicitly – that he takes this phenomenological character to be a necessary condition for a pseudo-process. Let us consider his example more closely. Imagine a beacon of light in uniform rotation placed in the centre of a hollow cylinder. The beacon produces a light ray directed towards the lateral wall of the cylinder that produces a light spot. We can now distinguish four series of events, each of which constitutes a world line, and which allow us to illustrate the classification we need.

First, the flow of photons produced at the source of light and flowing towards the wall constitutes a causal process. As far as processes are concerned, Salmon's theory, just as our own, aims to construct a purely objective concept: in this sense, what is decisive for whether something is a process or not, is not whether the flow of photons possesses or not a certain phenomenal or observable character that is manifested along its world line. We shall see that pseudo-processes share precisely this property with certain processes: the manifestation of a salient phenomenological trait. According to our own account, the criterion that justifies our judgment that this is a process, is the transference of amounts of conserved quantities: each photon contained in the light ray bears an amount of energy and momentum and propagates them with the so-called group velocity:

$$v_{gr} = d\omega/dk$$

where  $\omega$  is the angular frequency  
and  $k$  is the wave number.

As far as the flow of photons from the beacon to the wall is concerned, both theories reach the same result. It is a world line because there is no spatio-temporal gap between the position occupied by each photon at successive moments. The theory of causation in terms of transference produces the result that this is a causal process on the basis of the fact that an amount  $h\nu$  of energy is carried by each photon along its trajectory. In Salmon's theory, the same result is obtained by the criterion of the capacity to transmit a mark. What this theory considers as decisive is the *capacity* of a process to transmit a character: it does not matter whether it actually transmits one; however, according to Salmon it is not sufficient to *manifest* a character, but (capacity of) *transmission* is required. Now, photons are capable of transmission: if a coloured filter is introduced somewhere on the path of the light ray between the beacon and the wall, the photons will transmit this modification to the wall.

Second, consider the series of events determined by the position of the phase of a light wave. Suppose that the wave in question is uniform inside the cylinder in the sense that it possesses everywhere the same wavelength and the same amplitude, and that it propagates through a uniform medium (for example air). In that case, the phase of the wave which is determined by a fixed position relatively to the shape of the wave, for example by the position of a "summit", propagates with the so-called phase velocity  $v_{ph}$ , which is determined according to the equation

$$v_{ph} = \lambda v = \omega/k$$

where  $\lambda$  is the wavelength,  
 $v$  the frequency,  
 $\omega$  the angular frequency,  
and  $k$  the wave number.

This law governs the propagation of the phase of the wave, in the sense of expressing a universal proportion  $\omega/k$  that characterizes all events constituting together the wave.<sup>75</sup> However, the phase velocity is an apparent velocity. There is a well-determined property of the wave that propagates with this velocity  $v_{ph}$ , namely its phase, for example the position of a "summit". In a macroscopic wave such as a water wave, the summit is observable and appears to travel at the speed  $v_{ph}$ . By virtue of this appearance, the series is not only a continuous series of events without a gap, i.e. a world line, but also a pseudo-process.

Salmon's theory explains why this series of events is not a causal process by the fact that the phase is not capable of transmitting a mark: if a modification is introduced in the wave at a given point, by a unique interaction, this modification is not propagated with the phase velocity  $v_{ph}$ , but with the group velocity  $v_{gr}$ . The series of events characterized by a constant phase cannot transmit the modification. According to the transference theory, what determines that the series of events is not a causal process is the fact that there is no amount of any conserved quantity carried by the wave which is transferred along that series, and which is therefore carried with the velocity  $v_{ph}$  characterizing it.

Consider now as our third example of a series of events without a gap the apparent movement of the spot of light projected on the wall of the cylinder. The series of events localized at the point where the light spot appears on the wall constitutes for Salmon a paradigmatic pseudo-process. It satisfies the two conditions defining that concept: first, it is a world line exhibiting the structural uniformity that makes it appear as a causal process; second, this appearance is fallacious, in the sense that it is objectively no causal process.

Again, both criteria lead clearly to the correct result: the series of events constituted by the light spot on the wall is no causal process. To show it, we can reason in the following way. The speed of the apparent movement of the light spot ( $\omega 2r\pi$ ) is determined by two variables, first by the angular speed  $\omega$  of rotation of the light source, and second by the ray  $r$  of the cylinder. Therefore, for any given value of  $\omega$ , it is possible to find a size  $r$  so large that the product  $\omega 2r\pi$  represents a speed faster than the speed of light. According to the angular speed  $\omega$ , it may be necessary to use astronomical distances: for a beacon turning around its axis once in a second, it is necessary to consider a wall situated at a distance whose order of magnitude is that of the distance between the Earth and the Moon. However, the fact that the light spot thus projected moves faster than  $c$ , the speed of light in empty space, suffices to show that its movement cannot possibly be a causal process; this result is again reached both by our criterion and by Salmon's. The theory of relativity guarantees that nothing can be transferred at a speed higher than that of light in empty space, neither a mark nor an amount of a conserved quantity. Indeed no amount of a conserved quantity can be transferred between the events of a series whose spatio-temporal

coordinates are such that the spatial distance between two successive events of the series, divided by the temporal distance between them, is larger than  $c$ . Cases of this type provide the motivation for introducing the following criterion (N), which however constitutes only a necessary condition for a world line to be a causal process.<sup>76</sup>

(N) Two events  $c$  and  $e$  can be causally related only if the spatio-temporal distance between them is “timelike”.

The distinction between “timelike”, “spacelike” and “lightlike” distances between events can be understood in the following way. One constructs a Minkowski diagram (see Figure 1.2 above) around a point-like event  $O$ , both “light cones” having their origin in  $O$ . All events  $E$  situated within each of these cones<sup>77</sup> are in a timelike distance from  $O$ . This means that all these events  $E$  may be causally related to  $O$ . The speed necessary to go from  $O$  to  $E$  or from  $E$  to  $O$  is lower than the speed of light in vacuum,  $c$ . The surfaces of the cones constitute limiting cases: the events located on the surfaces can be causally linked to  $O$  only by processes spreading with the speed of light. Their distance with respect to  $O$  is such that it is possible to bridge it only at the maximal physically possible speed  $c$ . For this reason, the events situated on these surfaces are said to be at a “lightlike” distance from  $O$ . Finally, all events happening outside both cones cannot be causally related to  $O$ , by virtue of their spatio-temporal positions. Our criterion (N) is useful precisely with respect to these latter events: it states in terms of a purely geometrical criterion that certain pairs or series of events cannot be causally related: they cannot, simply because the spatio-temporal distance between them is too large. Their distance is said to be “spacelike”.

To see that (N) provides a criterion that is necessary but not sufficient for the relation between two events to be causal, we need only consider two neighbouring space–time regions that are isolated from one another, so that neither influences the other. Between two such regions causal relations are possible although none are actual. (N), taking into consideration only the spatio-temporal distance between the events concerned, cannot determine whether or not there are any causal relations between them.

The criterion (N) is useful as a heuristic tool to the extent that it allows for the detection of some cases of pseudo-processes. The example of the rotating beacon is a case in point. However, (N) is not sufficient to detect all pseudo-processes. One type of pseudo-process that cannot be recognized as such by the application of criterion (N) is the propagation of the phase of a wave if its speed  $v_{ph}$  is lower than  $c$ : *such a propagation remains of course a pseudo-process, independently of its particular speed  $v_{ph}$ .* Note that, although the transference theory and Salmon’s theory reach the same result concerning the apparent movement of the light spot on the wall, they do not take into account the same properties of that world line. Salmon’s theory classifies it as a pseudo-process because it ***appears to us as a causal process.*** By

**contrast, transference theory judges it independently of its appearance, only on the basis of the absence of transference of any amounts of conserved quantities between the successive illuminated spots on the wall. Contrary to Salmon's theory, it only makes use of observer-independent properties.**

Our fourth and last example of a series of events without any gap shows that the distinction established by Salmon between processes and pseudo-processes is not a dichotomy. Let us accept his criterion according to which world lines that are apparently but not really causal are pseudo-processes. Then the distinction between causal processes and pseudo-processes is not exhaustive and cannot therefore be ontologically fundamental. A whole type of world lines remains outside of Salmon's distinction, being neither processes nor pseudo-processes.

To give an example of this type of world line, let us consider a continuous series of events that is constructed in the following way. Let its constitutive events be circles of the same size as the spot on the wall illuminated by the beacon, but located at a certain fixed distance  $d$  from the illuminated spot (the direction being arbitrary but remaining fixed for the whole series), and let the temporal coordinates of these events be identical to those of the continuous series of light spots of the previously described pseudo-process.<sup>78</sup> On the basis of our definitions, the series of events thus characterized is not a causal process. As far as causality is concerned it is on the same footing with the apparent movement of the light spot: No amount of any conserved quantity is transferred from its constitutive events to their successors.

The fact that the series of events we have described lacks, contrary to the series of the preceding example, any phenomenal marks salient for human observers is not relevant for its classification with respect to causation. Our characterization of a process is purely objective in the sense of making no reference to observers. Therefore, the question whether a series of events appears as a causal process to human observers is not relevant for determining whether it is a causal process or not. By contrast, taking the distinction between causal processes and pseudo-processes to be the fundamental distinction of the theory of causation makes Salmon's theory partly subjective. The concept of a pseudo-process contains an essential reference to human observers. However, it is desirable to conceive of the distinction between causal and non-causal as independent of human observers, in order to avoid any anthropocentrism analogous to that of the theory that reduces causation to human action. To be perfectly general, the distinction between causal and non-causal must also be applicable to world lines that are not observed and even not observable. The distinction between process and pseudo-process is not so widely applicable.

We have tried to show two things: first, the concept of a pseudo-process is not apt to ground the fundamental distinction between causal processes and series of events that contain no gap but are nevertheless not causal. It is not because there are series of events that are neither causal processes nor pseudo-processes. Second, the concept of a pseudo-process introduces a

subjective aspect in the theory, which reduces the generality of its domain of application.

Salmon's theory encounters another major difficulty: the definitions of the "capacity to transmit a mark" and of a "causal process" form a vicious circle.<sup>79</sup> A process is something that manifests "consistency of characteristics", in other words continuity of properties. A causal process is a process that is capable of transmitting a mark. The circularity in the theory stems from the impossibility to define the notion of a mark without making use of causal concepts. According to Salmon, a mark is the alteration of a property that is introduced by a single local interaction. However, only *causal* interactions are capable of introducing a mark. For if a non-causal interaction, for example the intersection of two pseudo-processes, could introduce a mark, a pseudo-process could be a cause, but this contradicts the concept of a pseudo-process. If it is part of the concept of a mark that only causal interactions are capable of introducing it into a process, the concept of a causal process has not been analysed without using another explicitly causal concept. Salmon explains the notion of a mark with the help of the concept of causal interaction, and the notion of causal interaction with the help of the concept of a mark.

I do not intend to deny that causal processes are capable of transmitting a mark. The argument I have just offered is only supposed to show that this capacity cannot serve as the conceptual foundation of causation. On the contrary, it is necessary to call upon causal concepts in order to explain what the capacity to transmit a mark consists in. However, the theory of causation in terms of transference allows us to explain *why* causal processes and only causal processes possess that capacity. In an ontology in which events are more fundamental than processes, a mark can be taken to be a property of an event. However, such a property can be transferred to other events if and only if it is the manifestation of a conserved quantity. For example, a sound signal consists either in a wave package or in the local modification of a uniform wave. Both are manifestations of energy. To simplify, imagine that the signal is propagated in a medium without dispersion in which the signal is transmitted without any alteration. If the mark is transmitted without any new external intervention, this is explained by the fact that the amount of energy underlying the mark is transmitted. The process is causal because its constitutive events all contain the same amount of energy. According to our theory, the identity of this quantity of energy explains both the causal character of the process and its capacity to transmit a mark. The transmission of a mark is just a special case of the transmission of an amount of a conserved quantity.

I do not deny the conceptual utility of the capacity to transmit a mark. This capacity provides a criterion, complementary to our criterion (N), allowing us to establish the non-causal character of certain series of events, i.e. of pseudo-processes.



Before I examined the distinction between causal processes and pseudo-processes, I had put forward a thesis it is now time to defend. We have just seen how the transmission of amounts of conserved quantities allows us to explain the transmission of a mark. By virtue of its capacity to provide this explanation, the transmission of amounts of conserved quantities is the more fundamental criterion. What is more, it turns out to be impossible to defend the distinction between processes and pseudo-processes if one denies that amounts of conserved quantities possess an identity through time. Therefore, the alternative account proposed by Dowe (1992b), which grounds the concept of causation on conserved quantities but denies the persistence of individual amounts of conserved quantities, must fail.

Dowe defines a causal process as “a world line of an object which manifests a conserved quantity” (Dowe 1992a, p. 210; 1995a, p. 326). However, the notion of “manifestation” is too weak to justify on its own the distinction between processes and pseudo-processes. Who denies the identity of individual amounts of conserved quantities through time encounters the following problem: there is no way to distinguish “the regular appearance of energy”, for example in the pseudo-process described above where a light spot passes across a wall, from the energy that is “manifested” along a world line that consists of a causal process. First, the spatio-temporal coordinates of events constituting the paradigmatic case of a pseudo-process in which a light spot travels across the wall define a world line. Second, the spot manifests energy whose amount is equal at each stage of its movement – even if the particular amount of energy arriving from the light source at each moment differs from the particular amount that has arrived at the preceding moment: the spot does not consist in a unique particular amount. However, Dowe’s theory does not allow us to make this distinction between a particular identical amount and amounts which are different as particulars but of equal size. Third, concerning the word “object” that figures in Dowe’s definition, it must be interpreted in a very general sense different from its ordinary meaning. This is because Dowe must deny that the persistence of objects be grounded on anything stronger than the continuous possession or manifestation of certain properties. One cannot deny, as does Dowe, every kind of diachronic identity to particular amounts of conserved quantities and at the same time attribute such a diachronic identity to the material objects bearing those quantities. The matter which these objects consists of is after all just one form of appearance of one of the conserved quantities: mass-energy. However, if one conceives as an object every entity that manifests some regularity of any form whatsoever or which, according to Dowe’s criterion, manifests a certain amount of a conserved quantity, one must count the spot on the wall as an object.

One could try to avoid the conclusion that denying the possibility of transmission leads to considering pseudo-processes as causal processes, by imposing stronger restrictions on the notion of an object. Indeed, Dowe tries to make more precise the concept of an object appearing in his

definition of a causal process. He says that “*an object must be wholly present at a time* in order to exist at that time” (Dowe 1995a, p. 329; italics in original). This criterion is supposed to exclude arbitrary temporal aggregates from the class of objects. One could construct such an aggregate by taking the mereological sum of temporal parts of different objects (i.e. objects intuitively considered as such) and using a relational description, such as “the closest ball to the black ball” (cf. Dowe 1995a, p. 328) in a snooker game. The definite description “the closest ball to the black ball” can be satisfied by the pink ball before  $t_1$  and by the red ball after  $t_1$ . The requirement that, in order to count as an object, something must be entirely present at every moment of its existence is supposed to exclude from the category of objects all kinds of relationally defined mereological sums. However, this proposal is circular: what does it mean to say of an arbitrary assemblage of two temporal parts of two different objects, that it is not “entirely present” for example at  $t_0 \leq t_1$ ? Awaiting a more thorough explanation of this notion, it seems to mean simply that the red ball is not the same object as the pink ball, and that it is for this reason that, if you have – with Dowe’s words – the pink ball at  $t_0$ , you don’t have the whole assemblage at that moment. As long as the notion of being entirely present at a moment is not grounded on an independent criterion, it cannot ground the concept of an object in a sense that would allow the exclusion of arbitrary assemblages. The result of this reasoning is that, by refusing to admit transmission, a theory of causation such as Dowe’s lacks the conceptual resources needed to exclude the spot on the wall in Salmon’s thought experiment from the category of objects.

One last reply seems to be open to Dowe: he may try to distinguish the light spot from the portion of the wall on which it appears. Indeed, Dowe (1992c, 1995a) insists on the importance of this distinction. He tries to overcome the problem we have just raised by arguing that it is not the spot that bears conserved quantities but rather the series of portions of the wall on which the spot successively appears. This may seem to allow him to exclude the spot from the category of causal processes for the reason that it does not satisfy the second of the previously listed requirements, i.e. the requirement to carry or manifest conserved quantities. But this move really only pushes the problem one step further rather than solving it because it reappears in a strictly analogous form concerning the series of portions of the wall: let us suppose that each element of this series possesses an equal amount of some conserved quantity. It is clear that there are just as good reasons to consider the series of illuminated portions of the wall as a pseudo-process as there are reasons to consider the light spot itself as a pseudo-process. However, on Dowe’s theory it comes out as a causal process: it is a series of events forming a world line; each event in the series manifests an equal amount of energy because, the rotational movement of the beacon being uniform, an equal number of photons reaches the wall per unit of surface and time. It is of course not “the same” energy that appears

in one event and the following but, according to Dowe's account, the question whether two amounts of energy manifested by two different events are identical or not makes no sense. Finally, the series of illuminated portions of the wall constitutes an object *according to the weak criteria that a theory denying transmission has at its disposal*,<sup>80</sup> given that such a theory denies that individual amounts of mass possess any diachronic identity.<sup>81</sup>

Given the failure of all efforts to avoid the disastrous consequence of classifying the paradigmatic pseudo-process in the category of causal processes, we may consider Dowe's theory of causation as refuted.

### Causation and laws of nature

We have seen that the notion of transference presupposes the existence of a particular class of laws of nature: to be capable of being transmitted, an amount of energy must remain invariant with respect to its change of position, in other words it must be conserved. It is really only because it is subject to a conservation law that energy can play its role in our theory: each time an amount of it is transferred between events  $c$  and  $e$ , they are related as cause and effect.

In order to clarify the relation between causation and laws of nature, we must answer several questions. The first is whether any kind of nomic relation between two events is sufficient for the existence of a causal relation between them or whether a conservation law is required; the second question is whether the instantiation of a conservation law is sufficient for the existence of a causal relation and, if not, what other conceptual element is required.

The answer to the first question is that the instantiation of a law of nature is not sufficient to distinguish causal processes from non-causal world lines, and in particular from pseudo-processes. For example, the series of events determined by the spatio-temporal position of the centre of a wave packet is a causal process, whereas the series of events determined by the position of a particular phase of the wave is only a pseudo-process. An example of the latter type is the series of successive positions of the summit of a wave. However, in both these cases of series of events, one consisting in the successive positions of the centre of a wave packet, the other in the successive positions of the summit (or any other determinate phase) of a wave, there is a law governing the speed of the series. The first propagates with the group velocity  $v_{gp}$ , the second with the phase velocity  $v_{ph}$ .<sup>82</sup> Therefore, the instantiation of a law is not sufficient for a given series of events to count as a causal process.<sup>83</sup> There are laws relating the properties of series of events that fall into the category of pseudo-processes.

Let us now consider the second question concerning the relation between laws of nature and causation. We have already justified our thesis that the instantiation of a conservation law is a necessary condition for two events being causally related, and that the instantiation of an arbitrary kind of nomic relation is not sufficient. We must now ask whether the instantiation

of a conservation law is also sufficient for the existence of a causal relation. It remains to show that the concept of causal relation is richer than the concept of conservation law,<sup>84</sup> so that the first does not reduce the second.

In this respect, our position differs from two alternative proposals as to the relation between conservation laws and causation. On the one hand, Dieks (1981) has noted that each of the concepts of energy, momentum and the other conserved quantities is logically dependent on the notion of a conservation law. From the fact that the concept of law is more fundamental, Dieks concludes that the theory of causation in terms of transference is just a variant of the nomological theory of causation: according to this theory, the existence of a nomic relation between two events is necessary and sufficient for these events to be related as cause and effect.

On the other hand Aronson (1982) holds first that causation is reducible to transference of energy and second that “transference can take place and, hence, energy can flow, even if conservation laws were false” (Aronson 1982, p. 301). Aronson seems to give the conservation laws for energy and momentum a purely epistemic role: “We use them to keep track of a quantity throughout an interaction” (Aronson 1982, p. 301). Fair (1979) also suggests that conservation laws play only an epistemic role: “The conservation laws play the role of supplying criteria for identifying energy and momentum through time” (Fair 1979, p. 239).

What we have said before allows us to show that none of these positions is correct. The conservation of transferred quantities is a necessary condition without which the very idea of transference would not make any sense. However, Aronson correctly criticizes Dieks when he notes that from this “it doesn’t follow that causation is simply a matter of laws, viz., those laws involving conservation” (Aronson 1982, p. 301–302). The reason is, says Aronson, that the laws of nature are symmetrical whereas the relation of cause and effect is asymmetrical. This implies that the concept of causation contains an element that goes beyond mere conservation. However, Aronson does not pursue any further the question as to what is the conceptual element contained in transference that is not reducible to laws of conservation.

The simplest way to identify this further conceptual element is to consider the ontological status of the *terms* of the causal relation, the nature of cause and effect. According to the account developed here, amounts of conserved quantities are transferred between *events*. The most general concept of an entity that can both carry an amount of energy and lose it by transferring it elsewhere without thereby losing its identity is the concept of event. For only a concrete particular<sup>85</sup> can remain identical to itself while undergoing a change in one of its properties. This leaves us, as possible terms of the transference of amounts of conserved quantities, with two categories of entities: objects and events. However, we have already reached the result that the category of events is more fundamental. Every transference between two objects is equivalent to a transference between events, to

the extent that departure and arrival take place at a certain moment: the terms of the causal relation are really, not the objects but the events constituted by these objects at the moment at which they lose or receive an amount of a conserved quantity, for example of energy. Events are the minimal temporal parts of objects required for a transference. According to this reasoning, the choice to consider as terms of the causal relation the temporal parts of objects directly implied in the causal transference, and not the entire objects, is justified by the requirement of parsimony: an adequate theory of causation should identify the simplest possible entities that are capable of playing the role of the terms of the causal relation. For there to be transference and therefore causation, there must be entities between which the transference can take place. These are events. However, the concept of an event is not contained in the concept of a conservation law. In this sense, the concept of causation goes beyond the concept of a conservation law.

We have seen that, contrary to what Dieks (1981) holds, transference theory does not entail the reduction of the concept of causation to that of a conservation law. The instantiation of a conservation law is not sufficient for the existence of a causal relation. But it is also important to insist, against Aronson (1982), on the fact that the validity of conservation laws is necessary for transference. The validity of these laws is a necessary condition for causation. Contrary to what Aronson holds, it is impossible to say of something that is not conserved that it is transferred.

Note that the relation we have just established between laws of conservation and causation, is the consequence of a hypothesis about the notion of causation that goes beyond the commonsense concept of causation. Our aim is to *discover* the particular kind of relation that underlies causation in our world. This implies that causation is a theoretical relation some properties of which may not correspond to any properties of the corresponding commonsense concept. Aronson (1971a) and Fair (1979) pursue a different project: they try to analyse the commonsense concept of causation. It is from this perspective that Fair rejects the idea that causation requires the conservation of the transferred quantities. "There does not seem to be any reason that the commonsense notion of causality should require exact conservation" (Fair 1979, p. 239). It is indeed undeniable that exact conservation is a scientific notion that one would look for in vain as an ingredient of commonsense concepts.<sup>86</sup>

The conceptual dependence between the notions of transference and conservation laws also provides a clear answer to the question of what the exact quantities are of which the transference is sufficient for causation. It is precisely those quantities that are conserved in virtue of a law of nature. We owe our knowledge of these quantities to physics; we cannot pretend to discover them by analysing our prescientific concepts. Every attempt to do this encounters the obstacle of the imprecision of these concepts, which led Hume to the conclusion that the concept of causation cannot be given a

purely objective explication: “I begin with observing that the terms of *efficacy, agency, power, force, energy, necessity, connexion,* and *productive quality,* are all nearly synonymous; and therefore ’tis an absurdity to employ any of them in defining the rest” (Hume 1955, p. 157). We know today that energy is the only conserved quantity in this list. The fact that we have an independent theoretical access to energy and to force, allows us to try to accomplish what Hume considered impossible. Hume is right in saying that it is impossible to *define* causation with the help of one of the terms in his list, because these terms are more or less synonymous; in other words, because the concepts they express are not sufficiently distinct from each other. However, this does not prevent us from trying to produce a theoretical reduction of the notion of causation. For this project it is legitimate to use the *scientific* concepts that correspond to some of the concepts expressed by the words in Hume’s list. Among them, only the word “energy” names a quantity that is subject to a conservation law. Aronson (1971b, p. 145) is wrong to consider that “force” is also capable of being transmitted and that the “transference” of force can constitute a causal relation: there is no scientific concept corresponding to the commonsense concept of force that is subject to a conservation law. It would therefore be wrong to add force to the list of quantities the transference of which is sufficient for causation.<sup>87</sup>

Krajewski (1982, p. 225) makes the mistake of thinking that the transference of *information* is sufficient for causation in the same way as the transference of energy. If it is said that a given transmission conserves information, what is meant is that the transmission channel adds no noise to the signal. However, there is in general no physical quantity that is transmitted. In general, the transference of information is not a physical process although such a physical process can accompany it in certain cases. Information can be transferred by virtue of a simultaneous correlation at some distance between two series of events. However, in such a situation nothing physical can be transmitted between them; indeed, information, insofar as it can be said to be transferred simultaneously at a distance, is not a physical quantity subject to a conservation law.<sup>88</sup>

Others have made the opposite mistake not to take into consideration enough such quantities: Salmon who rejects the criterion of transference as the most fundamental criterion distinguishing causal processes from non-causal world lines, holds that “all and only causal processes transmit energy” (Salmon 1984, p. 146). We need only observe that in an elastic collision, which is a paradigmatic causal relation, only the momentum but not the kinetic energy of the participants in the collision is modified. If two perfectly elastic billiard balls make a central collision, they conserve their kinetic energy and exchange their momentum. In this case, energy is not the relevant quantity.

We therefore agree with Fair (1979) that momentum should be added to the list of quantities transference of which is sufficient for causation. However, it is Dowe (1992a, p. 214) who has proposed the adequate

generalization to conserved quantities – although he rejects the principle of a theory of causation in terms of the *transference* of amounts of such quantities. To the question which quantities are those whose successive manifestation in two events is sufficient for a causal relation between them, Dowe replies that it is those whose amount is conserved by law. What this general point of view acknowledges particularly is the causal character of the transference of electric charge, even if it is not accompanied by the transference of energy or momentum. In this respect, Fair's version which recognizes only energy and momentum is still too restrictive.

Once the conceptual dependence between transference and the conservation of that which is transferred has been clarified, it appears that the correct way to formulate the condition of transference is to make reference to "quantities whose amount is conserved". To ask what exactly these quantities are, beyond energy, momentum, and electric charge, is an empirical and open question. The answer that should be given depends on accomplishing a scientific project that aims at exhaustive knowledge of the class of conservation laws.<sup>89</sup>

## **Conclusion**

We have chosen to undertake the analysis of the concept of causation by following the lead of Hume's account. According to Hume, to judge that an object *A* is the cause of another object *B*, it is necessary and sufficient that *A* is contiguous to *B*, that *A* precedes *B* and that the sequence of these two "objects" regularly recurs in experience. We have however resisted Hume's conclusion that the idea of necessary connection or causation cannot be purely objective. We have therefore rejected the condition of regular repetition, which is central to Hume's subjectivist position. Some causal relations never recur. Objectively, their existence cannot depend on the existence of similar relations that have been observed earlier or later. Objectively, it is not because a sequence of two events is regularly repeated that it is causal; on the contrary, it is because a sequence is causal that it tends to repeat itself.

We have analysed and rejected two accounts that have been proposed as alternatives to Hume's. According to the "singularist" analysis, singular causal relations are ontologically prior to the regularity arising from their repetition; however, it does not provide any account of singular causal relations. The anthropocentric analysis judges that the objective foundation of causation is the human capacity to act: it reduces the notion of effect to the result of a human action. We have found it wanting, to the extent that it does not account for the existence of objective causal relations which are, in principle, outside of the reach of human actions.

This critical analysis has finally led us to a new proposal for the analysis of the causal relation. Causation can be reduced to a physical relation that depends on the transference of an amount of a conserved quantity.

According to our analysis, condition (S) is necessary and sufficient for the existence of causal relations between the events  $c$  and  $e$ .

(S) Two events  $c$  and  $e$  are related as cause and effect if and only if there is at least one physical quantity  $P$ , subject to a conservation law, exemplified in  $c$  and  $e$ , and a determinate amount of which is transferred between  $c$  and  $e$ .

(S) is supposed to provide an a posteriori reduction of the causal relation. This relation is not conceptually identical to the relation of transference. Rather, if our analysis is correct, causal relations form a natural kind whose members share the property of being transmissions of amounts of conserved quantities.

Then we have developed the most important consequences that follow from our reduction (S) of the causal relation. First, the reduction of causation to transference justifies the idea that the cause must be either contiguous with the effect or related to it by an uninterrupted chain of intermediate events. We have tried to defend this consequence in the light of certain results of contemporary physics. In particular, with respect to quantum correlations at a distance, we have proposed to consider these relations as non-causal, and thus to stay with the traditional idea that there are no causal interactions at a distance.

Second, we have shown that our notion of transference is not an intrinsically asymmetrical relation. The causal relations in the actual world are asymmetrical, but not all are intrinsically asymmetrical. One can make sense of this by supposing that all causal relations are part of a single network. It is a contingent fact that a large part of the causal relations in the actual world are intrinsically asymmetrical: these are the processes in which entropy increases. Through their common inclusion in the network, those causal relations that are intrinsically symmetrical inherit the asymmetry of the intrinsically asymmetric processes. This explanation allows us to reconcile the following ideas. First, *symmetrical* causal relations are conceptually possible; they may exist in causally isolated regions of the actual world or in other possible worlds. Second, even in the actual world there may be “backwards”, or “reverse”, causal relations, i.e. causal relations in which the effect precedes the cause. If they exist, such causal relations have a direction of transference that is opposed to the direction of transference characterizing the vast majority of causal relations in the network. Finally, our account leaves us free to construct a theory of time that explains its asymmetry on the same basis as the asymmetry of causation. In particular, our analysis does not need to presuppose the asymmetry of time in order to explain the asymmetry of causation, something that has been objected to earlier versions of transference theory.

Third, our analysis justifies Hume’s thesis that there is no simultaneous causation, although for reasons different from Hume’s. Our claim is that, in



all cases of apparently simultaneous causation, either it turns out that the effect is not *strictly* simultaneous with the cause, or it turns out not to be a causal *relation*, in the sense that “the cause” and “the effect” are really only one event. In other words, the simultaneity is only apparent in those cases in which the effect is spatio-temporally separated from the cause, and in those cases in which cause and effect coincide spatio-temporally, there is strictly speaking no causation. The justification of the latter thesis depends on the specification of the nature of the terms of the causal relation. We conceive them as events, understood as the content of space–time zones.

On this basis we could then reconsider Hume’s position according to which cause and effect are *objects*. On the one hand, it is often enough to say, in an imprecise but satisfying way, that objects are causes (and effects), to the extent that the precise and correct way of stating this is: causes and effects are most of the time *temporal parts* of objects. Each causal interaction engages an object only for a moment. However, the object at that moment – a “temporal part” of it – is an event. On the other hand, the choice imposed by transference theory to consider events as the terms of the causal relation allows us to explain the persistence of objects through time on a causal basis. Objects persist because each of their temporal parts is related by the transference of mass to its neighbouring temporal parts. This conception of genidentity as a causal process can take over the role of the traditional concept of substance.

We have tried to reply to the objection that individual amounts of conserved quantities lack identity through time. To make sense of the idea that an individual amount is transferred, it is necessary that the amount remains the same during its transmission. The fact that the quantity of which it is an amount, is subject to a local conservation law allows us to conceive its identity through time according to the model of a *flow* of this quantity. Outside of interactions bringing into play different amounts of the same quantity, the amount underlying a given flow persists through time as the same individual amount.

We have also rejected an epistemological version of this objection, according to which it is impossible to *recognize* that two appearances of an amount of a given conserved quantity at different moments are indeed appearances of the same individual amount. We can acknowledge that it may be difficult or even impossible to know whether a given process is really causal or only *appears* as such. Our reply to this *epistemic* difficulty is that it does not throw doubt on the well-grounded character of the *objective* distinction between causal processes and processes that are only apparently causal.

We have obtained a further result concerning the latter type of processes that have been called “pseudo-processes” by Wesley Salmon. On the one hand, our criterion (S) allows us to justify and explain the conceptual distinction between (causal) processes and pseudo-processes, whereas the rejection (among others, by Salmon himself) of the idea that there can be transference of conserved quantities destroys the foundation of this distinc-

tion. On the other hand, we have been able to show that the account of causal processes offered by Dowe is inadequate because it leads to categorizing certain pseudo-processes as causal processes.

Finally, we have established that the instantiation of a law of nature is necessary but not sufficient for the existence of a causal relation. The instantiation of a conservation law is necessary because it is required by the condition that what is transferred from cause and effect must be conserved during the time of transmission. However, the instantiation of a law is not sufficient to distinguish pseudo-processes from causal processes. The propagation of the phase of a wave provides a counterexample to this hypothesis. For this reason, our analysis does not fall into the category of the “nomological theory of causation” which constitutes one of the most important contemporary paradigms of the theory of causation. According to this account, the existence of a *nomic* relation between two events is necessary and sufficient for these events to be related as cause and effect. Later (in Chapter 5) we shall raise detailed objections against the nomological theory of causation. However, the theory we have defended can be said to be “nomological” in the weak sense that it presupposes that what is transferred in a causal relation must be conserved by virtue of *law*.

## 2 Laws of nature and universal generalizations

The aim of this chapter is to clarify the notion of a law of nature, by opposing it to that of an accidental universal correlation. We shall analyse the semantic difference between statements of law and other statements of universal scope. It will result from this comparison that it is neither valid to infer the existence of a nomological relation from the existence of a universal correlation, nor to infer from a law that the properties it relates are universally coinstantiated. The first claim is widely accepted but the second is incompatible with most currently defended accounts of laws: in effect, it is equivalent to the claim that most laws of nature allow for exceptions under certain circumstances.

### **The semantic difference between nomological statements and accidental universal generalizations**

The analysis of the concept of law of nature requires the clarification of the distinction between universal generalizations that are accidentally true and others that express a law. The difference between these categories of universal generalizations is not manifest in the *logical form* of their statements.<sup>1</sup> The difference must therefore be due to their *interpretation*. Statements of universal form can both express laws or generalizations whose truth is due to a coincidence. The simplest logical form of such a universal statement

(U1) All *As* are *Bs*,

is the following:

(U)  $(x) (Ax \rightarrow Bx)$ .<sup>2</sup>

Here is an example of a universal generalization:

(1) All electrons are attracted by positive charges.

The universal scope of the statement need not be made explicit by a quantifier expression such as “all”, as can be seen from this example:

(2) Beavers build dams.

We shall take it for granted that the generic form

(U2) *As are Bs.*

is logically and semantically equivalent to the explicitly universal form (U).<sup>3</sup>

The main thesis I shall defend is this. In the usual interpretation of universal statements of form (U), in which the domain of the variable bound by the universal quantifier is taken to consist of *concrete objects*, they do not express laws of nature. As we shall see in detail, the existence of a law according to which “all *A* are *B*” does not guarantee that *every object* possessing property *A* also possesses property *B*. One can express this fact simply by saying that the validity of a law does not exclude the possibility of *exceptional situations*. The difference between universal generalizations and laws can be understood in the following way: if the domain of the variable *x* is taken to be the domain of concrete objects, (U) says *of all objects* that if they possess the antecedent property *A*, they also possess the consequent property *B*. However, if the domain of *x* is taken to consist of the *particular instances* of properties *A* and *B*, the statement receives a specifically nomological interpretation, in other words it is understood as expressing a law of nature. Then it is interpreted as saying that all sufficiently pure instances of *A* are accompanied by an instance of *B*. The main reason for this thesis is that most laws, but not accidental generalizations, allow for exceptions. Very few laws are true if one takes the variable that corresponds, in their logical form, to *x* in (U), to refer to concrete objects: we may call such laws *strict*. In the rest of this chapter and the next we shall elaborate this strategy for interpreting law statements and accounting for exceptions.

A word of terminology first. We say of a yellow lemon – which is a concrete<sup>4</sup> object – that it *possesses* or *exemplifies* the property of being yellow, and that it is constituted by a certain number of property *instances*, such as this instance of yellow, this instance of being round, etc. This conception allows us to say that if it is a law that all lemons have seeds, and if this is a lemon, it is nevertheless possible that it has no seeds. In such a case we say that the lemon exemplifies the property of being a lemon in an insufficiently pure way, in the sense that other properties of the lemon or of the circumstances interfere and prevent the lemon from having the consequent property of having seeds.

We shall evaluate the conception thus sketched by its capacity to take account of the properties of laws of nature. For example, law statements, contrary to accidental universal generalizations, do not conserve their truth value under substitution of coreferential predicates. This chapter has two

aims: first, to establish the distinction between laws – or nomological generalizations – and accidental generalizations; second, to use this distinction for explaining the specific properties of these two types of generalization. The aim of the following chapter will be to show the capacity of this conception of laws to account for the phenomenon of exceptions: it is possible and even common that a law is valid although it has exceptions.

To simplify the analysis proposed in this chapter let us provisionally suppose that we deal with exceptionless laws. This allows us to focus our attention on the explanation of the difference between laws and accidental generalizations. The first point to note is that the *logical form* of a universal statement of type (U) contains no indication of whether the correlation it expresses is accidental or nomic. In other words, (U) admits both interpretations: by virtue of its logical form, the correlation it expresses can be either accidental or nomic.

It is of course possible to state explicitly that a generalization is intended to express a law of nature, with a statement of the following form:

(L) It is a law that *As* are *Bs*,

For example:

(1a) It is a law that electrons are attracted by positive charges.

The explicitly nomic form (L) allows us to ask our question more precisely: what is the difference between the relation expressed by (U) and the relation expressed by (L)? Within the framework of an empiricist position, according to which laws are nothing else than universal regularities, the answer is very simple: there is no semantic difference between (U) and (L). In the tradition of logical empiricism,<sup>5</sup> a law of nature is considered to be a universal regularity. In other words the content of the law statement (L) is considered to be the same as the content expressed by the corresponding universal generalization (U).

This still leaves open two options. The more radical possibility is to deny any difference between (U) and (L), in which case clarity demands to avoid the redundant form (L) and to use only the universal formula (U). However, it is difficult for someone adopting this position to explain the difference between accidental universal correlations and nomic relations that is acknowledged on all hands, and some of whose aspects we shall analyse in a moment. There is, in particular, a consensus to accept that there is a difference in type between the following two statements:<sup>6</sup>

(3) All objects in massive gold have a volume of less than a km<sup>3</sup>.

(4) All objects in massive uranium-235 have a volume of less than a km<sup>3</sup>.

Statements (3) and (4) may illustrate the distinction between nomic and accidental universal generalizations, because the former is true by virtue of a coincidence, and thus expresses a contingent fact whereas the latter is true by virtue of a law of nature. It is appropriate to interpret the former by assigning the variable bound by the universal quantifier to the domain of concrete objects: it so happens that all objects that have the property of being of massive gold also have a property of having a volume smaller than a  $\text{km}^3$ . However, the truth of (4) is not due to a coincidence. (4) can not only be interpreted in the same manner as (3), as a universal statement bearing on objects. Rather, it is also possible to show that the universal regularity expressed by (4) is grounded on a law. The law itself can be expressed either by prefixing to (4) the operator “it is a law that”, as in (4’), or by specifying that the domain of universal generalization is constituted by the properties themselves, as in (4’’):

(4’) It is a law that all objects in massive uranium-235 have a volume of less than a  $\text{km}^3$ .

(4’’) The property of being an object composed of massive uranium-235 is subordinate to the property of having a volume of less than a  $\text{km}^3$ .

Otherwise widely differing positions such as those of Goodman (1955) and Braithwaite (1959),<sup>7</sup> on the one hand, and those of Armstrong (1983) and Tooley (1987), on the other, agree in acknowledging the difference between laws and accidental universal correlations. However, these authors offer radically divergent explanations of the origin of that difference. The nominalist position of Goodman and Braithwaite consists in acknowledging the difference but denying its absoluteness: according to them, the dichotomy between laws and other universal generalizations is theory-relative. The difference between laws and accidental relations stems from the roles their respective statements play in science. The nominalists acknowledge the existence of a difference between the law expressed by (L) and the universal correlation expressed by (U) but they take the difference to be grounded neither in logical form nor in interpretation. Rather, they take it to have its origin in theoretical context. The criterion that singles out laws among true generalizations is that the statement of the former but not of the latter is contained in, or may be deduced from, a theory actually accepted by the scientific community. This position has the advantage of allowing us to express nomological statements in a purely extensional language. However the price to pay is that it requires abandoning the idea of an objective, actual (i.e. not only counterfactual) and absolute foundation of the distinction between laws of nature and universal but accidental regularities. For some nominalists, such as Goodman (1955) and Braithwaite (1959), the theories used as a criterion to distinguish laws from other universal generalizations are those that are accepted by the present scientific community.

Others, such as Ramsey (1928) and David Lewis (1973; 1983; 1994) hold that the distinction must be grounded on an *ideal* scientific theory.<sup>8</sup> Against this position I shall argue that, in the case of non-strict laws, there is a difference in truth value between a statement of type (L) and a statement of type (U) if both contain exactly the same predicates.

But there are also other conceptions of laws that take the semantic content of a law statement to go beyond the content of a universal generalization. One important proposal is to say that a law statement expresses that the consequent is *necessarily* implied by the antecedent;<sup>9</sup> according to the realist conception of laws,<sup>10</sup> this necessary implication can be deduced from an effective relation between two universals, referred to respectively as the antecedent and the consequent predicate. The realist theory of laws tries to account for the difference between laws and accidental relations by making the hypothesis that the former create *necessary* correlations between the instances of the properties referred to by the antecedent and the consequent. Rather than postulating directly (following Kneale 1949; 1961) that in the case of laws the antecedent necessarily implies the consequent, the realist approach tries to explain the origin of this necessity.

What distinguishes all these accounts from the nominalist position is the idea that statements of law and statements of accidental universal correlations have objectively different content. According to nomological realism, the difference between law statements of type (L) and universal generalizations of type (U) is semantic, not only epistemic or pragmatic. (L) has a richer content than (U) in the sense that

(5) (L) implies (U)

but

(6) (U) does not imply (L).

The aim of this chapter is to make (6) plausible and to look at some reasons pleading for (5). In the next chapter we shall throw doubt on the plausibility of (5).

Here is an argument establishing that there is a difference in content between (L) and (U).<sup>11</sup> If we compare statements (1) and (1a), we find that the universal correlation expressed by (1) is not accidental but is rather the consequence of the nomic relation expressed by (1a). In other words, (1) is true because (1a) is true. Now let us compare (1) to (7):

(7) All particles having a mass equivalent to 511 keV are attracted by positive charges.

The universal correlation expressed by (7), however real, is not due to a nomic relation; on the contrary, the transformation of (7) into (7a), by

prefixing the expression “it is a law that” to (7), leads to the false statement (7a):

(7a) It is a law that all particles having a mass equivalent to 511 keV are attracted by positive charges.

The example of (7) shows that one cannot in general infer the truth of a statement of type (L) from a statement of type (U) with the same predicates. However, (1) and (7) have the same truth conditions because they have the same logical form and because the predicates they contain are coreferential (in other words, express coextensive concepts).<sup>12</sup> How then, can we explain that (7a) is false whereas (1a) is true? This difference of truth value can be explained in terms of the difference between statements of type (U) and statements containing the same predicates, but of type (L). One way to explain why statements like (1a) and (7a) differ in truth value although they differ only by coreferential predicates, is to say that these predicates are, in a statement of type (L), embedded in an *opaque context*.<sup>13</sup> The opaque context is created by the operator “it is a law that ...”.

One can find this idea clearly expressed by Dretske (1977b).

The class of universal truths is closed under the operation of coextensive predicate substitution. Such is *not* the case with laws. [ ... ] We may say that the predicate positions in a statement of law are *opaque* while the predicate positions in a universal truth of the form  $(x)(Fx \supset Gx)$  are *transparent*. [ ... ] To refer to something as a statement of law is to refer to it as an expression in which the descriptive terms occupy opaque positions.

(Dretske 1977b, p. 250, italics Dretske’s)

To make the idea behind Dretske’s analysis more explicit, let us begin by considering the concept of a *transparent context*.<sup>14</sup> The predicates contained in a universal generalization of type (U) are in a referentially transparent context. Take example (1). The truth of this general statement implies the truth of all statements obtained by (1), through substitution of coreferential terms. Let us accept that all and only electrons in fact possess a mass equivalent to 511 keV. Then, by virtue of the transparency of statement (1), its truth guarantees the truth of (7). The validity of the inference from (1) to (7), which differs from (1) only by the substitution of a coreferential predicate, is exclusively due to the logical form of (1). This means that all statements of type (U) allow this kind of inference. The validity of such inferences is one of the classic criteria for the transparency of the context in which the predicates undergoing substitution are embedded: a context is transparent if it allows us to make substitutions of coreferential terms *salva veritate*, i.e. without possible change in truth value.



It is different with nomological statements of type (L). We have seen that substituting a coreferential predicate for one of the predicates contained in the true nomological statement (1a), yields the false statement (7a). Dretske's hypothesis is that universal generalizations of the form (U) are transparent whereas law statements of the form (L) are opaque. This accounts for the fact that the former but not the latter always preserve their truth value when their predicates are replaced by other coreferential predicates. One can express this point by saying that the class of universal truths but not the class of law statements is closed under the substitution of coreferential predicates.<sup>15</sup> However, existential quantification is a valid inference neither for universal generalizations nor for law statements. To the extent that universal statements express material implications, their truth does not entail the *existence* of any objects satisfying their antecedent predicate. None of these types of statement can therefore be called "transparent" in the strict sense.<sup>16</sup>

### **The Fregean conception**

The distinction between opaque and transparent contexts, which we have used to ground the distinction between laws and accidental generalizations, comes from Frege. Thus it seems promising to see whether Frege's philosophy contains a satisfying explanation of the origin of the opacity of nomological statements. For Frege, they are *second-level* statements. Indeed, it is possible to explain certain properties of law statements by the hypothesis that they are statements expressing the *subordination of concepts*. Instead of attributing, in the way of first-level statements, a property to an object, that is, saying of an object that it falls under the concept expressed by the predicate, a nomological statement expresses a relation of subordination between the concepts themselves. Frege's conception of the logical relation between predicates, concepts and objects allows us to explain this important property of law statements: in such statements, the substitution of a predicate for a coreferential predicate does not in general yield a law statement.

In a singular statement, the subject makes reference to an object and the predicate makes reference to a concept. If one says that "the Atlantic ocean is salty", one says that that ocean – a particular object – falls under the concept of being salty, or satisfies the predicate "is salty". However, Frege endeavours to prevent the misunderstanding that this analysis is adequate for all statements. In particular, it is inappropriate for the universal statements we are concerned with here. In their case, it would be a mistake to think that the word playing the role of the grammatical subject denotes an object: rather, it denotes a concept, in the same way as the grammatical predicate. Frege calls the concept denoted by the grammatical subject "subject-concept" (Frege 1983, p. 130; transl. p. 119). What characterizes general statements, and in particular nomological statements, is the fact that the grammatical subject has a predicative nature, in other words expresses a concept.

In “Sense and Meaning” (Frege 1892b), Frege introduces the distinction between the sense and the denotation (or reference) of proper names. He further generalizes (Frege 1892a) the concept of denotation so that it can be applied also to predicates: A predicate expresses or denotes a concept. “A concept is what is meant by a predicate; an object is something that can never be the total meaning of a predicate, but can be what a subject means” (Frege 1892a, p. 72; transl. pp. 47–48). Finally, in a posthumously published text he devotes to this topic, Frege (1983) analyses the logical role played by predicates and the concepts they express by distinguishing between sense and denotation of a predicate. “In an article (*Über Sinn und Bedeutung*) I distinguished between sense and meaning in the first instance only for the case of proper names (or, if one prefers, singular terms). The same distinction can also be drawn for concept-words.” (Frege 1983, p. 128; transl. p. 118). As long as the logical relation between predicates and concepts is not sufficiently clear, one may be tempted to content oneself with a purely extensional conception of logic and to confuse “the division into concepts and objects with the distinction between sense and meaning, so that we run together sense and concept on the one hand and meaning and object on the other” (Frege 1983, p. 128; transl. p. 118). Frege undertakes to remedy this confusion.

A proper name is referentially complete, in the sense of being alone sufficient to denote an object. The expression “evening star” denotes a certain planet, namely Venus. It also has a sense, which differs from the sense of the expression “morning star”, although both share their reference. As a proper name, the “word ‘Venus’ [ . . . ] can never be a proper predicate, although it can form part of a predicate.” (Frege 1892a, p. 69; transl. p. 44). To form a predicate, one must at least associate the copula “is” to the term. “Is Venus” is a predicate; it is “unsaturated” (*ungesättigt*, Frege 1892a, p. 80; transl. p. 54) by virtue of its predicative nature. The concept denoted by the predicate inherits its unsaturated character. Therefore, says Frege, “a concept (as I understand the word) is predicative” (Frege 1892a, p. 67; transl. p. 43).

The concepts expressed by predicates share the logical property of being unsaturated with *functions* (Frege 1891; 1904). “Statements in general, just like equations or inequalities or expressions in Analysis can be imagined to be split up into two parts; one complete in itself, and the other in need of supplementation, or ‘unsaturated’. [ . . . ] Here too I give the name ‘function’ to what this ‘unsaturated’ part stands for” (Frege 1891, p. 29; transl. p. 31). More directly still: “What in the case of a function is called unsaturatedness, we may, in the case of a concept, call its predicative nature” (Frege 1983, p. 129; transl. p. 119). As soon as the expression of the concept or function is saturated by a proper name (or by another singular term; cf. Frege 1891, p. 29; transl. p. 31) the complex proposition thus obtained denotes one of the two truth values. Either the object to which the proper name makes reference falls under the concept, in which case the proposition is true, or it does not fall under the concept, in which case the

proposition is false. Statements of this type are *first-level*, which means that they bear on an object, of which they say that it falls under, the concept expressed by the predicate.

It is fundamental for understanding nomological statements that there are also *second-level* statements. They express a relation between two concepts: this is, according to Frege, a “second-level relation” (Frege 1983, p. 131; transl. p. 120; Cf. Frege 1884, §§ 46, 53). Frege distinguishes three cases: the relation between a second-level concept and a first-order concept falling under it, the relation of subordination of a first-level concept under another (Frege 1892a, p. 73; transl. p. 48) and the (second-level) relation of equality between two first-level concepts (Frege 1983, p. 131). As an example of the first case, he offers the statement “the concept *square root of 4* is realized” (Frege 1892a, p. 74; italics Frege’s; transl. p. 50). Here, the predicate “is realized”, which has the same sense as “is not empty”, or more explicitly “does not have an empty extension”, expresses a second-level concept. The only “objects” which this concept can apply to are first-level concepts, in Frege’s example, the concept *square root of four*. However, we must put the word “object” between quotes, because Frege takes great pains to warn against the danger of speaking in this way: first-level concepts are not, and can never become, objects, precisely because they are unsaturated, in other words because they have a predicative nature. However, Frege himself sometimes speaks in this misleading way: “In my way of speaking expressions like ‘the concept F’ designate not concepts but objects” (Frege 1892a, p. 73; transl. p. 48). This usage can be justified by simply informing the reader of the possibility of being misled (cf. Frege 1983, p. 133). In the example quoted above, of a statement expressing a subordination relation between concepts, a first-level concept (the concept *square root of four*) falls under a second-level concept of a logical type.<sup>17</sup> The second-level concept expresses the relation between a (first-level) concept and the objects falling under it: it denotes the relation of realization or satisfaction.

The logical structure of universal statements is different: these statements form a second type of second-level statements: they express the subordination of the first-level concept under another first-level concept. “This relationship [between a concept and a second-level concept under which it falls], however, should not be confused with the subordination of species to genus” (Frege 1884, p. 65; transl. p. 65e). For Frege, all general statements containing quantifying expressions such as “all”, “every”, “no”, or “some”, belong to this kind of statements of subordination. Frege distinguishes a subclass of them as a third type of second-level statements: they say of two first-order concepts that they have the same extension. All such statements belong to the second-level because they do not bear on the objects satisfying the predicates they contain, but on the concepts expressed by these predicates. They express a relation between two concepts, either the relation of subordination of a concept under another (Frege 1884, p. 60f.; 1892a, p. 72f.; transl. p. 48) or the relation of equality (i.e. coextensionality) between

concepts (Frege 1983, p. 131). In the *Foundations of Arithmetics*, Frege considers the following example of a universal statement:

(8) All whales are mammals.

According to Frege, this statement does not, contrary to appearances, bear on objects, i.e. whales, but on *concepts*. It expresses the subordination of a concept under another, i.e. the subordination of the concept of whale under the concept of mammal. In the context of our enquiry, the following aspect of Frege's analysis is crucial: the fact that statements of type (8) do not bear on objects but rather on concepts, is exclusively due to their generality. As a consequence, it is independent of whether the general statement under consideration is nomological or true by accident. The crucial feature of Frege's analysis is the idea that, as far as general statements are concerned, not only the grammatical predicate, but also the grammatical subject, has a predicative nature and thus expresses a concept. Of the grammatical subject Frege says:

But even in such cases [i.e. cases in which the concept occupies the subject position within the proposition; M.K.], e.g. in the sentence "all mammals have red blood", we cannot fail to recognize the predicative nature of the concept [expressed by the grammatical subject; M.K.]; for we could say instead: "whatever is a mammal has red blood" or "if anything is a mammal, then it has red blood".

(Frege 1892a, p. 72; transl. p. 47)

The following passage contains the explicit statement that it is by virtue of the element of generality due to the presence of a "subject-concept" in a statement that it bears on concepts rather than on objects: "In universal and particular affirmative and negative sentences, we are expressing relations between concepts; we use these words ['all', 'any', 'no', 'some'; M.K.] to indicate the special kind of relation" (Frege 1892a, p. 72; transl. p. 48). Furthermore, the arguments Frege offers in favour of this thesis are only based on the generality of the statement, but are independent of its nomological or accidental character. He says of statement (8):

It is true that at first sight the proposition "All whales are mammals" seems to be not about concepts but about animals; but if we ask which animal we are speaking of, we are unable to point out any one in particular. Even supposing a whale is before us, our proposition still does not state anything about it. We cannot infer from it that the animal before us is a mammal without the additional premise that it is a whale, as to which our proposition says nothing. [ . . . ] The word "whale" is not a name for any individual creature.

(Frege 1884, p. 60; transl. p. 60e)

Once we have clearly brought out into the open that Frege's analysis attributes the same logical form to all universal statements, it is obvious that it can provide no answer to our problem. The analysis Frege offers succeeds in explaining the opacity of nomological statements: it is an immediate consequence of the fact that they do not bear on objects but rather on the concepts expressed by their predicates, that the substitution of a predicate by another denoting a different concept, even if it is coextensional, does not preserve the truth value of the statement: two concepts can be different even if they share their extension (cf. Frege 1983, p. 131).

Frege's analysis cannot solve our problem for two reasons. First, it treats accidental generalizations exactly in the same way as nomological propositions. Indeed, it attributes the same logical structure to statements (1) and (7): both express a general proposition and are therefore opaque with respect to substitution. This does not give us any indication of the origin of the difference between (1) and (7) we are looking for: (1) but not (7) expresses a nomic correlation; (1) but not (7) can be transformed into a true and explicitly nomological statement (1a), by prefixing "it is a law that". Worse, the result of Frege's analysis seems to contradict our result according to which the generalization (1) becomes opaque only at the moment in which its nomological character is made explicit by the expression "it is a law that". Contrary to what Frege's analysis implies, we have found that (1) and (7) *are* transparent, as long as one does not take into consideration their nomological status, but only their truth: indeed, (7) can be obtained from (1) by replacing the predicate "is an electron" by the predicate "has a mass equivalent to 511 keV" that denotes a different but coextensional concept. This shows its transparency with respect to the substitution of coreferential predicates.

The second reason for which Frege's analysis cannot satisfy us concerns an aspect that it shares with most other analyses of laws. In effect, if a law statement expresses a relation of subordination between concepts, it implies a general statement of the form (U). Indeed, Frege *identifies* statements of form (U) with statements expressing a relation of subordination between concepts. However, this consequence disqualifies this analysis with respect to laws that have exceptions. In fact, as I shall argue in Chapter 3 below, most laws of nature have exceptions; this constitutes a reason to deny that their validity implies the truth of a universal generalization of the form (U).

### **Dretske's account of laws of nature**

We have already introduced the fundamental idea of Dretske's (1977b) account of laws. Its most important aim is to overcome the first difficulty we have just mentioned: Frege's analysis does not explain the difference between accidentally true universal generalizations and law statements. The analysis proposed by Dretske contradicts Frege's analysis with respect to the following crucial point. Dretske defends the "extensionalist" point of view

Frege (1892a; 1983) criticizes, according to which, in general statements of the type “all *F*s are *G*s”, the predicates refer to the objects that satisfy them. We have just seen that Frege rejects this account, his own analysis being incompatible with it. As we have seen, Frege takes “*F*” and “*G*” in “all *F*s are *G*s” to refer to concepts, not to the objects falling under them.

It is important to underscore the incompatibility between Frege’s and Dretske’s analyses, given that the latter introduces his own by an explicit reference to Frege, beginning his presentation with the words: “taking our cue from Frege” (Dretske 1977b, p. 252). This remark is justified by the fact that the fundamental idea on which Dretske builds his analysis is to take advantage of Frege’s (1892b) distinction between opaque and transparent contexts. Frege has introduced this distinction in the context of a different problem: to analyse the semantics of statements containing indirect, or “oblique”, discourse. However, Frege does not apply this distinction to statements of laws of nature. Dretske’s idea is to take up Frege’s distinction between opaque and transparent contexts, and to make the hypothesis that in nomological statements, the predicates are embedded in an opaque context, whereas they occupy transparent positions in accidental universal generalizations. Dretske argues in the following way. A non-nomological universal generalization can be analysed according to the form (U).

$$(U) (x) (Fx \rightarrow Gx).$$

According to Dretske, a universal material implication of this form bears on the objects in the extension of the concept *F*, and not, as in Frege’s analysis, on the concepts *F* and *G*. This explains why the substitution of a predicate “*K*” to “*F*”, where both are coextensional, does not alter the true value of the generalization. If (U) is true, then

$$(U^*) (x) (Kx \rightarrow Gx).$$

is also true. This is according to Dretske equivalent to saying that the predicates “*F*” and “*G*” occupy in (U) transparent positions.<sup>18</sup>

By contrast, the operator “it is a law that” creates an opaque context in which the predicates do not have their usual reference; therefore the substitution of coreferential predicates within such a statement does not necessarily preserve its truth value. To say that it is a law that *F*s are *G*s, is to say that the statement “all *F*s are *G*s” must not be understood as a universal statement bearing on the extension of the concepts *F* and *G*, but rather as a singular statement expressing a relation between the universal properties “*F*-ness” and “*G*-ness”.<sup>19</sup> In other words, Dretske (1977b, p. 252f.) takes the statement

$$(L) \text{ it is a law that } F\text{s are } G\text{s}$$

to be equivalent to the following statement bearing explicitly on properties, which are interpreted as universals:

(P)  $F\text{-ness} \rightarrow G\text{-ness}$ .

Dretske remains faithful neither to Frege's analysis nor to Frege's terminology. Comparing his analysis with Frege's, we can say that Dretske takes over Frege's analysis precisely and exclusively for law statements, whereas he rejects it for accidental universal generalizations. He elaborates this differential treatment of the two types of statement, by making use of Frege's concept of an opaque context. If a universal statement is preceded by the clause "it is a law that", it works as an operator creating an opaque context. Within it, the predicates " $F$ " and " $G$ " do not have their usual reference, i.e. according to Dretske, the objects in the extension of the concepts  $F$  and  $G$ . The predicates undergo the same semantic modification as the embedded statement as a whole, and with it all the words it contains: in an opaque context, they refer to what is otherwise their *sense*.

Recall what is, according to Frege, the semantics of expressions embedded in an opaque context. In a transparent context, the substitution of a singular term for a coreferential expression does not change the truth value of the statement. This is also true for the predicates of first-level statements: in a transparent context, the substitution of a predicate for a coreferential predicate does not change the truth value of the statement. However, it does not hold with respect to subordinate clauses embedded in an opaque context. Frege shows that the truth of a statement containing a subordinate clause in an opaque context is independent of the truth value of this subordinate clause. Frege takes the reference of a statement to be its truth value. In certain contexts, one cannot substitute a subordinate clause for another having the same denotation (i.e., according to Frege, the same truth value). These are the contexts Frege calls opaque.<sup>20</sup> If all statements were transparent with respect to all their constituents, then, says Frege, the truth value of a statement containing another statement as a part would not change under the substitution of a subordinate clause for another clause with the same truth value (Cf. Frege 1892b, pp. 50–51; transl. pp. 66–67). However, this is not so in indirect discourse and in contexts of propositional attitudes.

As soon as it is embedded in an opaque context, such a clause does not anymore make reference to its truth value as it does usually, but to the thought (*Gedanke*) it expresses (Frege 1892b, p. 51; transl. p. 67). This implies, according to Frege, that both the whole subordinate clause and the words it contains change their reference: in an opaque context, the reference of an expression is what constitutes, in ordinary circumstances, its sense (Frege 1892b, p. 43; transl. p. 59).

Let us take up the thread of our analysis of Dretske's account of laws and adopt Dretske's "extensionalist" and explicitly anti-Fregean interpretation,

according to which predicates refer to the *objects* in the extension of the concepts they express. In this framework, the concept expressed by a predicate constitutes its sense. The coherence of the analysis of the semantics of phrases embedded in opaque contexts then requires drawing the conclusion that, in an opaque context, the semantic properties of predicates are modified as well. The concept that is in a transparent context the *sense* of a predicate, becomes in an opaque context its *reference*. According to Dretske, this is exactly what happens to the predicates of a statement of universal implication of type (U) if this statement is embedded in the context created by the operator “it is a law that”, so as to form a statement of type (L).

We have reached the following result: we retain for nomological statements and only for them, the analysis Frege proposes for all universal statements. This is possible if we adopt the extensionalist position with respect to the reference of predicates, and make the hypothesis that the form of nomological statements is given by (L). This form contains the operator “it is a law that”, which creates an opaque context. In the case of nomological statements, Frege’s realist interpretation is vindicated: predicates make reference to properties (concepts), which are interpreted as universals existing objectively, independently of the predicates expressing them. (Cf. Frege 1844, p. 60f.)

However, this entails that if Frege’s analysis is appropriate for statements such as

- (8) All whales are mammals

this is so simply by virtue of their universal form. Rather, we have found that this analysis is appropriate to the extent that the statement is given an explicitly *nomological* meaning. In other words, a universal statement *insofar as it is interpreted as nomological* bears on properties, not on the objects in the extension of the properties. Insofar as (8) is taken to express a law, the universal correlation stems from the subordination of the property of being a whale under the property of being a mammal.

The contrast between this conception and Frege’s own comes out more clearly if we compare the two analyses with respect to a universal statement that is true only by coincidence, in other words that is *not* nomological. Let us suppose that whales become extinct at some moment. Let us also suppose that, by coincidence, all oceans in which whales have lived since the origin of the species, contain traces of a salt *S* that has neither beneficial nor noxious effects on whales but gets superficially attached to their skin. In this case it is true that

- (9) All whales have traces of salt *S* on their skin.

However, this is a pure coincidence, which is not grounded on any relation between the properties themselves, whose instances are correlated in this



way. Frege's analysis is not sensitive to the difference between the accidental generalization (9) and the preceding nomological generalization (8). According to Frege, the generalization (9) bearing on salt *S* expresses a relation of subordination between concepts just as well as (8), because both bear on concepts, not on the objects in their extension. However, Dretske's analysis allows us to express the difference in the following way. According to their superficial form, both universal statements bear on objects. They state that all objects exemplifying the property of being a whale also exemplify a second property: to be a mammal and to have traces of salt *S* on one's skin. The difference between the two generalizations appears through the possibility to reinterpret (8) but not (9) nomologically: (8) remains true if one considers that the variable linked by the universal quantifier bears on the properties themselves. This change in the domain of the variable can be made explicit with the expression "it is a law that". Thus,

(8') It is a law that all whales are mammals

is true. In this case, there is a relation between the properties referred to, that underlies the coinstantiation of the two properties by all objects that possess the first property. However no such reinterpretation is possible in the case of (9). It is not true that

(9') It is a law that all whales have traces of salt *S* on their skin.

Within the framework of Dretske's analysis, this can be explained by pointing out that there is no relation between the antecedent and consequent properties themselves. It is for this reason that it is impossible to reinterpret (9) so as to give it a nomological meaning: the universal quantifier of the logical form of (9) must be interpreted as ranging over the domain of concrete objects. One makes a mistake if one interprets it as bearing on the *properties* instantiated in these objects, which is the interpretation required by the logical form of (9'). The mistake consists in confusing what happens by coincidence with what happens by virtue of a law of nature. At the level of objects the correlation expressed by (9) between the exemplification of the two properties is universal. But it is not due to the existence of a law of nature linking the second property to the first, subordinating the property of having traces of salt *S* on one's skin to the property of being a whale.

Dretske expresses the result of his analysis in the following way. Instead of saying, like Frege, that law statements express relations of subordination between concepts, Dretske says that they express factual relations between properties. However, Dretske's ontology leads to consequences incompatible with Frege's analysis. In particular, Dretske considers nomological statements, which he calls "law-like statements", as singular statements of fact that express a relation between properties and quantities (Dretske 1977b,

p. 253). Frege insists much on the impossibility of conceiving of the relation between concepts as a relation between objects (cf. Frege 1892a; 1983). In particular, the relation of subordination between two concepts *A* and *B* is not, for Frege, a singular relation between two objects. In Dretske's ontology<sup>21</sup> however, universals are conceived of in such a way that nomological statements making reference to them can be considered as factual and singular. Furthermore, Dretske interprets the relation expressed by:

(P) *F*-ness  $\rightarrow$  *G*-ness

as "an *extensional* relation between *properties* with the terms '*F*-ness' and '*G*-ness' occupying *transparent* positions in (P)" (Dretske 1977b, p. 263; italics Dretske's). In other words, the substitution of predicates making reference to the same property conceived of as a universal, preserves the truth value of a nomological statement. Any term making reference to the same property as "*F*-ness" can be substituted for this expression in (P), without modifying its truth value or its nomological status. "Coextensive terms (terms referring to the same *quantities* and *qualities*) can be freely exchanged for '*F*-ness' and '*G*-ness' in (P) without jeopardizing its true value" (Dretske 1977b, p. 263; italics Dretske's). According to Dretske, laws are factual relations between properties.<sup>22</sup> Rather than referring to the objects in their extension, the predicates embedded in the context of a law make reference to properties understood as universals.

### **Laws as universal relations between property instances**

The logical price the realist has to pay in order to reconcile the requirements that the difference between laws and accidental regularities be objective and that their logical form must be given an extensional analysis, may seem high. In the "platonian" variant of property realism grounding the realist theory of laws of nature defended by Dretske (1977b) and Tooley (1987), part of the price to pay is the necessity to postulate "universals" whose existence is independent of their instantiations. The "Aristotelian" variant defended by Armstrong (1983) avoids such a postulate by supposing that only instantiated universals exist. Armstrong defends a "Principle of Instantiation" according to which "a property must be a property of some real particular; a relation must hold between real particulars" (Armstrong 1983, p. 82). Therefore, neither laws nor the universals these laws link to each other, exist independently of their instantiations. To justify his rejection of non instantiated universals, Armstrong makes clear that "a major reason for accepting the Principle of Instantiation is my desire to uphold, along with Realism about universals, the logically independent doctrine of Naturalism. Naturalism I define as the view that nothing else exists except the single, spatio-temporal, world, the world studied by physics, chemistry, cosmology and so on" (Armstrong 1983, p. 82).

We do not need to follow Dretske with respect to the idea that law statements possess an “extensional” character of a new type, where the relevant extensions do not consist of first-level objects but of universal properties, conceived of as second-level objects. We need not follow him either in saying that law statements are “singular” statements in the sense of expressing a relation between universals that are “second-level objects”.<sup>23</sup> However, I shall stick to what I take to be the crucial idea of Dretske’s analysis: statements of type (L), but not universal statements of type (U), contain opaque contexts. This allows us to provide a clear solution to our problem of explaining the origin of the semantic difference between nomological statements and accidental universal generalizations.

We must reject yet another aspect of Dretske’s analysis, for reasons that will become clear in Chapter 3. Dretske’s analysis implies that a law guarantees not only its universal instantiation, but also the necessary character of each instantiation, where I use “instantiation of a law” to mean the instantiation of the consequent property together with an instantiation of the antecedent property. This consequence is in contradiction with the existence of exceptions. The interpretation of nomological statements according to which they contain an opaque context does not by itself entail that the validity of a law implies the truth of the corresponding universal generalization. It remains possible to accept this interpretation without therefore accepting Dretske’s thesis that every statement of form (L) entails the corresponding statement of form (U). Indeed, accounting for the existence of exceptions to valid laws requires rejecting this thesis.

There is no doubt that it is impossible to derive (L) from (U), and in particular (1a) from (1). In fact, we have seen – with the example of (7) and (7a) – that the truth value of a generalization is not always preserved if one prefixes the expression “it is a law that” to it. However, and at this point the variant developed here departs from other realist theories of laws, and from Dretske’s in particular, the inference from (L) to (U) is not valid either.<sup>24</sup> This entailment would only be valid on the basis of a crucial premise which is in general not true (as we shall see in detail in Chapter 3): to deduce from the law that electrons attract positive charges the universal statement that *all* electrons attract positive charges, it is necessary to make the hypothesis that there are no factors present that interfere so as to prevent the consequent from being true. Such a deduction would be valid only for “strict” or, as Armstrong (1983) calls them, “iron” laws.<sup>25</sup> However, we shall see that most laws of nature do not fall into this category. In the case of our example, the consequent is not true if a screen is placed between the electron and the positive charge; similarly, in atoms containing several electrons, the presence of the internal electrons may neutralize the attraction exerted by the positive charge of the nucleus, so that the outermost electrons may have an exceptional behaviour with respect to the law (1a).

If the circumstances are such that no interfering factor is present, the following deduction is valid. If a particular object *a* satisfies the antecedent

predicate of (1a) because it possesses property *A* (the property of bearing the elementary negative electric charge), one can conclude that the particular statement (11) is true:

(10) *a* bears the elementary negative electric charge.

(1a) It is a law that objects bearing the elementary negative electric charge are attracted by positive electric charges.

Therefore, (11) *a* is attracted by positive electric charges.

According to the particular premise (10), *a* has property *A*; according to the general premise, i.e. the law (1a), property *A* entails property *B*, i.e. the property to be attracted by positive electric charges. This allows us to conclude that, *if the situation is not exceptional*, *a* also possesses property *B*.

It is crucial to note that the law (1a) has consequences for the behaviour objects falling under it, not because it happens to be this or that particular object, but because the object possesses the property denoted by the antecedent predicate of the law. According to the conception defended here, a law applies to a given object not directly, but only through the properties it possesses. In the present case, the law (1a) imposes a constraint on object *a* only by virtue of *one* of its many properties, i.e. by virtue of a constraint imposed on property *A* that happens to be exemplified by *a*.

Anticipating on the result of the analysis of exceptions in Chapter 3, I propose to consider laws of nature as universal correlations between property instances.<sup>26</sup> This conception allows us to account for the specificity of nomic relations in a simple way, and yet to avoid the unwanted consequences of Dretske's position. We have found that (L) is semantically different from (U): it is in general not possible to deduce an explicitly nomological statement from a universal statement of the form (U). We explain this fact in a way similar to Dretske's: it is characteristic for nomic relations to bear directly on properties rather than on the objects possessing these properties. The domain of the variable *x* bound by the universal quantifier in

$$(U) (x) (Ax \rightarrow Bx)$$

are concrete objects. (U) says that, for all objects *x*, if *x* has the property *A* then it also has property *B*.

However, the operator "it is a law that" modifies the domain of the universal quantifier following it. In the explicitly nomological form (L), the function of this operator is to let the variable bound by the universal quantifier range over the domain of property instances. In this way, our account inherits the strength of Dretske's: our account clearly implies that the deduction of (L) from (U) is not logically valid. The change in the domain of the variable *x*, which consists of concrete objects in (U) and of

property instances in (L), is a change in *content*. This explains why the truth value is in general not preserved if one prefixes the operator “it is the law that” before a statement of form (U), to transform it into a statement of form (L). In other words, from the fact that all objects that exemplify property *A* happen to exemplify also property *B*, it does not follow that this universal coexemplification is due to any relation *between the properties A and B themselves*. In the case of *accidental* universal correlations, there is no such relation between the properties themselves.

We can sum up this result by saying that the truth of (L) is not necessary for the truth of (U). So far we agree with Dretske’s analysis. However we disagree with respect to a fundamental point. According to Dretske and other realists of universals, the truth of (L) is not necessary but sufficient for the truth of (U). However, I shall argue that the inference from (L) to (U) is not deductively valid either. The truth of:

(L) It is a law that *As* are *Bs*

is neither necessary nor sufficient for the truth of (U). (L) contains a universal quantification that does not bear on objects but on property instances. However, any object exemplifies, as a particular entity, many other properties, apart from the property *A* in virtue of which it falls under the law (L). (L) is, according to our account, equivalent to:

(LI) Every pure instance of *A* is accompanied by an instance of *B*.

In a regular situation, the validity of (L) guarantees that if an object *a* exemplifies property *A*, then it also exemplifies property *B*. However, our account of laws is designed to explain not only regular cases, i.e. cases in which a law is indeed instantiated, but also the possibility of exceptions to valid laws. It is possible that the antecedent of a valid law is exemplified but not its consequent.

Anticipating on the following chapter, we shall say that some property instances are “impure” with respect to a given law. This is the case if: 1) an object possesses property *A* but in an atypical manner. Such an object is not within the domain of validity of (L). It constitutes an exception to it to the extent that it satisfies the antecedent of the corresponding generalization (U) but not its consequent. 2) An object may possess property *A*, but exemplify at the same time *other properties* that interfere with the effects of the instance of *A* in such a way as to prevent property *B* from being also exemplified. This constitutes the second type of “impure” instantiations, where an object is an exception to (U) although it is subject to the law (L). 3) Finally, suppose object *a* possesses property *A*. Instances of the same property *A* by other objects located nearby can also prevent property *B* from being exemplified. These other objects can therefore make *a* into an exception to (U), although *a* falls under the law (L) by exemplifying *A*. We shall conclude from the analysis of cases of impure instantiations that an

instance of  $A$  is pure *relative to the law* that “all  $A$ s are  $B$ s” if neither the object containing this instance of  $A$  nor any other object nearby contains another property instance that prevents  $B$  from being exemplified.

The inference from the law (L) to the generalization (U) bearing on the *objects* exemplifying the antecedent of the law is only valid if it is *impossible* that any instances of the antecedent be impure. This is the case with “strict laws”: if (L) is strict, then *every* instance of the antecedent  $A$  is pure. However, it turns out that the class of strict laws is much more limited than one might think. A “non-strict” law (L) imposes only a constraint on the (pure or impure) instances of  $A$ , to produce an instance of  $B$ . To clarify the meaning of the word “pure” we need to examine in detail the three types of “impure” property instances that we have just mentioned. This way to proceed is preferable to the solution of substituting for the vague expression “pure” a clause that makes the truth of (U), given the law (L), depend on “the circumstances”. This would not yield any gain in precision. In particular, the following formulation is as vague as ours, but has the defect of lacking any indication of the way in which greater precision can be obtained: The inference from a law to its application in particular circumstances, is valid only *ceteris paribus*, i.e. under the proviso of premises that cannot be explicitly stated.

However, distinguishing pure and impure instances provides a guideline for the specification of the conditions under which a given law  $A \rightarrow B$  applies, i.e. the conditions under which it is possible to conclude from the fact that a given object  $a$  possesses  $A$ , so that  $Aa$  is true, that  $Ba$  is also true. Above all, this formulation has the advantage of helping to explain why a causal relation does not reduce to the instantiation of a law. The terms of causal relations are *particular events*. However, the terms of laws are property instances contained in these events. Now, nothing in the law (L) prevents an object  $a$  exemplifying its antecedent  $A$  from having an exceptional behaviour, in particular by virtue of other properties it exemplifies, in the sense that it does not exemplify the consequent  $B$  of the law. The effect of a law on the exemplification of the properties it relates can be considered as a constraint:<sup>27</sup> if an object is an  $A$ , and there is a law linking  $A$  to  $B$ , then the object is under the constraint of exemplifying  $B$ . However, the force of the constraint imposed by the law  $A \rightarrow B$  is in general (i.e. for non-strict laws) not absolute, so that, in exceptional circumstances, other constraints can overcome it. The possibility of interfering factors is the reason why such a law does not entail that *all* objects exemplifying  $A$  also exemplify  $B$ . This point will be more thoroughly developed in the following chapter.

As for the ontology required to make sense of this conception, the existence of property instances is essential, but not necessarily the existence of universal properties. Armstrong (1992, p. 26f.) mentions the possibility of accounting for laws of nature in terms of property instances, calling them, following Williams (1953), “tropes”.<sup>28</sup> However, Armstrong rejects it without trying to develop it any further, for the reason that it is incapable of

explaining the regularity with which each instance of the antecedent property is accompanied by an instance of the consequent property. His argument is the very argument that realists oppose to theories that reduce laws to universal regularities: they are incapable of *explaining why* these regularities exist, and they do not account for the existence of an absolute difference, in the sense of a difference independent of theoretical context, between nomic and accidental regularities. According to our account, a law is a universal *regularity* between property instances. It refrains from postulating something ontologically more fundamental grounding that regularity. Armstrong (1983) and Dretske (1977b) propose just such a grounding: according to them, it consists in a *singular* relation between universals, in other words in a second-level state of affairs. We can counter Armstrong's argument by showing that his theory, according to which a law is a relation between universals, encounters important problems that ours avoids.<sup>29</sup> Armstrong's theory has the erroneous consequence that every law (L) entails the corresponding generalization (U) bearing on the objects exemplifying the universals related by the law (L). We must of course also show that the theory in terms of property instances succeeds in solving the problems encountered by its rival. This is the topic of Chapters 3 and 4.

What we have just said makes it clear that our analysis of the concept of law forces us to abandon the idea that the relation between a law and each of its instantiations is a kind of *necessitation*. Once we have reached the conclusion that (L) does not guarantee the truth of (U), we must a fortiori reject the idea that a law (L) guarantees that the implication in the generalization (U) has the force of necessity: according to the main defenders of the realist theory of laws of nature, a law linking properties *A* and *B*, lends a necessary status to its instances, i.e. to the coexemplification of properties *A* and *B*.<sup>30</sup> According to these versions of the realist theory, a law of nature expressed by

$$(P) A \rightarrow B,$$

or, in Dretske's terms,

$$A\text{-ness} \rightarrow B\text{-ness},$$

does not only entail the universal generalization

$$(U) (x) (Ax \rightarrow Bx),$$

relating the exemplifications of these properties, but it also entails that each particular implication  $Aa \rightarrow Ba$  has the force of necessity. A law statement does not only say that the properties it refers to are regularly coexemplified, but also that this regularity is the consequence of an underlying *necessity*. However, if there are (possible) exceptions, a law does not even guarantee

the universal coexemplification of the properties it links, and a fortiori does not guarantee that this coexemplification is necessary.

## Conclusion

We have tackled the project of construing the notion of a law of nature on the basis of the hypothesis that the nomological statements expressing such laws have the logical structure of universal generalizations. Our main task has been to explain the origin of the difference between those universal generalizations that are accidentally true and those that are nomological, i.e. express a law of nature. Given that these two types of generalizations have the same *logical form*, the difference in their content must be due to a difference in *interpretation*. Abstracting away from the complex forms of actual scientific laws and concentrating on the analysis of the nomological link in its simplest form, we have taken the logical form of a universal generalization, nomological or not, to be:

$$(U) (x) (Ax \rightarrow Bx).$$

Here is how we have explained the difference between accidental universal generalizations and laws: interpreting (U) as accidental rather than nomological is equivalent to considering that the variable  $x$  ranges over the domain of concrete objects, rather than over the domain of property instances. In other words, if a generalization (U) is interpreted as bearing over concrete objects, it says *of all objects* that if they possess the antecedent property, they also possess the consequent property. However, a statement receives a specifically nomological interpretation, in other words is considered to express a law of nature, if the domain of the quantified variable is considered to consist of *property instances*, in this case of type  $A$  and  $B$ .

We have reached this conclusion on the basis of the following argument, borrowing Dretske's idea for explaining the fact that nomological statements do not allow for substitution of coreferential predicates *salva veritate*, with the help of the Fregean distinction between opaque and transparent contexts. We have begun by noting that accidental universal generalizations, but not law statements, permit the substitution *salva veritate* of predicates by others that make reference to coextensional properties. This can be explained by the hypothesis that these predicates are, in nomological statements, embedded in an *opaque context*, whereas they are embedded in a *transparent context* in accidental generalizations. Therefore we have suggested that the expression "it is a law that" works as an operator creating an opaque context. Nevertheless, our analysis differs both from Dretske's and from Frege's.

According to Frege, all universal generalizations express relations of subordination between concepts. They never bear on the objects belonging to the extension of the concepts but rather on the concepts themselves.



However, it turns out that Frege's analysis does not allow us to do justice to the difference we are focusing on, between accidental and nomological generalizations, for Frege's analysis gives these two types of propositions the very same analysis.

There is a second reason for modifying Frege's analysis of law statements. Frege identifies statements of the form (U) with statements expressing a relation of subordination between concepts. This disqualifies his analysis with respect to laws that allow for exceptions. According to the thesis defended in Chapter 3, most laws of nature have exceptions; this constitutes a reason for denying that their validity entails the truth of universal generalizations of the form (U).

To remedy the first difficulty, we have taken up Dretske's idea according to which nomological statements, but not accidental universal statements, contain an opaque context that modifies the reference of the predicates embedded in it. We part company from Frege by supposing that *accidental* generalizations do not bear on the concepts expressed by their predicates, but rather on the *objects* belonging to their extension. However, within the opaque context specific to nomological statements, the reference of their predicates is modified. The predicates *A* and *B* in the nomological statement (L) refer to the properties themselves.

However, we do not follow Dretske in his idea that law statements possess an "extensional" character of a new type, where the extension does not contain usual first-level objects, but universal properties conceived as second-level objects. This has led us to also reject his thesis according to which nomological statements are "singular" statements expressing a relation between two "second-level objects", i.e. universals.

Rather, we have proposed that the nomological statement

(L) it is a law that  $(x) (Ax \rightarrow Bx)$

has the meaning of a universal generalization bearing on the *property instances* *A* and *B*.

(LI) Every pure instance of *A* is accompanied by an instance of *B*.

This analysis differs from Dretske's in that we give nomological statements a *general* form, whereas Dretske reduces them to singular statements bearing on second-level objects, i.e. the universals *A* and *B*. No one doubts the impossibility of deducing (L) from (U). The truth value of a generalization is not always preserved if one prefixes the expression "it is a law that" to it. However, and this is where we part company from Dretske, our analysis entails that the inference from (L) to (U) is not valid either. This inference is not valid because, as we shall see in more detail in Chapter 3, it lacks a crucial premise. To deduce from the law that electrons are attracted by positive charges, the universal statement that all electrons are attracted by

positive charges, we must suppose that there are no factors that interfere to prevent the consequent from being exemplified. The deduction would go through only if the law in question were “strict”. However, we shall see that this supposition is not justified for most laws of nature.

In a regular, i.e. not exceptional, situation, (L) makes it the case that an object possessing the antecedent property *A*, also possesses the consequent property *B*. However, the main advantage of our conception of laws is that it allows us to explain not only “regular” cases, i.e. cases in which the law is indeed instantiated, but also the possibility of exceptions to valid laws. There are possible circumstances in which the antecedent of a law is exemplified but not the consequent. By making the distinction between regular and exceptional cases of application of a law in terms of pure and impure instances, we have a guideline that we can follow in trying to specify more precisely the conditions under which a law applies to a particular situation. This is the aim of the following chapter.

### 3 Applicability conditions and the concept of “strict law”

#### Introduction

One of the reasons that have led us, in the preceding chapter, to propose an account of the concept of a law of nature opposed to more traditional accounts is the following: we have begun by observing that a law, even if it is valid, may have exceptions. By an exception I mean a situation in which, to take the simplest case of a law whose form is that of an implication between two properties, the antecedent property is exemplified but not the consequent property. In this chapter I shall try to justify the hypothesis already introduced in the last chapter, that the validity of a law does not guarantee the universal exemplification of the consequent in all cases in which the antecedent is exemplified. The validity of a law (L) is not sufficient for the truth of the corresponding universal generalization (U).

I do not pretend to do justice to all extant theoretical alternatives that may be able to explain the possibility of what I call exceptions to valid laws. In particular, I shall leave to one side the “semantic approach” according to which scientific theories do not directly bear on the nomological aspects of reality but must rather be interpreted in terms of their models.<sup>1</sup> This omission can be justified by our aim of providing a realist reconstruction of the concept of causation. For the same reason we shall ignore the “normative approach” to laws, according to which nomological statements do not express propositions, but rather prescriptions or “licences” to form propositions.<sup>2</sup> Insofar as they do not express propositions, they are not capable of being true, whereas the realist conception is based on the idea that a law is what makes a nomological statement true.

Both realist positions we have analysed, Frege’s and Dretske’s, share the idea that (L) is sufficient for (U). However, both disagree with the traditional empiricist thesis that a law of nature (L) is *identical* with a universal regularity (U). The realists’ main argument against this identification has been explored in the last chapter: some universal correlations are not nomic but accidental, which is why (U) does not generally entail (L). In other words, the validity of a law (L) is *unnecessary* for the truth of a universal

generalization (U). However, the exceptional situations that we shall analyse in more detail in what follows, provide another reason for rejecting the identification of laws with universal correlations. They show that the validity of a law (L) is also *insufficient* for the truth of a universal generalization (U). This does of course not prevent the generalization from being *accidentally* true. By contrast, on both Frege's and Dretske's accounts, if it is a law that all *F*s are *G*s, it follows that all objects exemplifying the antecedent *F* also exemplify the consequent *G*. Their failure to account for exceptions to valid laws is our main reason for seeking a better account.

Doubts concerning the idea that the validity of a law is sufficient to guarantee universal exemplification of the consequent with each exemplification of the antecedent, first appeared in the debate on the so-called "deductive-nomological" (D-N) model of scientific explanation.<sup>3</sup> Some explanations that were traditionally considered as paradigmatic deductive explanations, have been shown not to be deductively valid.<sup>4</sup> We shall come back to this debate shortly.

In this situation, some authors<sup>5</sup> have noted a tension between, on the one hand, the conviction that in paradigmatic cases of laws of nature, the consequent is always exemplified with the antecedent and, on the other hand, the observation that many sciences, in particular the so-called special sciences, contain laws that have exceptions. This tension has led them to postulate the existence of two distinct classes of laws, "strict" and "non-strict". The validity of the latter is said to be protected by a so-called "ceteris paribus" clause, according to which a law does not by itself entail a universal generalization; rather, it entails only that the consequent is exemplified if the antecedent is exemplified *and* a certain number of additional conditions are satisfied, which can however not be explicitly stated. Instead of trying to spell out those conditions, one remains satisfied with the somehow mysterious condition that the consequent is exemplified "if all else is equal", which is the meaning of the "ceteris paribus" clause.

The article "Mental Events" (Davidson 1970) in which Donald Davidson develops an original theory of the causal efficacy of mental events,<sup>6</sup> puts forward a trilemma in the form of three theses which are all intuitively plausible but apparently incompatible. One of these theses says that the existence of a causal relation between two events entails the existence of a "strict and deterministic" law: according to Davidson, this follows from the "Principle of the Nomological Character of Causality" (Davidson 1970, p. 208); another thesis says that there is no strict and deterministic law by virtue of which mental events can be predicted and explained.<sup>7</sup> According to Davidson, the reason for which the generalizations bearing on types of mental events are never "strict" is that they are "only roughly true, or they are explicitly stated in probabilistic terms, or they are insulated from counterexample by generous escape clauses" (Davidson

1970, p. 219). Clearly, Davidson considers (at least part of) the generalizations of physics as strict: he says first that “all events are physical” (Davidson 1970, p. 214), which is to say that all events are “describable in a purely physical vocabulary” (Davidson 1970, p. 210), and second that all “events related as cause and effect fall under strict deterministic laws” (Davidson 1970, p. 208). However, Davidson restricts his attention to generalizations formulated in physical vocabulary and generalizations formulated in the vocabulary of intentional psychology; hence it seems clear that he takes the “strict deterministic laws” that underlie causal relations to be physical laws.<sup>8</sup> This leaves open the possibility that physics contains, in addition, non-strict laws that are either probabilistic or contain “generous escape clauses”.<sup>9</sup> Davidson’s distinction between strict and non-strict laws allows him to show that the three theses of the apparent trilemma are after all compatible, a demonstration which is crucial for Davidson’s solution to the mind-body problem, provided by a conception known as “anomalous monism”. Davidson’s distinction between two types of laws (or of universal generalizations, Davidson does not make this distinction) has been very influential and is often accepted without discussion: “strict” laws apply to events under their physical description; their validity does not tolerate any exception. “Non-strict” laws contain “generous escape clauses” that make them compatible with exceptions; these laws apply to events under their mental or psychological description.

Fodor has developed the same distinction with the help of the concepts of strict law and law holding only “*ceteris paribus*”.<sup>10</sup> It is worth noting that this distinction seems to be independent of the important difference between Davidson’s nominalism and Fodor’s property realism. In Fodor’s realist framework, the central idea of the nomological approach to causation can be expressed in the following way: a relation between two events is causal, if and only if, there is a strict law relating one or several properties of these events. Strict laws can be defined by their universal applicability, in the sense that their applicability does not depend on the satisfaction of *ceteris paribus* conditions.

Fodor’s conception of “strict laws” is a notable advance with respect to Davidson’s. Davidson distinguishes between strict and non-strict in a dualist framework, in which there are only two types of predicates, i.e. physical and mental. Davidson uses it mainly as a criterion to ground the distinction between physical laws and the laws of the psychology of propositional attitudes. Fodor puts the distinction in a more general perspective, so that it can ground the distinction between fundamental science and “special sciences”. A science is “special” to the extent that the validity of its laws is limited to a restricted domain of objects and events, whereas fundamental physics is distinguished by its universal domain of application. Everything real falls under the laws of fundamental physics; however, certain objects and events fall, by virtue of their specific properties, also under the laws of

certain special sciences. From this point of view, cognitive psychology with its generalizations implying mental predicates appears as one “special science” among others. The criterion of the size of the domain of applicability of a law provides a firm ground for Davidson’s distinction between strict and non-strict laws.

Fodor argues that these two distinctions, between strict and non-strict laws and between laws having universal and restricted domains of applicability, ground the same dichotomy. On the one hand, the laws of fundamental physics are strict and their domain of applicability is universal; on the other hand, the laws of the special sciences are non-strict and have a limited domain of application. I do not doubt that the classification of sciences according to the criterion of the generality of their domain of application is well grounded. However, I shall offer a general argument according to which the dichotomy between physics on one hand and the special sciences on the other does not coincide with the dichotomy between sciences whose laws are strict and sciences whose laws are applicable only if certain provisos are satisfied. Indeed, most laws of physics are also applicable only under a certain number of conditions. The nature of these conditions comes close to what has traditionally been thought to be expressed by a “*ceteris paribus* clause”. I shall however try to avoid the use of this expression because it may be misleading: *ceteris paribus* clauses are usually understood as characterizing a class of non fundamental or non-strict laws, in particular those of the psychology of propositional attitudes. However as I shall try to show, such conditions limit the application of most laws of physics, too.

Schiffer (1991) who contests the existence of “*ceteris paribus* laws”, in particular in psychology, raises the following challenge for the advocates of this concept: “the first task of someone wishing to maintain that folk psychology has *ceteris paribus* laws is to say what’s meant by *ceteris paribus* sentences. What, exactly, is the nature of the proposition expressed by ‘*Ms* cause *Bs* all other things being equal?’” (Schiffer 1991, p. 2). I shall not try to meet this challenge but rather join Schiffer in the idea that the expression “all other things being equal” does not contribute to the constitution of a complete nomological proposition.<sup>11</sup> My strategy is to reformulate what the so-called “*ceteris paribus* laws” are supposed to be: the *ceteris paribus* clause does not add any content to the law itself, but rather bears on the conditions of application of the law. To say that the law is valid only *ceteris paribus*, or that it applies only under a proviso, has the following meaning: the law is valid, but it does not entail the corresponding universal generalization (U). It follows that deductive arguments in which a law appears as a premise are incomplete. Such an argument becomes valid only if the law (L) is replaced by the corresponding universal generalization (U). However, in a given situation, this replacement is legitimate only if certain applicability conditions are satisfied. The *ceteris paribus* clause makes reference to these applicability conditions.

The challenge we shall try to meet is the thesis according to which these applicability conditions, i.e. the content of the *ceteris paribus* clause, is essentially vague and resists detailed analysis. Anscombe expresses this widespread idea in the following passage:

Suppose we were to call propositions giving the properties of substances “laws of nature”. Then there will be a law of nature running “The flash-point of such a substance is . . . ” [ . . . ]. This law of nature has not the form of generalization running “Always, if a sample of such a substance is raised to such a temperature, it ignites”; nor is it equivalent to such a generalization [ . . . ]. [T]he point here is that “normal conditions” is quite properly a vague notion. That fact makes generalizations running “Always . . . ” merely fraudulent in such cases; it will always be necessary for them to be hedged about with clauses referring to normal conditions; and we may not know in advance whether conditions are normal or not, or what to count as an abnormal condition.

(Anscombe 1971, p. 94)

### Explanatory and predictive deductions as enthymemes

The argument according to which a law (L) does not entail the corresponding universal generalization (U) bearing on the domain of concrete objects can be expressed in several ways. The simplest proceeds by the following *modus tollens*. Against the thesis one intends to establish, one supposes that “the law that all *F*s are *G*s” entails the universal generalization that all objects that are *F* are also *G*:  $(x) (Fx \rightarrow Gx)$ . Then (U) is used in a deductive-nomological argument: One shows that the conclusion of the following deductive inference, in *modus ponens* form, is not in general true: one of its premises is (U), the other is the proposition that a given object *b* possesses the antecedent property *F*; the conclusion is that *b* possesses the consequent property *G*. If the conclusion is false although there can be no doubt about the truth of the particular premise, (U) must be false.

From this result, several lessons can be drawn. Either one considers that the law that *F*s are *G*s is thereby refuted. Or one considers the deduction as an “enthymeme”,<sup>12</sup> in other words a deduction that is not strictly valid because its premises are incomplete. In this case, one is led to abandon both the idea that the content of laws of nature of the form (L) is identical to the content of the corresponding universal generalization (U) and the idea that (L) entails (U). I suggest that this is indeed the correct conclusion to draw. It leads to the thesis that a nomological statement of the form “all *F*s are *G*s” expresses a universal relation between *instances of the properties F* and *G*. All objects exemplifying property *F* are under the *constraint* of the law

that “Fs are Gs”. However, an object can fall under the law relating property  $F$  to property  $G$  and nevertheless be a *counterexample* to the corresponding universal generalization. For the object to fall also under the latter, an additional condition must be satisfied: the situation must not contain any other property instance exercising a counteracting nomological constraint against the exemplification of  $G$ . A law on its own does not guarantee that an object exemplifying its antecedent also exemplifies its consequent. It will do so only if certain conditions of applicability are satisfied.

Let us take a simple example<sup>13</sup> of a law expressing a relation between different properties of an object. Let  $Dx$  be the predicate “ $x$  is a copper wire of diameter  $d$ ”. In realist language, we shall say that  $D$  denotes the *property* of being a copper wire of diameter  $d$ . Let  $Fx$  be the predicate, belonging to an object also possessing  $D$ , “an iron weight of 2 kg is fixed to one extremity of  $x$ ”. Let  $Bx$  be the predicate “ $x$  breaks”. Finally, let  $L(W)$  be the law that all  $Ds$  that are  $Fs$  are  $Bs$ . If  $L(W)$  were equivalent to the universal generalization:

$$U(W) (x) ([Dx \ \& \ Fx] \rightarrow Bx),$$

where  $x$  ranges over the domain of concrete objects, we could construct the following deductive argument (I) concerning a particular copper wire  $c$  of diameter  $d$ , at whose end an iron weight of 2 kg is suspended. All objects having these two properties break. Therefore  $c$  breaks.

$$\begin{array}{l} \text{(I)} \\ Dc \ \& \ Fc \\ (x) ([Dx \ \& \ Fx] \rightarrow Bx) \\ \therefore Bc. \end{array}$$

However, the conclusion of this argument can be false even if the particular premise is true. Therefore the general premise  $U(W)$  must be false. In other words, the following deduction (I\*) is not valid. To be valid it would have to be completed by an indefinite number of premises.<sup>14</sup>

$$\begin{array}{l} \text{(I*)} \\ Dc \ \& \ Fc \\ L(W) (D \ \& \ F) \rightarrow B \\ \therefore Bc. \end{array}$$

More precisely, the substitution of the premise  $L(W)$  in (I\*) for the premise  $U(W)$  in (I) is legitimate only if the law  $L(W)$  is applicable to the particular situation, which requires that its applicability conditions are satisfied. The premises of (I\*) are all true, but (I\*) is an enthymeme. However, (I) is deductively valid, but contains a premise that is according to our thesis false.



In the following situation one of the implicit premises of (I\*) is violated. Let  $Mx$  be the predicate (applied to an object satisfying  $D$ ) “a magnet of magnetisation  $M$  is suspended at  $h$  cm above the lower end of  $x$ ”. Let  $L(M)$  be the law that “the magnetic force that a magnet of strength  $M$  exercises on an iron weight of 2 kg located at a distance  $h$  below it is equal to 2 Newton”.<sup>15</sup>

Always following our hypothesis that a law implies the corresponding universal generalization bearing on the objects exemplifying the antecedent, one can construct the following argument (II). It is deductively valid, just as (I).<sup>16</sup> It contains as a premise the generalization  $U(M)$  corresponding to the law  $L(M)$ .

(II)  
 $Dc \ \& \ Fc \ \& \ Mc$   
 $U(W) \ \& \ U(M)$   
 $\therefore \neg Bc$ .

In this situation, in which an iron weight of 2 kg is attached to the copper wire  $c$ , but in which a magnet located above this weight cancels the gravitational force it exerts on  $c$ , the wire does not break. However, if both arguments (I) and (II) were valid, we would end up with a contradiction, because the premises of argument (II) entail the premises of argument (I). We could therefore deduce from the premises of (II), both  $Bc$  and  $\neg Bc$ .

It is quite easy to find circumstances in which the conclusion of a deduction following schema (I) is false, in the sense that it is contradicted by the conclusion of a deduction following schema (II), which takes the circumstances into account. In each case, a deduction with the contradictory conclusion can be constructed. Take the law linking the period of a classical pendulum to its length.

$L(CP)$  The period  $P$  of a classical pendulum is  $T = 2\pi\sqrt{L/g}$ ,

where  $L$  is the length of the pendulum and  $g$  the acceleration due to the gravitational field of the Earth.

Let  $CPx$  and  $PEx$  be defined in the following way:

$CPx$ : “ $x$  is a classical pendulum of length  $L$ ”

$PEx$ : “ $x$  has a period of  $P = 2\pi\sqrt{L/g}$ ”.

The generalization corresponding to the law  $L(CP)$  is:

$U(CP)$ :  $(x) (CPx \rightarrow PEx)$ .

However, once more the conclusion of the deductive argument (III):

(III)  
 CPb  
 U(CP)  
 ∴ PEb

is false in each of the following circumstances: 1) the pendulum is not in a vacuum, so that friction forces act on it; 2) the pendulum is made of iron and a magnet is nearby; 3) the pendulum swings at a large amplitude. In circumstances of these types, some property of the pendulum other than its length and its being massive interferes with the constraint imposed on its movement by the interaction of its mass with the gravitational field of the Earth, given its form and length. In situations of type (1) its macroscopic surface, which is not taken into account by the law L(CP), interacts by virtue of a law determining the forces of friction with the surrounding air. In situations of type (2), the ferromagnetic properties of the pendulum give rise, by virtue of the external magnetic field, to a force superimposed on the force exercised by the gravitational field of the Earth. The pendulum law L(CP) imposes a constraint on the movement of the pendulum only by virtue of two of its properties, the property of being massive and of having a certain length, without taking into account any other of its properties, in particular the properties of being made of ferromagnetic metal and of having a macroscopic surface. Finally, the law L(CP) determines the movement of a pendulum only to the extent that the amplitude of its swinging is small. If this amplitude becomes larger, in situations of type (3), components of the gravitational force ignored by L(CP) are not negligible any more. In each case, the situation is such that the pendulum does not exemplify the consequent of the law L(CP) although it exemplifies its antecedent. In each case, a constraint exercised on the pendulum and created by some of its properties different from its mass and length, interferes with the constraint exercised by its mass and length by virtue of the law L(CP), with the result that the effective period of the pendulum in this particular situation is not the one imposed by the consequent of L(CP). Therefore, in these three kinds of circumstances, it is possible to construct an argument of the same type as (II) leading to the conclusion that the pendulum does not satisfy property PE. This conclusion contradicts the conclusion of argument (III), just as the conclusion of argument (II) contradicted the conclusion of (I), because the period of the pendulum is not in the proportion to its length indicated by PE.

We can draw the following lesson from this observation. Let there be a law of type L(CP) and a particular object *c*, and let us ask *whether* it satisfies predicate PE (prediction) or *why* it satisfies predicate PE (explanation). A deduction of type (I\*) (deductive-nomological explanation or prediction) is valid

only if certain applicability conditions are satisfied. Let us call them “the applicability conditions of the law  $L(CP)$  to the object  $c$ ”. The conclusion that  $PEc$  can be validly drawn only in situations in which these conditions are satisfied. In other words, the universal generalization  $U(CP)$  that figures as a premise in argument (III) is false: the regularity it expresses is not strictly universal but holds only under the condition that no interfering factor is present.

We have seen in two paradigmatic cases why explanations of deductive nomological form using a law ( $L$ ), as in schema ( $I^*$ ), are enthymemes in the sense of having incomplete premises, so that the premises that are explicitly given do not guarantee the truth of the conclusion. This fact is at the origin of a fundamental problem: what is lacking to establish the truth of the conclusion, the satisfaction of the applicability conditions, cannot be expressed both generally and explicitly, in the form of supplementary premises. In such a deductive-nomological explanation, the applicability of the laws it contains as premises, to the particular case one is dealing with, requires the satisfaction of certain conditions. Let us look more closely to these applicability conditions.

There are several ways of classifying such applicability conditions. One is to distinguish two groups, one of which contains two subgroups:<sup>17</sup>

1) The first group of conditions restricts the attribution of the antecedent predicate to particular objects. In realist language, one can speak of the conditions under which the object exemplifies the property expressed by the antecedent. This type of condition is analysed in the next section.

2) The second group consists of conditions under which the object satisfies the consequent, given the validity of the law and the object’s satisfying the antecedent. Within this category we may distinguish two kinds of conditions:

2a) The first kind is relevant if one applies a law relating properties that a first object  $b_1$  possesses by virtue of its relation to a second object  $b_2$  (for example laws of attraction and repulsion). Such laws have a “binary” form: they express the influence one object  $b_2$  exercises on a given object  $b_1$ . The deduction made on the basis of the corresponding universal generalization may have a false conclusion if there is a third object  $b_3$  also acting on  $b_1$  by virtue of the same type of relation, which constitutes the antecedent of the law.

2b) The second subgroup of conditions concerns the absence of other properties possessed by the object under consideration, which prevent it from having the property expressed by the consequent of the law. We have already seen cases where this kind of condition violated, in

the examples of the copper wire and of the pendulum. We shall return to the second group of conditions on pages 114–17.

### **Applicability of the antecedent**

One type of applicability conditions is common to mathematical theorems and law statements<sup>18</sup> of empirical science, in particular physics. A mathematical theorem typically contains a clause specifying the domain of objects to which it is applicable. For example, the Bolzano-Weierstrass theorem in analysis says that “every bounded infinite set of real numbers has an accumulation point”.<sup>19</sup> The domain of objects to which this theorem applies is constituted by infinite limited sets of real numbers.

Nomological statements of empirical science also contain sometimes a clause specifying explicitly their domain of validity. Let us take as an example the Dulong-Petit law of solid state physics:<sup>20</sup>

L(DP) All solid insulators possess the same specific heat  $c_v=3nR$ , and this specific heat has a constant value independent of the object’s temperature, where  $n$  is the quantity of the molecular constituents of the object, in moles, and  $R$  is the ideal gas constant.

The formal analogy with the mathematical theorem of Bolzano-Weierstrass mentioned before is obvious. Just as the latter, the statement of the Dulong-Petit law contains a clause specifying the domain of objects for which it is valid: solid insulators. However, there is a fundamental difference between mathematical theorems and physical laws: the conditions of validity of mathematical theorems can be explicitly spelled out. However, the statement of physical laws does not in general contain a specification of sufficient conditions under which the law is applicable to a given particular object. For example, the Dulong-Petit only holds if the temperature of the particular solid object lies in an interval specific for each type of solid. At high temperatures, the so-called harmonic approximation implicit in the derivation of the law is no longer valid. To obtain a better approximation of the observed values of specific heat, it is then necessary to consider so-called non-harmonic terms. At very low temperatures, it is necessary to use a completely different theory, i.e. quantum mechanics, in which the vibrations of the solid (on which the specific heat depends) are considered as a boson “gas”. Calculating the specific heat within this theory, one finds that the specific heat is a function of the cube of temperature, instead of being a constant function independent of temperature, as is the case at moderately high temperatures according to the Dulong-Petit law.

What is important for our analysis is that there is no unique general formula indicating the specific heat of an arbitrary solid object, applicable to

all objects under all conditions. There is no general criterion independent of the particular type of solid, which would indicate which of the three formulas corresponding to the three ranges of temperature, applies at a given temperature to an arbitrary solid object. In particular, there is no general criterion specifying explicitly the conditions under which an arbitrary isolating solid instantiates the law of Dulong-Petit.

However, within the framework of a realist conception of laws, the fact that it is impossible to find such a general criterion of applicability is comtible with the possibility that there is a well-determined class of property instances that objectively constitute the antecedent of the law. The limitation of applicability conditions we are presently concerned with, concerns *explicit statements* of laws of nature. The antecedent predicate of the law statement must be applicable to the objects whose behaviour the law is supposed to explain or predict. However, the fact that a given object satisfies the predicate and nevertheless does not behave in accordance with the law statement, appears as an exception to the law only within a nominalist conception of laws. The realist conception distinguishes between the antecedent *property* of the law and the antecedent *predicate* of the statement expressing the law: in this framework the apparently exceptional behaviour of certain particular objects can be explained by the inappropriateness of the predicate chosen to denote the antecedent property of the law. In such an apparently exceptional case, the object does not exemplify the antecedent property of the law, although the antecedent predicate of the law statement applies to it. In short, the law statement is an inadequate expression of the law.

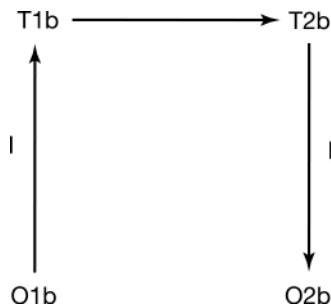
If we choose to describe the situation in a realist framework, an object may have an exceptional behaviour with respect to a given law although it satisfies the antecedent predicate of its explicit formulation. Let us consider the following example. To predict the period  $P$  of a pendulum  $b$  on the basis of the law  $L(CP)$ ,  $b$  must satisfy the antecedent predicate of  $L(CP)$ , which relates the period of a classical pendulum to its length. However, this is a theoretical predicate whose applicability depends on the satisfaction of a set of criteria that do not reduce to the object's possession of immediately observable properties. The nature of these criteria is determined by theory. In the case of the pendulum  $b$ , for the predicate  $CP$  to be applicable to  $b$ , so that  $CPb$  is true, the amplitude of the movement of pendulum must be small (the sine of the maximum angle reached by its swinging movement must be approximately equal to that angle itself), and it must be located in a constant and uniform gravitational field acting in the direction perpendicular to its movement. In a realist framework, we can say that situations of type (1) mentioned in the last section do not constitute exceptions to the law itself, but only to approximate expressions of this law, in which its antecedent property is designated in an inadequate way. Pendulums whose amplitude is too large do not fall under the law  $L(CP)$  and cannot therefore constitute exceptions to it.

Within Hempel's (1988) nominalist framework, the situation looks different: it does not distinguish the law from its statement, nor the properties of the object from the predicates it satisfies. Therefore the objects that are the focus of explanation or prediction are considered exclusively under certain descriptions, insofar as they satisfy certain predicates. In the case of apparently exceptional behaviour, instead of saying that the predicates of the law statement do not adequately express the properties contained in the law itself, Hempel says that the objects are not directly known by the predicates of the theory used in the explanation or prediction but, in the simplest case, by observational predicates not belonging to that theory.

Let us illustrate this framework with Hempel's own example. There is a law of solid state physics according to which:

L(H) Parts of a magnet are themselves magnets, with the same vector magnetization per unit of volume.

This law allows us to explain the following phenomenon. Consider a uniformly magnetized metal bar. If it is cut in two halves and the halves are freely suspended in air, they will orient themselves so as to fall into a straight line.<sup>21</sup> The antecedent of L(H) contains the predicate "magnet". The question we are presently concerned with is whether this predicate refers precisely to the property contained in the law. The law is valid for all magnets, where "magnet" expresses a *theoretical* concept of solid state physics. In a concrete case, the theoretical concept is judged to be applicable on the basis of a certain number of *observable* phenomena characteristic of magnets. Suppose we have observed that a given object attracts iron filings. In Hempel's terms, this makes legitimate a "theoretical ascent" from the fact that the observational predicate<sup>22</sup>  $O_1$  applies to the bar  $b$ , to the fact that the theoretical predicate  $T_1$  applies to  $b$ , in short from the statement  $O_1b$  to the statement  $T_1b$ . Anticipating a little, the structure of the explanation of an observable phenomenon with the help of a theoretical law is the following:



where  $O_1b$  is the statement:

$O_1b$ : “ $b$  is a metal bar that attracts iron filings”.

In virtue of “operational criteria” collected in set I, one can infer that:

$T_1b$ : “ $b$  is a magnet”.

Next, calling upon one or several laws of solid state physics including L(H), one can infer that:

$T_2b$ : “if  $b$  is cut in two pieces  $b_1$  and  $b_2$ , these two pieces are magnets and their poles attract or repel each other”.

Finally, and once again according to the operational criteria in I, which relate the vocabulary of the theory of solid objects to a vocabulary understood before the introduction of the theory (for simplicity called “observational” vocabulary), one can derive the statement:

$O_2b$ : “if  $b$  is cut in two shorter bars  $b_1$  and  $b_2$  and these are suspended, by long thin threads, close to each other at the same distance from the ground, they will orient themselves so as to fall into a straight line”.

The step we are particularly concerned with here is the first inference, from  $O_1b$  to  $T_1b$ . In Hempel’s terms, the transition from a statement in “observational” language to the statement in theoretical language, requires an “inductive or theoretical ascent” (Hempel 1988, p. 21). Hempel notes that the theory of magnetism does not contain any principle that would allow the deduction that an object that attracts iron filings is a magnet, just as it does not contain any principle allowing the possibility to draw the converse deduction that all magnets attract iron filings. The theory applies to objects on the condition that they are magnets. It does not provide any general principle by which one could judge whether a given object belongs to the proper domain of the theory, on the basis of statements in an extra-theoretical language, in particular an observational language. There is no simple equivalence between objects satisfying phenomenological criteria such as the fact that they attract iron filings, and magnets. The inference from  $O_1b$  to  $T_1b$  is not valid in general. One case in which its conclusion is false although its premise is true is when the observed object contains only a small magnetic nucleus, without therefore *being* a homogenous magnet. Another such case is when a magnet is hidden under or behind the observed object, in such a way that the observable situation is just as if the object itself were a magnet. Yet another type of case is when the observed object attracts iron filings for a different reason, for example because it is electrostatically charged.

Within the realist conception of laws defended here, cases of these types do not constitute exceptions to the law L(H): this law relates instances of the property of being a magnet to instances of properties of separate pieces of a magnet. What those cases show is that there are objects that are not really magnets although their observable behaviour shares important characteristics with magnets, or that there are objects that are not magnets in the sense required for the application of the theoretical law at issue. In one case, an object *contains* a magnet, without therefore *being* a magnet; in another case, an electrostatically charged object “simulates” the behaviour of a magnet, without being a magnet. It turns out that the observational criteria employed are too broad to distinguish between these possibilities. For this reason, an object satisfying the predicate “is a magnet” – where this predicate is applied on the basis of observational criteria – may seem to behave in a way that is exceptional with respect to the law. But this is only because its explicit statement contains the predicate “is a magnet”, which does not refer precisely to the antecedent (theoretical) property of the law, to the extent that its applicability conditions are at least partially observational. In a realist conception, this does not show that the situation is an exception to the law itself.

As a second example, consider the application of Newton’s law determining the force  $F$  acting on two bodies of mass  $m_1$  and  $m_2$ , as a function of their mutual distance  $d$ :

$$L(GR)F = \frac{g \cdot m_1 \cdot m_2}{d^2}$$

where  $g$  is the gravitational constant.

In order to apply this law in a paradigmatic situation – the movement of celestial bodies – it is necessary to switch from an observational description to a theoretical description. The conceptual distance that has to be bridged separates an extended body with an irregular shape and composition from a massive *point* without extension that has no intrinsic properties other than being massive. In order to be able to apply the law to particular objects whose evolution is to be explained or predicted, these objects must be *considered* as point masses. It is only on this condition that the notion of the distance between two bodies is well determined, and only on this condition that the law allows us to calculate the attractive force and indirectly the bodies’ movements. There is no general method for applying theoretical concepts to concrete objects. However, this is an epistemic difficulty. To find out whether a law linking theoretical properties applies to a given concrete situation, one must begin by inferring the possession of theoretical properties on the basis of the possession of properties that are accessible to observation. To apply a theoretical law to a given situation, a step of “theoretical ascent” is indispensable.

The problem of justifying such a theoretical ascent in a given case persists even if the statement of the law contains a clause explicitly specifying its



domain of validity. The function of such a clause is to indicate one condition limiting the applicability of the law. However, in the case of empirical laws bearing on concrete objects, no such clause can spare us the necessity of theoretical ascent. In the case of a mathematical theorem, there is no gap to bridge between the predicate specifying the domain of validity of the law and the antecedent property. In the case of the theorem of Bolzano-Weierstrass, the predicate “is an infinite bounded set of real numbers” is both the predicate determining its domain of validity *and* determines exactly the antecedent theoretical property. Therefore there is no logical space in which the epistemic problem we are faced with could arise: the problem of knowing whether a given object that satisfies the antecedent predicate of the explicit *statement* of the law, also possesses the real antecedent property of the law, given that the applicability conditions of the predicate determine that property only imperfectly.

In the case of empirical laws, we can in general not be certain that there is a predicate referring precisely to the antecedent property of the law. Let us take an example already mentioned: the Dulong-Petit law bearing on the specific heat of insulating solids. The predicate referring to the antecedent property of the law, “being an insulating solid” does not specify that property exhaustively: this is equivalent to saying that it is not sufficient for an object to satisfy this predicate in order to fall under the law. The law applies to insulating solids only under the condition that a certain number of *other* requirements are also fulfilled. It applies in particular only within a certain interval of temperatures, specific for each type of solids. The fact that each type of solids has its own interval in which the law applies is the reason for which the antecedent property of the law cannot be explicitly specified in complete generality. The discrepancy between the explicit statement of the law and the law itself is therefore inevitable.

This discrepancy is responsible for the apparently exceptional behaviour of certain objects. Suppose that the solid object *o* possesses at temperature *T* a specific heat differing from the value predicted by the law of Dulong-Petit although at that temperature, “ordinary” solids behave in perfect accordance with that law. Given the particular structure of *o*, temperature *T* is too high *for it* to justify the harmonic approximation. The object *o* appears as an exception to the law because its explicit statement does not express the law with sufficient precision. Object *o* satisfies the antecedent predicate “is an insulating solid” of this imperfect statement, although it does not fall under the law the statement is supposed to express. However, in our conception *o* does not appear as an exception, because it does not fall under the law itself. This object does not contain an instance of the property contained in the law. If we tried to get rid of this difficulty by designating the antecedent property of the law by the predicate “is an insulating solid in the sense of the Dulong-Petit law”, we would be caught in a circle, for this predicate does not provide an *independent* specification of the antecedent property, given that it makes explicit reference to the Dulong-Petit law.

By contrast, within the framework of Hempel's conception, an object such as *o* appears as an exception to the Dulong-Petit law, because it satisfies the "observational" predicate  $O_1$  but not the consequent predicate  $O_2$ , "has a specific heat of  $3nR$ ". The exceptional behaviour of *o* can be explained with the help of the concept of theoretical ascent. In *o*'s case, the theoretical ascent from  $O_1$  to the theoretical predicate  $T_1$  is not justified. Within this conception, it is difficult to characterize in a satisfying way the nature of the problematic inference from  $O_1$  to  $T_1$ . In the context of our analysis, it is enough to know that it is not a deduction. However, this does not justify Hempel's suggestion that theoretical ascent is a form of *inductive* inference.<sup>23</sup>

Whatever may be the exact nature of Hempel's theoretical ascent, the results he obtains are unsatisfactory: in Hempel's framework, the behaviour of solid *o* appears as exceptional. However it is hardly compatible with the concept of an exception that all solids of the same type as *o* have exactly the same behaviour as *o*.<sup>24</sup> A "systematic exception" is paradoxical: to the extent that the behaviour at issue is in fact part of a regularity, it is not strictly speaking exceptional. This suggests an argument in favour of a realist conception of laws and against nominalist theories taking only explicit law statements into account. Hempel's framework, which belongs to the latter category, makes it inevitable to conclude that solids of the type of *o* constitute an exception to the Dulong-Petit law, insofar as this law is identified with the generalization  $O_1 \rightarrow O_2$  in terms of observable predicates. In a nominalist framework, it is difficult to consider the law to be the theoretical generalization  $T_1 \rightarrow T_2$ , because it is not possible to indicate in any independent way the extension of the predicate  $T_1$ ; there is no general criterion determining which objects satisfy this predicate. The only possible "criterion" is tautological, to the extent that it mentions the law  $T_1 \rightarrow T_2$  itself: according to this pseudo-criterion, the objects falling under the concept expressed by  $T_1$  are those that fall under the law  $T_1 \rightarrow T_2$ . Therefore, it seems that only a realist position allows us to hold that the behaviour of solids of the type of *o* is regular although it appears as exceptional with respect to every explicit law statement. The realist position allows us to say that objects of the type of *o* obey to a law that is not exactly, but at best approximately, expressed by any explicit law statement. Their behaviour is not exceptional in reality; it is apparently exceptional because the explicit statement of the law is only imperfectly adequate to the objective law.

### **The deduction of the consequent**

Some restrictions to the application of laws of nature are independent of the problem we have just discussed, of the inadequacy in principle of the predicates figuring in explicit law statements, with respect to the antecedent properties involved in the real laws. We have already illustrated some such conditions of applicability with the examples of the copper wire and the pendulum.

Even if a law is applicable to a given object in the sense that this object possesses the antecedent property of the law, it is nevertheless possible that the object does not possess its consequent property. In that case we are faced with a genuine exception. This is different from situations of the type analysed in the preceding section where an object appears to have an exceptional behaviour because it satisfies the antecedent predicate of an explicit law statement without satisfying its consequent predicate. Now we suppose that the object really contains an instance of the property concerned in the law itself. If the object does nevertheless not possess the consequent property, we must conclude that the law does not by itself guarantee that every object possessing its antecedent property also possesses its consequent property.

We can once again use the example of the magnet to illustrate the possibility that an object has an exceptional behaviour with respect to a law although it falls in its domain of validity, i.e. although it really possesses the antecedent property of the law: as a concrete object, a magnet does not only possess the property of being a magnet but also the properties of being a massive and extended body. It is therefore subject to laws bearing on friction with air. In the situation described in the preceding section, we need only expose the two pieces of the magnet to a sufficiently strong air current to prevent them from falling into a straight line parallel to the direction of the vector of magnetization: in such conditions they may well orient themselves parallel to the direction of the air current.

In such a case we shall say that the law at issue does not apply to the object in that particular situation. We can distinguish two sorts of reasons for which a law may not apply to a given particular case: one type of exceptions has its origin in the fact that every concrete object or event possesses many properties, although any particular law governs its evolution and causal interactions, only by virtue of *one or a small subset* of its properties. Other laws bearing on other properties of the object also constrain its evolution and interactions.

Another limitation of applicability has its origin in the fact that even with respect to just one property of a given object, one application of a law (expressed in a form that has variables for only two objects, as in the form  $L(GR)$  of the law of gravitation mentioned above) determines the evolution of the object only with respect to one other object. Insofar as the relevant law for magnetized solids  $L(H)$  takes into account only the interaction between the magnetic fields exerted by the two halves of the magnet on each other, the following case is exceptional with respect to the law  $L(H)$ : in the presence of an external magnetic field having its source outside the two bodies considered in this first application of the law, the magnets do not align according to the constraint imposed by the law although they are magnets (first condition) and although their other properties do not exercise any causal influence on their movements and their mutual position (second condition). In such a case, it may be necessary to apply the same law repeatedly, in order to be able to explain or predict the behaviour of the two halves of the magnet.

What we have called “conditions of applicability” of a law correspond to what Hempel (1988) calls “provisos”. However, Hempel’s model suggests a distinction in subcategories of provisos that is not equivalent to our own. Furthermore, Hempel’s classification is not well grounded. Consider once again the steps that are, according to Hempel, necessary to apply a theoretical law in order to predict or explain a phenomenon:

If all three inferential steps were deductively valid we could say, within the framework of Hempel’s model, that the theoretical generalization:

$$(x) (T_1x \rightarrow T_2x).$$

deductively entails the following generalization in “observational” terms:

$$(x) (O_1x \rightarrow O_2x).$$

It would be true that:

$$(x) [(T_1x \rightarrow T_2x) \rightarrow (O_1x \rightarrow O_2x)].^{25}$$

There would be no exceptions in the sense that the validity of the theoretical law alone would guarantee that all objects falling in the domain of the universal generalization  $O_1 \rightarrow O_2$ , possess, without any exceptions, the consequent property  $O_2$  if they possess the antecedent property  $O_1$ .

The fact that the first inference  $O_1 \rightarrow T_1$  is not deductively valid constitutes, as we have seen, the first source of “exceptions” to the law  $T_1 \rightarrow T_2$  (more precisely, they are exceptions to the universal generalization  $(x) (O_1x \rightarrow O_2x)$ ). Now Hempel holds that the two other inferences are not deductively valid in any strict sense either. For each of them, he thinks there are exceptional situations in which the conclusion is false although the premise is true. However, as far as I can see, in the case of the inference  $T_1 \rightarrow T_2$ , there are no reasons for thinking that it rests on implicit premises which, if they are not satisfied, may make its conclusion false.

Hempel offers two examples of violations of implicit premises underlying the inference from  $T_1$  to  $T_2$ . First, if the magnet is broken into pieces at a high temperature, the pieces may lose their magnetization so that their behaviour will no longer be in accordance with the law L(H).<sup>26</sup> Second, if one first infers the gravitational attraction that two objects exercise on each other, from their masses and the distance between them, and then infers from this attraction, the acceleration they undergo, the second step of the inference is grounded on the implicit premise that no other forces act on the bodies.<sup>27</sup>

In the first example, here relevant for the application of the law, its antecedent  $T_1$  is not satisfied any more. In other words, this situation appears as a source of exceptions in Hempel’s model, but not in a realist framework: the inference  $O_1 \rightarrow T_1$  is not justified in that case because the heated magnet does not satisfy the antecedent  $T_1$  of the theoretical law. We should simply say

that, as soon as it has been heated, the object no longer falls in the domain of validity of the law, and no longer possesses its antecedent property.

In the second example, in which other forces act on two bodies in gravitational interaction, the invalid inferential step is not, as Hempel holds, intra-theoretical. Rather, it is an inferential step of the type  $T_2 \rightarrow O_2$ . Both the gravitational attraction exercised by  $m_2$  on  $m_1$  and the acceleration caused by this attraction are really those predicted by the laws L(GR) and

$$L(\text{Newton}) F = M \times A.$$

However, if the real acceleration  $A(m_1)$  is not identical to the acceleration  $A$  resulting from the inferences  $O_1 \rightarrow T_1$  and  $T_1 \rightarrow T_2$ , this is not because the conclusion  $T_2$  of the theoretical inference  $T_1 \rightarrow T_2$  is false. Rather, we must distinguish the real acceleration  $A(m_1)$  undergone by the concrete object  $m_1$  from the acceleration due to its attraction by  $m_2$ . The acceleration  $A(m_1)$  is jointly produced by the attraction due to the presence of  $m_2$  and by all other forces to which  $m_1$  is exposed.<sup>28</sup> This means that the conditions of applicability that are violated in this case, correspond, in Hempel's model, to implicit premises of the deduction  $T_2 \rightarrow O_2$ , and not to implicit premises of the theoretical inference  $T_1 \rightarrow T_2$ .

### **Are its conditions of applicability part of a law of nature?**

It may be objected that we have privileged a particular presentation of the realist position, which is responsible for our result that a law (L) is not equivalent to the universal generalization bearing on the objects possessing the properties related by the law. We have so far presupposed without any justification that the conditions of applicability are not part of the law itself because it is in principle impossible to spell them out explicitly. For a given law, it is impossible to draw an exhaustive list of conditions that must be fulfilled in order for the law to apply to a particular object in a given situation, in other words, in order for the law to guarantee that, if the object exemplifies the antecedent property, it also exemplifies the consequent property.

It is not only impossible to state explicitly those conditions of applicability, but the premises that could, for a given law, guarantee the validity of the problematic deductive inferences we have examined above, do not constitute a well-determined set. This is important for the evaluation of a thesis that we shall consider in a moment, which aims at vindicating the dichotomy of strict and non-strict laws. Our analysis implies that laws of nature are in general non-strict, in the sense that not all objects possess the consequent property of a given law even if they possess its antecedent property and therefore lie in its domain of validity. Armstrong (1983) also insists, in a way characteristic for nomological realism, on the importance of distinguishing clearly between laws and law statements. For realists, the objective

validity of a law is independent of its explicit statement. Armstrong makes use of this distinction when he proposes that its conditions of applicability are part of a law although they are not part of the nomological statement explicitly expressing it. The explicit statement of a law can contain only a finite number of clauses; however, it is possible to consider, with Armstrong, the possibility that this not does not prevent *the law itself* from containing an infinite number of them. This implies that the law itself is strictly speaking ineffable. “Here are possible laws which, if they hold, can never actually be stated in full” (Armstrong 1983, p. 28).

Let us suppose, with Armstrong, that there are laws that apply always and everywhere, for which the conditions of applicability of our second type are always satisfied, so that no property interferes with the antecedent property of the law. In the case of laws of this type, which Armstrong calls “iron laws” (Armstrong 1983, p. 147), it is guaranteed that every object satisfying the antecedent of (U) satisfies also its consequent, independently of its other properties and of the circumstances. In the case of such laws, the inference from the law statement (L) to the corresponding universal generalization (U) is valid. However, the laws in the category Armstrong calls “oaken laws” (Armstrong 1983, p. 148) have the now familiar property of having exceptions. For an oaken law, there are in fact circumstances in which the antecedent is exemplified but not the consequent.

However, Armstrong’s way to draw this distinction appears to be ad hoc. In the form in which Armstrong introduces it, the distinction is not based on any properties of the laws themselves: Armstrong’s criterion for establishing it is the accidental fact that, until now – as one is tempted to add – no exception has yet appeared. In the absence of any reason for which these laws might, by themselves, prevent the presence of any interfering factors, Armstrong’s distinction does not appear to bear on the laws themselves, but rather on an accidental property of their instantiations.

There is a thesis of Armstrong’s that may seem to provide the means for countering this objection: he holds that oaken laws, though not their explicit statement, *contain* an indication of the exceptional circumstances in which they are not applicable. According to this thesis, an oaken law has the following form:

“It is a law that *Fs* are *Gs*, except where *Fs* are *Hs*, are *Js*, and *Ks* . . . and so on for an *infinite* set of distinct properties” (Armstrong 1983, p. 28, italics Armstrong’s).

Here the distinction between the law taken in the realist sense and the statement expressing it is fundamental. If a law is oaken, then according to Armstrong, it contains an exhaustive list of the conditions necessary for its applicability, something that would be impossible for its explicit statement. However, this thesis is problematic insofar as it does away with what it was meant to explain: the origin of exceptions. Conceived in Armstrong’s way, oaken laws turn out in the end to be iron laws. A particular situation only satisfies the antecedent of an oaken law if no potential interfering factors are

present in it. The antecedent is only satisfied to the extent that the situation *is not* exceptional. In light of this, it appears that Armstrong's conception of the laws of nature has some undesirable consequences. First, it eliminates the difference between strict laws and laws that may be applied only under certain conditions. Armstrong himself is forced to admit that the distinction between strict laws (or iron laws) and non-strict laws (or oaken laws) is not based on a conceptual criterion drawn from distinctive properties of the laws themselves. Rather, it turns out to be a distinction in terms of contingent circumstances: between laws for which exceptional circumstances happen to exist and laws for which exceptional circumstances happen not to exist actually, although they are possible. However, this accidental difference hardly seems capable of grounding a distinction between two fundamental kinds of laws.

Furthermore, there are two general reasons for doubting that iron laws (or oaken laws in the logical form proposed by Armstrong) are necessarily strict in the sense explained above. First, we have said that an iron law is a law for which there does not actually exist any property that interferes with its antecedent property. This excludes exceptions due to a violation of conditions of the first type we have distinguished above. However, it seems impossible to exclude violations of the second type of conditions: it always seems to be possible that an object behaves exceptionally because of the interference of other objects that exemplify the very antecedent property of the law. This can happen in the situation of the example analysed in the preceding sections: the two halves of the cut magnet do not line up because there is another magnet whose field is superposed on the fields of the two observed halves. The law does not apply to those two parts, in the sense that their behaviour is not that dictated by the consequent of the law. To predict or explain their behaviour it is necessary to take into account another object; but this other object interferes by virtue of the same property that is responsible for the regular behaviour of magnets. Therefore it seems that there are always exceptions of this kind; this is enough to show that there is no particular class of "iron laws".

Second, by trying to include the whole set of possible circumstances that may produce exceptions, Armstrong seems to be trapped on a "slippery slope". Canfield and Lehrer (1961, p. 207f.) propose an argument to this effect. Let us return to their example of the iron weight suspended on a copper wire. The initial deduction we considered above has the form

- (I)  
 $Dc \ \& \ Fc$   
 $U(W) \ (x) \ ([Dx \ \& \ Fx] \ \rightarrow \ Bx)$   
 $\therefore Bc.$

However, to show that the use of the general premise  $U(W)$  of (I) is justified, it is necessary to examine whether there are conditions of the type expressed by the predicate  $Mx$  ( $Mx$  was the predicate – applied to an object  $x$  satisfying  $Dx$  – “a magnet of magnetisation  $M$  is suspended at  $h$  cm above the lower end of

$x$ ”), in other words, whether there are properties whose exemplification by  $c$  is both compatible with the exemplification of  $D$  and  $F$  and incompatible with the exemplification of  $B$ . The answer to this question is affirmative if there is some other law  $L(M)$ , such that the corresponding universal generalization  $U(M)$  allows for an argument with the conclusion  $\rightarrow Bc$ , on the basis of the very premises of (I) whose conclusion is  $Bc$ . Such an argument has the form:

(II)  
 $Dc \ \& \ Fc \ \& \ Mc$   
 $U(W) \ \& \ U(M)$   
 $\therefore Bc$ .

However, there now arises, with respect to the law  $L(M)$ , a question precisely analogous to the question relative to the law  $L(W)$ . To know whether the conditions of application of  $L(M)$  are satisfied so that the corresponding universal generalization  $U(M)$  can be used in deduction (II), it is necessary to know whether there is a property  $Z$ , such that  $Zc$  is both compatible with  $Dc \ \& \ Fc \ \& \ Mc$  and incompatible with  $\neg Bc$ , by virtue of a law  $L(Z)$  whose corresponding universal generalization is  $U(Z)$ . Once again, this question can only be answered if we already know whether  $L(Z)$  is applicable and therefore whether  $U(Z)$  may be used in the relevant deduction. But this depends in turn on whether there is a property  $Z'$ , such that  $Z'c$  is compatible with  $Dc \ \& \ Fc \ \& \ Mc \ \& \ Zc$  and incompatible with  $Bc$ , and so on.

Finally, if this reasoning is correct, in order to consider with Armstrong that laws contain their own conditions of applicability, one would have to accept the idea that each law contains a large number of other laws. By virtue of the fact that properties belonging to the domains of different sciences may interact to produce a certain effect, it is not possible to limit this conclusion to the laws relevant for the domain of the first law we started with,  $L(W)$ . However, a conception of laws of nature according to which each law contains a large number of other laws bearing on different types of properties, seems paradoxical: an essential part of the concept of law is the idea that a law determines a regularity across the diversity of situations. The difficulty is most obvious in the extreme case in which the slippery slope stops only after having included *all* laws: if there is only one law that applies always and everywhere, the notions of regularity and diversity and ultimately the notion of a law of nature, lose their meaning.

## Conclusion

We have proposed to explain the possibility of exceptions to empirical laws in a realist framework – even if an object falls within the domain of validity of a law because it possesses its antecedent property, two sorts of conditions must be satisfied in order for the law to be applicable to a concrete situation: 1) *no other property of the object itself* must interfere with the con-



straints imposed by the law, thereby preventing the exemplification of the consequent property, by virtue of laws bearing on those other properties; 2) *no other instance of the same property by another object* must interfere, by virtue of the very law at issue – this condition is meant to exclude interferences by instances of the *same property*. On the basis of the hypothesis put forward in Chapter 2 according to which (L) differs semantically from (U), we can express the role of the conditions of applicability in the following way. Let us suppose that object  $c$  falls under the law

(L) It is a law that  $(x) (Fx \rightarrow Gx)$ ,

by virtue of its possessing an instance of property  $F$ . This means that (L) imposes a constraint on  $c$ , to possess also an instance of the property  $G$ . However, to be able to conclude from this that  $c$  obeys in a given concrete situation also the universal generalization

(U)  $(x) (Fx \rightarrow Gx)$

corresponding to the law (L), or whether it behaves in this situation in an exceptional way, the two conditions (1) and (2) must be satisfied.

We may express the result of our analysis by saying that the conditions of application play the role of implicit premises in a deductive argument whose conclusion is  $Gc$ , from the premises  $Fc$  and (L). In order to establish that (L) applies to the concrete object  $c$ , it must be shown that the constraint imposed by (L) encounters no obstacle. Only then is it possible to substitute (U) for (L) in the deductive argument. To be able to conclude  $Gc$  from  $Fc$ , it is not only necessary that the law (L) be valid but also that it applies to the concrete situation, so that it is legitimate to use the corresponding universal generalization (U) in the deduction. The substitution of (U) for (L) is not legitimate in general but depends on the satisfaction of the conditions of application in each particular case.

## 4 Consequences

The distinction between laws and accidental universal generalizations we have elaborated in the last two chapters has important consequences for the role laws play in causation. Most of these consequences result from our thesis according to which most physical laws are not strict, in the sense that their validity alone does not guarantee that their conditions of application are satisfied in particular situations.

### Laws and counterfactual conditions

The following criterion, which is in general considered to be fundamental for the distinction between laws and accidental universal generalizations, turns out to be unacceptable in light of the results we have reached so far. The possibility of exceptions makes it strictly speaking false that laws, contrary to accidental universal correlations, are able to ground the truth of counterfactual conditionals.<sup>1</sup> It is true that accidental universal correlations are not sufficient to ground the truth of counterfactual conditionals. Goodman's (1955) famous example makes this obvious: from the fact that all coins that are now in my pocket are of silver one may not conclude that, speaking of some other coin, if that coin *were* in my pocket it *would be* of silver. I do not deny that there is an important difference between possible instantiations of laws and of accidental universal generalizations. It is more reasonable to expect the instantiation of a law, for example that this coin falls if I lose my grip on it, than to believe that this copper coin would be of silver if it were in my pocket.

However, it is incompatible with our analysis to stick to the thesis that the existence of a law  $A \rightarrow B$  is *sufficient* to guarantee the truth of a non-tautological counterfactual conditional whose antecedent refers to a possible instance of the antecedent property of the law:

If  $c$  were  $A$ , then  $c$  would be  $B$ .

It follows from the analysis given in the last chapters that the law does not by itself entail the truth of the counterfactual. The general strategy for the

evaluation of a counterfactual conditional<sup>2</sup> consists in considering the truth value of the consequent in a possible world differing as little as possible from the actual world, given that the antecedent of the counterfactual is true in it, which is not the case in the actual world. However, nothing guarantees that this counterfactual situation in which *c* is *A*, is not exceptional, in the sense explained in Chapter 3. We may try to add a clause to the counterfactual that explicitly excludes this possibility: “if *c* were an electron *and if the circumstances were not exceptional*, then *c* would be attracted by the positive charges in its proximity”. However, this makes the statement formally tautological, to the extent that it is impossible to make the clause excluding exceptions explicit in a noncircular way (cf. Chapter 3). On the other hand, if the counterfactual does not mention the possibility of an exception, it is strictly speaking false: the counterfactual conditional states that the consequent is true in the nearest possible world where the antecedent is true. However, the possibility of exceptional situations means nothing else than the existence of possible worlds in which the antecedent of the counterfactual is true whereas its consequent is false. One of these may well be the nearest possible world in which the antecedent is true: this is the case if the actual world contains circumstances whose counterparts in the counterfactual situation interfere with the consequences of the law given the counterfactual instance of its antecedent.

### **Homonomic and heteronomic laws**

Our analysis of the concept of law has led us to deny that there is a basis for the dichotomy distinguishing the strict laws of physics from the non-strict laws of special sciences whose applicability is subject to provisos. According to this widely accepted dichotomy, there is a special class of laws of nature holding only *ceteris paribus* characteristic of the special sciences, as opposed to the class of laws of physics.<sup>3</sup> With respect to the laws of special sciences, the advocates of this dichotomy can accept the conclusion of the last chapters that their applicability depends on a certain number of conditions, which are not part of the laws themselves and which cannot explicitly be stated.

Following Davidson (1970, p. 219ff.), the distinction that is supposed to establish the dichotomy is often grounded on the following criterion: the difference between strict laws and non-strict laws lies in the fact that the expression of the conditions of applicability requires, in the case of the latter, the use of a theoretical vocabulary belonging to sciences *different* from the science to which the law at issue belongs. The content of the conditions of applicability of laws of this sort is said to be *heteronomic*. However, there is thought to be a privileged class of strict laws whose conditions of applicability are *homonomic*: in their case, these conditions can be explicitly stated in the vocabulary of the science to which the law belongs.

This criterion, supposed to ground the dichotomy between strict and non-strict laws, faces three problems. The first stems from the fact that the

conditions of applicability cannot be exhaustively spelled out. However the proposed distinction between homonomic and heteronomic laws requires such an explicit statement, for only an explicit statement allows to distinguish cases in which the theoretical terms appearing in the conditions of applicability all belong to the same science as the law at issue, from cases in which their statement requires the use of theoretical terms belonging to other sciences.

The second problem is that there is a whole category of restrictions according to which the applicability of a law depends on whether the evolution of the object under consideration also depends on properties other than the antecedent property of the law (cf. Chapter 3, pages 103–15). However, we have no reason to expect that there are events or objects all of whose properties belong to the domain of only one science, at least as far as macroscopic objects and events are concerned. When we have analysed conditions of applicability (in Chapter 3), we have chosen examples among the laws of physics. This is the domain where the concept of a homonomic law has the greatest chance of being applicable, for the objects falling under laws of other sciences also fall under the laws of physics. The fact that all objects fall under the laws of physics provides a source of exceptions for the laws of other sciences. Therefore, if we can throw doubt on the existence of perfectly homonomic laws in physics, we shall thereby have thrown doubt on the general idea that there are homonomic laws. The concept of a homonomic laws presupposes a clear criterion of the unity of scientific disciplines. However, there seems to be no principled reason to refuse geophysics and other subdisciplines of physics the status of sciences, and grant that status only to physics as such. But if one accepts that the subdisciplines of physics are full-blown sciences, the thesis of the existence of homonomic laws loses much of its plausibility, at least as far as macroscopic objects are concerned. Macroscopic objects always possess properties belonging to (and falling under the laws of) more than one subdiscipline of physics. For example, the rotation of the earth around itself is determined both by geophysical laws and by the laws of celestial mechanics: the Moon causes the rotational movement of the Earth to slow down because the movement of the tides leads to a loss of energy by friction. This effect is a consequence of a geophysical law, whereas the conservation of angular momentum and of energy which are also relevant for the determination of the rotational movement of the Earth are laws of mechanics. This example shows that the latter laws are not homonomic because their applicability can depend on circumstances that are governed by laws not belonging to mechanics.

It may be objected that if the properties of an object fall under the laws of several subdisciplines of physics, their conditions of applicability can nevertheless be considered to be homonomic to the extent that they are all subdisciplines of *physics*. But the fact that these conditions cannot be explicitly and exhaustively stated makes it impossible to be a priori certain that no chemical or biological property is capable of interfering with the relevant physical properties.

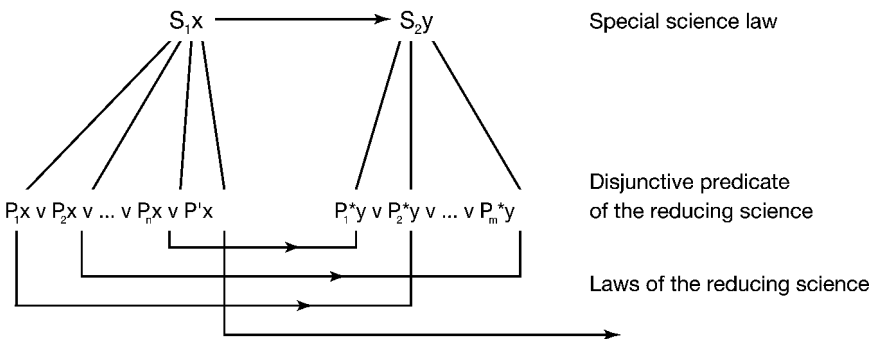
There is a third, and more fundamental, reason not to identify strict laws with homonomic laws. The proposition we are considering at present consists of *reducing* the concept of strict law to the concept of homonomic law. However, as our analysis of the concept of strict law in Chapter 3 shows, these concepts are independent. The law “As are Bs” is strict if and only if it entails the universal generalization ( $x$ ) ( $Ax \rightarrow Bx$ ). A law is homonomic if its conditions of applicability can be explicitly stated within the vocabulary of the science to which the law belongs. Even if we admit that the laws of physics are homonomic, this does not imply that they obey the criterion of strictness, insofar as they have nontrivial conditions of applicability. Being homonomic is not sufficient for being strict.

### Exceptions and multiple realization

Our explanation of the origin of certain types of exceptions is complementary to Fodor’s explanation of the possibility of exceptions to a law by the multiple realizability<sup>4</sup> of the properties related by this law. An exceptional situation to a law is a situation in which some of its conditions of applications are not satisfied. In Fodor’s terms, “an exception to a causal law is an event, which falls under the antecedent of a law and which is not the cause of an event falling under the consequent”. And he adds: “Presumably exceptions disconfirm strict law statements but not *cp* [ceteris paribus] law statements” (Fodor 1991, p. 24, note 7).

Fodor explains the origin of exceptions to laws of nature within the framework of an account of the reduction of special sciences to more fundamental sciences. A given law can have exceptions if and only if the properties involved in the law belong to a non-fundamental level, and are therefore realizable in many ways by properties belonging to a more fundamental level, in other words if these properties are “multirealizable”.

Let us consider the following schema (modified from Fodor 1974, p. 139; 1975, p. 20):



According to Fodor,

a nomologically sufficient condition for an exception to [the law]  $S_1x \rightarrow S_2y$  is that the bridge statements [i.e. statements linking the predicates of the reduced special science to the predicates of the reducing science] should identify some occurrence of the satisfaction of  $S_1$  with an occurrence of the satisfaction of a  $P$ -predicate which is not itself lawfully connected to the satisfaction of any  $P^*$ -predicate (i.e., suppose  $S_1$  is connected to  $P'$  such that there is no law which connects  $P'$  to any predicate which bridge statements associate with  $S_2$ . Then any instantiation of  $S_1$  which is contingently identical to an instantiation of  $P'$  will be an event which constitutes an exception to  $S_1x \rightarrow S_2y$ ).

(Fodor 1974, p. 142)

I do not doubt the explanatory value of this model. However, I contest that Fodor's account provides not only a sufficient but also a necessary condition for the existence of exceptions to a law. My claim is, in other words, that the type of exceptions Fodor's model accounts for is not the only type. However, even if he does not explicitly say so, Fodor at least strongly suggests in the following passage that he thinks it is:

We allow the generalizations of the special sciences to *have* exceptions [ . . . ]. [W]e expect both that there will be exceptions to the generalizations and that these will be "explained away" at the level of the reducing science. This is one of the respects in which physics really is assumed to be bedrock science; exceptions to *its* generalizations (if there are any) had better be random, because there is nowhere "further down" to go in explaining the mechanism whereby the exceptions occur.

(Fodor 1974, p. 143; italics Fodor's)

This passage clearly shows that Fodor thinks there is a nomologically fundamental level, which is the domain of strict exceptionless laws.<sup>5</sup> His remark that there may be "random" exceptions at the fundamental level is presumably to be understood as meaning that the fundamental level may contain nondeterministic laws. However, events whose probability is low according to a probabilistic law covering their appearance, are not exceptions in the sense we have given this concept in the context of deterministic laws.

Let us accept Fodor's thesis that a law can have exceptions on the condition that the properties it involves are multirealizable and therefore belong to a higher level with respect to a nomologically more fundamental level. However, our analysis of the conditions of applicability shows that there is an additional source of exceptions: they may arise from the violation of those conditions. The existence of this type of exceptions shows that the condition provided by Fodor is not necessary for the existence of an exception, in the sense that his model does not take this type of exception into

account; furthermore it is incompatible with Fodor's thesis that deterministic physical laws have no exceptions.

According to Fodor, exceptions are possible only in the case of special science laws involving multirealizable properties. This implies that the laws of physics, the ultimate reducing science, cannot have any exceptions. However, Fodor's model is in fact stronger than Fodor himself makes it appear. For within physics too there are multirealizable properties; laws bearing on such properties may therefore have exceptions Fodor's model can account for.

Let us take our example of the property of being a magnet and the law that the pieces of a magnet cut into pieces are also magnets with the same magnetization. This case illustrates the fact that Fodor's model is incomplete because there are exceptions of a type it cannot account for. Fodor's theory accounts for exceptions in cases in which an object falls under the concept of a magnet but in an atypical way. The concept of a magnet expresses a property belonging to an ontologically higher level: properties of that level are *realized* by properties belonging to a lower level; in our example, the property of possessing a permanent magnetization can be realized, among other possibilities, in an iron crystal by the fact that the magnetization vectors of all atoms are aligned in the same direction, or by a coil of copper wire in which an electric current flows, or by a circular superconductor in which a steady electric current flows. There are therefore exceptions to our law if (and according to Fodor only if) at least one of these three possible realizations is not nomologically related to a realization of the property expressed by the consequent of the law. This is indeed the case for the second realization that we might have already excluded insofar as our example requires the presence of a *permanent* magnet. For the other two realizations, the iron crystal magnetized by the alignment of the magnetic momenta of its atoms and the superconductor containing a steady current, no "Fodorian" exceptions seem to be possible: if a magnet (realized in one of these two ways) is cut into pieces, these pieces preserve the magnetic properties of the whole, and this guarantees the presence of the properties realizing the consequent of the law.

However, we have seen above (Chapter 3, pages 103–15) that there are sources of exception of a quite different kind. Even if an object possesses the antecedent property of the law it may behave in an exceptional way: even if a given object is a magnet it may not obey to a law L under which it falls by virtue of being a magnet. It can possess other properties that make its evolution and causal interactions diverge from those determined by the law L.

### **The inference problem**

Given our analysis of the relation between laws of nature and their instantiations, our account is not directly<sup>6</sup> vulnerable to an argument of van Fraassen's (1989) against realist theories of laws of nature, according to

which laws are relations between universals.<sup>7</sup> Van Fraassen objects to such theories that in spite of the satisfaction they may at first sight provide by making reference to universals, their explanatory force is only apparent because they fall victim to an infinite regress of the type of Plato's "Third man" (van Fraassen 1989, p. 103). We have introduced above the fundamental idea of Dretske's (1977b) version of the realist theory of laws. According to van Fraassen, the following inference leads inevitably in an infinite regression: the realist is not justified in concluding from a relation of fact between the universals *A* and *B*, of the type

(P) A-ness  $\rightarrow$  B-ness

to a material implication between the *exemplifications* of these universals, of the type

(U)  $(x) (Ax \rightarrow Bx)$ .

To justify this inference the realist needs to postulate universals of a new type: this is the first step of the infinite regress. Let us look more closely at this difficulty, which van Fraassen calls "the inference problem". The argument starts from the interpretation of the subordination of a universal property to another, according to which it entails a relation of extensional inclusion at the level of the exemplification of these properties. Let "R" designate a relation of subordination among properties, which is for the realist a law of nature. Then, says van Fraassen,

*R* is a *necessitation relation* exactly if it is a contingent relation, irreducibly and purely of some order  $>2$ <sup>8</sup>, and the statement that *R* holds between certain universals of order 1 entails that a certain other (corresponding) relation holds between particular instances of those universals. When that corresponding relation among particulars is exactly that all instances of the first universal are instances of the second, let us call *R* a *proper necessitation relationship*.

(van Fraassen 1989, p. 101; emphasis van Fraassen's)

The argument consists in asking under which conditions it is possible that the so defined relation *R* between universals entails a relation at the level of particulars. "How could a relation *R* with the initially noted features (contingent, irreducibly and purely of order *two*) also entail extensional inclusion?" (van Fraassen 1989, p. 102; emphasis van Fraassen's). *R* is supposed to be the nomic relation between properties the realist takes law statements (P) to refer to, and van Fraassen's extensional inclusion is the relation expressed by a universal generalization of form (U). Van Fraassen observes that, from a logical point of view, nothing can be deduced concerning the extensional inclusion (U) bearing on entities of order *zero* – i.e. particular



objects – from a premise that is purely and irreducibly of order *two*. The content of such a premise “will give no logical clue to what is true at orders below *two*. The entailment asked for cannot be logical entailment, just as a fact that is purely about the parents cannot logically imply anything about the children” (van Fraassen 1989, p. 102; emphasis van Fraassen’s). According to van Fraassen, the only way to obtain a valid inference having a premise at the level of universals and a conclusion at the level of particulars, is to introduce a new relation of order *three*. He points out that it is necessary to expand the word “entail” in our definition of necessitation as follows:

and the statement that *A* bears *R* to *B* entails that all *A* are *B*, in the following sense: there exists a(n impure) relation *R'* of order *three* which *R* bears to *A* and *B* and is such that, if any *X* bears *R* to *Y* and *R* bears *R'* to *X* and *Y* then it follows that all *A* are *B*, for the case of proper necessitation.

(van Fraassen 1989, pp. 102–103; emphasis van Fraassen’s)

Van Fraassen then argues that the word “it follows” in the last sentence raises the same difficulty as the word “entails” in the original definition of necessitation. It cannot refer to logical entailment; hence the only way to interpret it seems to require the introduction of a relation corresponding to a universal of order *four*; thus we are caught in an infinite regress of the “Third man” type. Since Plato’s idea to explain states of affairs at the level of particulars with the help of universals, the difficulty of explaining how universals can have consequences for their particular instances has provoked this criticism: explaining the relations existing between one order and the next seems to require postulating universals of still higher order.

The account we have presented in Chapter 2 above does not share with realism of universals the thesis that the law that all *As* are *Bs* is “a relation *R* irreducibly and purely of order two”, between universals *A* and *B*. It does therefore provide no target for the criticism that there is no valid inference from this relation *R* to a universal generalization of form (U) – van Fraassen’s “extensional inclusion”. According to our account, a law is a universal generalization bearing on *property instances*. This conception allows us to account for the possibility of exceptions to valid laws. The reason is that it is not legitimate to conclude from the validity of the law (L) the truth of the universal generalization (U) bearing on the *objects exemplifying those properties*. This account is not vulnerable to Fraassen’s objection because property instances are particular, although abstract entities, which are of the same logical order as the objects containing them. To explain both the relevance of a law (L) for the corresponding generalization (U) and the possibility of exceptions, we do not need to reason from states of affairs concerning entities of a logically higher order, i.e. of order two, to states of

affairs concerning entities belonging to a lower order, i.e. one. However, this step is the source of van Fraassen's "inference problem". According to van Fraassen, it is "entirely uncontroversial" (van Fraassen 1989, p. 29) that all realist theories of laws consider legitimate such an inference across different logical orders. However, this is so only for realism of universals. There is logical space for a theory of laws that is "realist" insofar as it considers the entities involved in a law to be real, without thereby taking them to be universals: they may be taken to be property instances. The fact that a property is never instantiated alone, but always in a concrete object exemplifying at the same time many other properties, is no reason to not consider each individual instance as real. In this framework, the application of a law to a particular situation does not require reasoning across logical levels. The conditions of applicability are stated at the same level as the laws themselves, i.e. the level of property instances.

Van Fraassen objects to the analysis of laws in terms of universals that "it simply does not seem that (irreducible higher-order) relations among universals can provide information about how particulars behave" (van Fraassen 1989, p. 126). Answering this challenge requires the elaboration of a satisfactory theory of properties and their instances. The mystery van Fraassen takes to be implicit in every theory making use of the concept of a universal, is not specifically related to *laws* construed as *relations* between universals. If there is indeed a mystery, it is already present in the very idea that a property is different from the class of objects possessing it.<sup>9</sup> But this mystery can be dispelled. Let us consider the following case. The property of bearing the elementary negative electric charge is different from the class of electrons, although it is exemplified by all electrons. Furthermore, the instances of the property of bearing the elementary negative charge are not identical with any of the objects containing them, i.e. electrons. The reason is simply that these objects have other properties, such as the property of having a mass equivalent to 511 keV. An object such as an electron *consists in* a set of property instances. No single property instance belonging to the set is therefore identical with the whole object. According to our account of laws, a nomic relation subordinates one of these properties to another, for example the property *A* of bearing the elementary negative charge, to the property *B* of being attracted by positive charges. The law *imposes a constraint* on every object exemplifying the first property, *A*. An object containing an instance of property *A*, in other words an object exemplifying *A*, also exemplifies *B*, if none of its other properties nor any other property exemplified by nearby objects interferes to prevent it. However, the possibility of exceptions makes the inference from the law to its instantiation in the case of this concrete object nontrivial. A simple law relates instances of only a few properties; but these properties are never exemplified alone, for the objects possessing them have many other properties independent of that law. The nomic subordination of the property *B* to

the property  $A$  has no analogue at the level of the *exemplification* of these properties by concrete objects. If a law relates the properties  $A$  and  $B$ , this does not entail that necessarily all or even most *objects* possessing  $A$  also possess  $B$ .

### Falsification

The thesis that most laws have exceptions may appear to threaten the possibility to refute, or to use Karl Popper's term, "falsify" hypotheses, insofar as they have empirical content. The falsification of a given nomic hypothesis

(L) It is a law that  $(x) (Ax \rightarrow Bx)$

may be thought to depend on one of the following two requirements. According to the first, the falsification of an empirical hypothesis is equivalent to the confirmation of a contrary hypothesis (first interpretation of the logic of falsification): in this case, there is no conclusive falsification without verification. This makes it impossible to insist, as Popper does, on the fact that inductive verification is strictly speaking impossible, without drawing the conclusion that falsification is impossible to the same extent and for the same reasons. According to the second, an empirical hypothesis can be falsified by a finite number of particular cases in which the antecedent of the nomic hypothesis (L) is satisfied but not its consequent (second interpretation of the logic of falsification). However, the fact we have established in Chapter 3 that the deduction of  $Bc$  from the premises  $Ac$  and (L) is not strictly valid, entails that the premises  $Ac$  and  $\neg Bc$  do not conclusively falsify (L). They only falsify (U) whose content is different from the content of (L).

It is not entirely clear which of these two accounts of falsification is Popper's own. When Reichenbach (1936, p. 269f.) argues against Popper's thesis that the verification and the falsification of a law obey a fundamentally different logic, his argument presupposes the second interpretation, according to which a finite number of singular propositions (in Popper's terminology, "basic statements") suffices to falsify a hypothesis of universal form.<sup>10</sup> However, when Carnap (1935) argues against Reichenbach and in defence of Popper, he adopts the first interpretation according to which a hypothesis of universal form cannot be falsified by singular statements, but only by another universal empirical hypothesis, which is confirmed independently: "A theory is said to be *falsified* if it is incompatible not only with a singular real event but with a repeatable effect; in formal terms: if it is incompatible with an empirical hypothesis that is confirmed"<sup>11</sup> (Carnap 1935, p. 290; emphasis Carnap's).

Popper himself seems to consider both types of falsification as legitimate. On one hand, the *Logic of Scientific Discovery* supports Carnap's interpretation:

We say that a theory is falsified only if we have accepted basic statements which contradict it [. . .]. This condition is necessary, but not sufficient; for we have seen that non-reproducible single occurrences are of no significance to science. Thus a few stray basic statements contradicting a theory will hardly induce us to reject it as falsified. We shall take it as falsified only if we discover a *reproducible effect* which refutes the theory. [. . .] If accepted basic statements contradict a theory, then we take them as providing sufficient grounds for its falsification only if they corroborate a falsifying hypothesis at the same time.

(Popper 1934, p. 54f.; transl. p. 86f.; emphasis Popper's)

In the same sense, Popper considers that a *theory*, containing hypothetical laws, can be falsified by another theory replacing it. He takes this to be the case with Newton's theory, which replaces Kepler's.<sup>12</sup> If this interpretation of falsification is correct, it would seem that Popper has not really established the logical superiority of falsification over verification: according to Popper, the first is logically valid because it is deductive whereas the latter is not, because it is inductive.<sup>13</sup> However, the passage quoted entails that, to be strictly valid, falsification would presuppose the *verification* of the falsifying hypothesis. This would give support to Reichenbach's (1936, p. 284) thesis that the falsification of an hypothesis is, from a logical point of view, not easier than its verification, for an hypothesis H is only "falsified" with the degree of corroboration that can be given to the contrary hypothesis required for the falsification of H.

On the other hand, in the following passage, Popper considers the falsification of an hypothesis with a conjunction of *singular statements*: "The 'problem of the demarcation' is what I call the problem of finding a criterion by which we can distinguish the statements of empirical science from non-empirical statements. My solution is the principle that a statement is empirical if there are (finite) conjunctions of singular empirical statements ('basic statements' or 'test statements') which contradict it" (Popper 1972, p. 12 note).<sup>14</sup> In his *Post-script*, Popper says in the same sense: "Hence the asymmetry is that a finite set of basic statements, *if true*, may falsify a universal law; whereas, *under no condition* could it verify a universal law" (Popper 1983, p. 185; italics Popper's). However, if our argument is correct, this interpretation makes conclusive falsification impossible. The refutation of a general hypothesis, in particular of a hypothetical law, depends directly on the entailment, by the law, of the universal generalization bearing on its exemplifications. The falsification results from the following modus tollens: the law (L) entails the generalization (U). However, there is an object (or a finite sample of objects), which exemplifies its antecedent but not its consequent. Therefore, (U) is false and, insofar as (L) entails (U), (L) is also false. In Popper's terms,

my proposal is based upon an *asymmetry* between verifiability and falsifiability; an asymmetry which results from the logical form of

universal statements. For these are never derivable from singular statements, but can be contradicted by singular statements. Consequently it is possible by means of purely deductive inferences (with the help of the *modus tollens* of classical logic) to argue from the truth of singular statements to the falsity of universal statements. Such an argument to the falsity of universal statements is the only strictly deductive kind of inference that proceeds, as it were, in the “inductive direction”; that is, from singular to universal statements. [ . . . ] The method of falsification presupposes no inductive inference, but only the tautological transformations of deductive logic whose validity is not in dispute.

(Popper 1934, pp. 15–17, transl. pp. 41–42)

In fact, singular observations incompatible with a given universal generalization are not sufficient to refute a nomological hypothesis. We have seen that Popper himself insists on this fact when he defends the first interpretation of the logic of falsification. Experimenters often find “counter-examples” in the sense of samples for which the proportion of two values taken by properties *A* and *B* does not coincide with the proportion given by the nomological hypothesis they are testing.

For example, the theory of electromagnetism contains a theorem deducible from Maxwell’s equations, according to which an electrical current *I* creates at distance *d* from the conductor through which it flows a magnetic field *B*, whose strength is given by:

$$2\pi \cdot I/d$$

But the measurements of current *I* and magnetic field *B* taken at distance *d* from the conducting wire, may yield a value for *B* that differs from  $2\pi \cdot I/d$ . The “falsificationist” theory must appeal to ad hoc reasons to avoid the conclusion that the law is “refuted” in each situation of this kind, because they are all “counter-examples” to the law. In fact, many other properties of the conductor and its environment have nomic effects on the two measured values. It is of course true that there is an objective difference between an exceptional situation in which some other property interferes and a case manifesting the falsity of the nomological hypothesis. However, contrary to the falsificationist thesis, it is illegitimate to conclude that a given nomological hypothesis is false on the basis of a case (or a finite number of cases) in which an object does not exemplify the properties related by a hypothetical law. This conclusion would have to be supported by an argument establishing that the situation cannot be explained by interfering factors.

It is important to note that this problem is different<sup>15</sup> from the problem raised by Duhem (1906, p. 303–328) and Quine (1961), according to which it is not possible to test scientific hypotheses one by one: their argument relies on the observation that it is in general not possible to derive observation statements from a single theoretical statement, but only from vast

sets of such statements. Therefore, if observation shows an observation statement to be true or false, this can confirm or refute only the *whole theory* required to deduce that observation statement, but not any single theoretical statement. As Quine famously puts it, “our statements about the external world face the tribunal of sense experience not individually but only as a corporate body” (Quine 1961, p. 41). The difficulty we have raised with falsification is different. The deduction of singular (and observable) effects from laws has the form of an enthymeme and is therefore not formally valid even if it includes in its premises the relevant theory as a whole. Therefore, the *modus tollens* corresponding to this deduction is not valid either.

Hempel also compares the difficulty raised by Duhem and Quine with falsification and confirmation with the difficulty stemming from the existence of nontrivial conditions of applicability (which he calls “provisos”), and concludes: “The argument from provisos leads rather to the stronger conclusion that even a comprehensive system of hypotheses or theoretical principles will not entail any  $V_A$  sentences [i.e. sentences in “observational” language] because the requisite deduction is subject to provisos” (Hempel 1988, p. 26).

However, it is controversial whether the two problems are distinct. Lakatos (1978) analyses the role of “*ceteris paribus* clauses” within the framework of the problem raised by Duhem, i.e. that one cannot test isolated theoretical hypotheses but only conjunctions of them.<sup>16</sup> Within Lakatos’ approach, the problem of factors limiting the applicability of a law to concrete situations appears indeed as a form of Duhem’s problem. According to Lakatos, in order to “test a *ceteris paribus* clause” in a situation in which one has observed an “anomalous” (Lakatos 1978, p. 26) phenomenon, one must suppose “that there *are* other influencing factors” (Lakatos 1978, p. 26; emphasis Lakatos’) at play, specify these factors and test the hypotheses stating that these factors are present. “If many of them are refuted, the *ceteris paribus* clause will be regarded as well-corroborated” (Lakatos 1978, p. 26). The inapplicability of a law in an exceptional situation is for Lakatos equivalent to the falsity of a *different* hypothesis, which he calls “the *ceteris paribus* clause” (Lakatos 1978, p. 26). But this means that Lakatos does not acknowledge what we have established in Chapter 3, i.e., that the conditions of applicability *cannot* be stated as explicit hypotheses.<sup>17</sup>

### **Perturbation theory**

After having so much insisted on the importance of exceptional situations, it may be useful to dispel a possible sceptical doubt. After all, the introduction of strict laws (or, in Armstrong’s (1983) terminology, “iron laws”) was first of all motivated by the wish to express the idea that regularity is indeed “the rule”, whereas exceptions are exceptional. However, the account we have presented seems to be incompatible with this simple idea. Almost all laws, and in particular physical laws, apply to particular situations only under the proviso that a certain number of conditions are satisfied; fur-

thermore, it seems as if there were few situations in which strictly no interfering factors are present. It might seem as if our account had the paradoxical consequence that exceptional situations are the rule. Does this consequence force us to abandon the fundamental conviction that nature is regular, a regularity without which there could be no science?<sup>18</sup> It does not. Perturbation theory provides the key to the solution of the apparent paradox according to which most concrete situations are exceptional with respect to any one nomic regularity.<sup>19</sup>

Perturbation theory plays for quantitative empirical sciences the role of a metatheory of the application of those theories to concrete situations. It allows us to classify the omnipresent interfering factors according to their order of magnitude, and then to compare them with the magnitude of the effect which is the object of inquiry. For example, the free fall of an electrically charged body is only partly determined by Newton's law of gravitational attraction: first, by virtue of its particular macroscopic form, the object is subject to friction that slows down its acceleration. Second, by virtue of its electric charge, it is subject to electromagnetic interactions imposing further constraints on its movement. As we have sketched the case, these factors, limiting the applicability of Newton's law of gravitational attraction (through the presence, in the object, of properties other than the mass involved in that law), result in forces whose magnitude is of a lower order of magnitude than that of the main force under study, determined by Newton's law. Perturbation theory allows us to classify the different components of the total force acting on the body, whose evolution one is concerned with explaining or predicting, according to their respective order of magnitude. For any desired degree of precision, there is a finite number of components the magnitude of whose contribution to the total force exceeds the threshold fixed by that degree of precision.

Perturbation theory allows us to cope quantitatively with the fact that, in a concrete situation, a given effect (such as the trajectory of an object) is in general determined by a large number of properties and interactions governed by different laws.

### **Conservation laws and transference theory**

The theory of causation defended in this book (in Chapter 1) reduces the causal relation to the transference of an amount of a conserved quantity between the cause and the effect. This reduction only makes sense if the amounts in question are indeed conserved, in other words if they are subject to a law of conservation. But given our result that almost all laws of physics are applicable only under the proviso of the satisfaction of a certain number of conditions that cannot be explicitly stated, one may wonder whether this is also true of conservation laws. However, conservation laws belong to the restricted class of "strict" laws; in other words, their application is not conditioned in the way we have explained.

There is a sense in which conservation laws are not “strict” either. According to a fundamental result of quantum mechanics, there is for each measurable quantity a complementary quantity, which is such that if the system is, with respect to one of these complementary quantities, in a so-called “eigenstate” (a state corresponding to an “eigenvalue” of this quantity), its state with respect to the other is completely indeterminate. For energy, the complementary quantity is time. The complementarity of energy and time entails that the energetic state of a system is not determinate up to an arbitrary precision, if the interval of time over which the energetic state is measured is itself finite. According to Heisenberg’s principle of uncertainty, the product of the indeterminations of the two quantities is always larger than (or equal to) the quantum of action:

$$\Delta E \cdot \Delta t \geq \hbar = h/2\pi$$

This entails that the energetic state of a system can, without any interaction, undergo spontaneous fluctuations of a size limited by their duration: the shorter its duration, the larger the fluctuation can be. Conversely, one can say that the energy contained in a system is conserved, in the sense that the degree to which it is determined is larger the longer the time over which it is measured. The conservation of energy is therefore also subject to a proviso, but of a type different from the conditions of applicability of laws we have examined in the last chapter. It is a limit of the intrinsic determination of the energetic state of a system, intrinsic in the sense of being independent of the particular circumstances.

This clarified, one can say that conservation laws are strict, in the sense that, for each such law, there is a true universal generalization bearing over the instances of the conserved quantity. In the case of the law of conservation of energy, the form of this generalization is:

$$U(\text{Cons}) (x)(t')(t) [(t' > t) \rightarrow (E(x, t) = E(x, t'))],$$

within the limits of precision imposed by the relation

$$\Delta E \cdot \Delta t \geq \hbar,$$

which means that for each isolated system  $x$ , and all moments of time  $t$  and  $t'$ , if the energy of  $x$  at time  $t$  is  $E$ , it is  $E$  at all later moments  $t'$ .

## Conclusion

One may express the result of our analysis by saying that a particular law of nature does in general not completely “cover” the situations in which its antecedent property is exemplified. A law is not equivalent to the universal



generalization according to which all particular objects exemplifying its antecedent property also exemplify its consequent property (chapters 2 and 3). In the present chapter, we have drawn several conclusions from that result.

1. The idea that laws but not accidental generalizations guarantee the truth of counterfactual conditionals is in general considered to be a fundamental criterion distinguishing laws from accidental universal generalizations. We have shown that the possibility of exceptions makes this criterion inoperative, at least in its simplest form. The truth value of counterfactual conditionals can only be determined with respect to particular counterfactual situations. Some situations being exceptional, the law is insufficient by itself to guarantee that the consequent of the counterfactual is true in the relevant (“nearest”) worlds in which its antecedent is true. The validity of the law (L)  $A \rightarrow B$  by itself is not sufficient to make true the counterfactual saying of object  $c$  that is actually not  $A$ , that “if  $c$  were  $A$ , it would be  $B$ ”, for in the nearest possible world in which  $c$  is  $A$ , the particular situation may well be exceptional with respect to the law (L).

2. A law is “homonomic” if its conditions of applicability can be explicitly stated within the vocabulary of the science to which the law belongs. Even if one accepts that physical laws are homonomic in this sense, this does not entail that they are strict, for the concept of strict law requires that there are no interfering factors that may possibly prevent the consequent from being true whenever the antecedent is true, whether those factors belong to the domain of the same or of different sciences.

3. Fodor’s account of the origin of exceptions is incomplete: it provides conditions that are sufficient but not necessary for exceptions. First, the laws of special sciences have exceptions of a type which Fodor’s model does not take into account. Second, even the laws of physics have exceptions, both exceptions of a Fodorian type and exceptions due to the violation of their conditions of applicability.

4. Our analysis of laws escapes the “inference problem”, which threatens according to van Fraassen (1989) all realist theories of laws. According to van Fraassen, the realist is not justified in inferring the truth of a universal implication between the exemplifications of two universals  $A$  and  $B$ , of the form (U), from a factual relation between these universals, of the form (P). However, according to our interpretation of the content of nomological statements, the inference from a law and the exemplification of its antecedent in a concrete situation, to the exemplification of its consequent is not valid independently of the particular situation. Exceptions are counterexamples to such inferences.

5. The fact that laws are in general applicable only under a proviso creates a problem for the falsification of nomological hypotheses, which differs from the problem raised by Duhem and Quine. If the falsification of an empirical hypothesis presupposes (or is equivalent to) the confirmation of a hypothesis incompatible with it, there is no falsification without verification,

and conclusive falsification is strictly speaking impossible for the same reason for which inductive verification is. However, our analysis shows that an empirical hypothesis cannot be conclusively falsified by a finite number of particular cases either: according to one interpretation of Popper's account of falsification, the relevant falsifying instances are situations in which the antecedent of the nomological statement (L) is satisfied but not its consequent. But we have shown that such situations do not refute the law because the deduction of the consequent  $Bc$  from the premises  $Ac$  and the law statement (L) is not valid.

6. Our analysis seems to lead to the paradoxical conclusion that exceptional situations are the rule. We have seen that perturbation theory can resolve this apparent paradox. Exceptions as we have understood them do not entail that nature is fundamentally and irreducibly irregular. On the contrary, exceptions to a law (L) are themselves due to laws. These laws are in general different from (L) and interfere with the constraints imposed by (L), thereby preventing (L) from being instantiated in the concrete situation. This accounts for the possibility of explaining exceptions scientifically.

7. The theory of causation defended in the first chapter makes an essential use of conservation laws. These laws are strict in the sense that their universal application is not subject to any proviso.

## 5 The nomological theory of causation and causal responsibility

In the first chapter of this book we have argued for the thesis that the ontological foundation of causation is to be found in the causal relation between *events*. We have suggested a reduction of that relation in physical terms. Two events are causally related if and only if an amount of a conserved quantity is transmitted from one to the other. This analysis is incompatible with the still very influential nomological theory of causation. There are two major differences between these rival accounts, regarding the role laws of nature play in causation. First, as we have seen in the first chapters of the book, the only type of laws appearing in the physical reduction of the causal relation are *conservation laws*. These laws are strict, in the sense that their validity excludes the possibility of exceptions. Second, the instantiation of a conservation law is, according to our analysis, not *sufficient* for the existence of a causal relation: in a closed system containing pseudo-processes, the global energy is conserved but this does not make the pseudo-processes causal.

### The nomological theory of causation

The nomological theory of causation, as it was defended by the logical empiricists, reduces causation directly to the instantiation of a law. The idea that all there objectively is to causation is the instantiation of a law has a long history in modern philosophy. One essential source, which we have considered in the first chapter, is Hume's analysis of causation. Expressed in contemporary language, its fundamental idea is this: the essential criterion for the truth of the claim that *c* is the cause of *e* is that there is a universal regularity according to which all events of the same type as *c* are followed by an event of the same type as *e*. At first sight, it might seem more appropriate to call this account "regularist" rather than "nomological". However, it turns out that both labels fit equally well because the advocates of the nomological account identify the *laws* of nature with universal *regularities*. Given this identification, we can express the nomological analysis in this way: to say that *c* is cause of *e* means that there is a law of nature that guarantees that an event of the same type as *e* occurs after every occurrence of an event of the same type as *c*.

Hans Reichenbach, a major representative of logical empiricism, expresses the nomological theory of causation just sketched in these terms: “Since repetition is all that distinguishes the causal law from a mere coincidence, the meaning of causal relation consists in the statement of an exceptionless repetition – it is unnecessary to assume that it means more” (Reichenbach 1951, p. 158). In the same vein, Rudolf Carnap says that “a statement about a causal relation [ . . . ] describes an observed regularity of nature, nothing more” (Carnap 1966, p. 201). To give a last example, Wolfgang Stegmüller sums up his analysis of causation in this way: “To say that *A* is cause of *E* is to say that there *exist* laws of nature which allow us, together with an exact description of *A*, to deduce a description of *E*” (Stegmüller 1969, p. 434; emphasis Stegmüller’s).

The analysis of causal relations and of laws of nature developed in the first chapter of this book permits us to evaluate the nomological account of causation. We shall begin this critical evaluation by recalling the most important reasons already stated for which it is impossible to *reduce* causation directly to the instantiation of a law. Then we shall add a new layer to the analysis given thus far, by introducing the concept of *causal responsibility*, which will allow us to integrate the explanatory strength of the nomological theory.

The following statement by Carnap is sufficiently precise to serve as a target for our critical analysis: “What is meant when it is said that event *B* is caused by event *A*? It is that there are certain laws in nature from which event *B* can be logically deduced when they are combined with the full description of event *A*” (Carnap 1966, p. 194). This statement shows clearly that what lies behind the idea of reducing causation to the possibility of subsuming a succession of events under a law, is the strategy to reduce causation to a form of *explanation*.<sup>1</sup> This project is, however, incompatible with realism with respect to causation. The reasons for which causation is not equivalent to explanation will show us how to modify Carnap’s analysis.

The first reason for not considering causation and explanation to be equivalent deserves particular attention. One cannot “logically deduce” a particular event but only a *fact* bearing on an event or an object. We shall later analyse more thoroughly the crucial concept of a fact, as opposed to that of an event. Reasoning on the deduction of facts from law statements and other facts, is reasoning about *explanation*. Facts have the structure of propositions, which can be expressed by statements. They are not concrete particular entities as events are. In these terms, our objection against Carnap’s analysis is that it reduces objective causation to an *epistemic* relation. According to Carnap, it is not only *our knowledge* of causal relations that allows us to explain and predict events, insofar as causation is itself subordinate to laws of nature. Rather, he defends the idea that causation only *exists* relative to the epistemic operations of explanation and prediction. Carnap sums up his analysis by this formula: “Causal relation means predictability”, where predictability is characterized in a way strictly parallel to

explainability. Both are analysed in terms of the logical deducibility of the fact to be explained or predicted from other facts (the “initial conditions”) and laws. “This prediction is a logical consequence of the facts and laws. In other words, there is a logical relation between the full description of the previous condition, the relevant laws, and the prediction of the event” (Carnap 1966, p. 193). Predictability is – as much as explainability – a logical relation between possible objects of knowledge. Hence Carnap’s analysis reduces causation to an epistemic relation.

Michael Scriven has also defended the idea that “a cause is *an* explanatory factor (of a particular kind)” (Scriven 1975, p. 11; emphasis Scriven’s) and that “causation is the relation between explanatory factors (of this kind) and what they explain” (Scriven 1975, p. 11).<sup>2</sup> But he has made it more explicit than others that this view implies that causation is reduced to an epistemic relation. Although he acknowledges the fact that there is an “objective correlative” to each cause, he suggests that what makes a given set of circumstances *a cause* are exclusively “psychological” factors: causes are determined by a “request for information” (Scriven 1975, p. 12). Scriven shows that the reduction of causation to an epistemic relation has a series of consequences:

1. If causation is irreflexive, in other words if *A* cannot be cause of *A*, this is only because the explanation of *A* by *A* is tautological, and lacks informative value.
2. *Events* cannot cause anything, for explanations do not, in general, provide knowledge of *events* but rather knowledge of *certain descriptions* of events. Scriven concludes that, according to the explanatory account of causation, causes are not events, but descriptions of events.
3. If birth is not the cause of death, this is only because the relation between them is known by everyone and can therefore not be the object of a request for explanation. In general only the most specific factor relevant for an explanation can be considered as the cause. It depends on the previous knowledge of the explanation seeker whether the presence of oxygen is or is not a factor causally responsible for the fact that the match burns.
4. Causality is not antisymmetric, in other words if *P* can explain, i.e. cause, *X*, the question as to whether *X* can at the same time explain and hence cause *P*, remains open.
5. The distinction between cause and effect, i.e. the peculiar asymmetry of the causal relation, is reduced to the asymmetry between explanans and explanandum.

The thesis that causation can be reduced to explanation must be rejected, for reasons to be given presently; however, it is also important to take into account the reasons that motivate this thesis. We shall do this by introducing the concept of *causal responsibility*: The truth value of a causal explanation does not only depend on the existence of a causal relation between two *events* but also on whether the *fact* designated as the cause is responsible for the *fact* designated as the effect. But let us first look at Scriven's account of causation as explanation. His analysis turns out to be unacceptable to us for the following reasons.<sup>3</sup>

Among the consequences mentioned by Scriven, (1) and (2) are incompatible with the realist account of causation developed in the first chapter of this book. Our defence of the realist account thus justifies the rejection of (1): *A* is not cause of *A* because nothing can be transferred from an event to that same event. This is independent of language, whereas the criterion of tautology applies only to linguistic expressions.

The same holds for (2): the truth of a causal statement presupposes the existence of a causal relation between two events. Scriven is wrong to identify the linguistic expression, the description of the cause, with what it refers to, i.e. the cause itself. However, what we shall retain from Scriven's account is that the content of a causal explanation exceeds in most cases the information concerning the existence of a causal relation between two events. We shall interpret this as meaning that the descriptions of cause and effect provide information about the *properties* involved.

Consequence (3) is also incompatible with a realist account. It is indeed a minimal requirement of a realist account of causation that whether *A* is objectively the cause of *B* does not depend on the particular interests of persons seeking an explanation of facts concerning those events.

(4) entails that there are causal relations in which each one of the two causally related events is cause and each one is the effect of the other. Then, given that causation is transitive, if *A* causes *B* and *B* causes *A*, *A* would also be a cause of itself. This is unacceptable insofar as *A* is a particular entity.<sup>4</sup>

(5) is incompatible with the fact that one can explain a cause in terms of its effects: the knowledge of *its effects* can contribute to the understanding of an event and therefore to its explanation. The logical structure of a deductive-nomological explanation does not provide any basis for the *asymmetry* of causation. On the basis of a law linking properties *A* and *B* one can as much explain an instance of *A* on the basis of an instance of *B* as one can explain an instance of *B* on the basis of an instance of *A*. This implies in particular that there are cases in which a law linking facts *A* and *B* allows explaining *A* by that law and *B*, but where *B* is not the cause of *A*. In general, that one fact *B* figures in the explanans of a correct explanation and another fact *A* in its explanandum, does not suffice to make *B* a cause of *A*. When the length of a shadow is mentioned in the explanation of the height of a flagpole, one explains the cause by the effect.<sup>5</sup> The explanans

does not refer to the cause and explanandum does not refer to the effect. Their logical and epistemic functions in the explanation are independent of the causal order.

This result can be generalized: not only is it possible to explain the cause on the basis of the effect, whereas the nomological account wrongly implies that only the opposite is possible, but it is quite clear that many explanations are not causal at all: in such explanations, neither the explanans nor the explanandum makes reference to a cause or to an effect of the other. Indeed, we have already encountered several types of such non-causal explanations. This provides a major reason to reject the idea that causality can be reduced to the relation holding between two items (facts or events) if and only if one can be explained in terms of the other by virtue of laws. Here are some important types of non-causal explanations:<sup>6</sup>

1. Mathematical explanations mention axioms and theorems. These may be considered as “logical laws”, but such explanations are completely independent of causation.
2. One can explain a property of an event by another property of the same event: the temperature of a gas is lawfully dependent on its pressure at the same instant, but one does not cause the other.<sup>7</sup>
3. The property of an object to be at temperature  $T$  can be explained in terms of its microscopic *reduction base*, i.e. by the property, of the components of the object, of having the mean molecular energy corresponding to  $T$ . With Cummins (1983, p. 20), we can call explanations of this type “analytical explanations” and observe that they are not causal: the constituents of a composite object, and their properties, are not among the causes of the whole object and its properties.
4. A lawful correlation between two effects of a *common cause* can ground an explanation of one effect by the other although there is no causal relation between them: if  $A$  causes  $B$  and  $A$  causes  $C$  by virtue of a law of nature, it is possible that there is a nomic correlation between  $B$  and  $C$ , without  $B$  being the cause of  $C$  nor  $C$  being the cause of  $B$ . This happens, e.g. if  $A$  is the emission of a radio signal and  $B$  and  $C$  are events of reception of the signal situated in opposite directions with respect to  $A$ . In that case, and if the signal spreads at the speed of light, the existence of a causal relation between  $B$  and  $C$  is impossible; the distance between  $B$  and  $C$ , in the terminology of special relativity, a “spatial” distance, is impossible to bridge by causal relations.<sup>8</sup>
5. An epiphenomenon of an underlying causal process can explain another epiphenomenon of the same process: Sober (1985) shows that the degree

to which a given trait is modified by natural selection can be explained in terms of the force of selection acting on the trait and the inheritability of the trait. Such an explanation, e.g. the explanation of the increase of the mean size of the animals in a certain population from one generation to the next, is valid without bringing the concept of genome into play. However, insofar as it is exclusively expressed in terms of phenotypes, the explanation is not causal: phenotypes can be understood as epiphenomena of an underlying causal process involving genotypes.

6. Pseudo-processes may serve as bases for explanations: when the phase of a wave spreads faster than the speed of light, the phases at different spatio-temporal positions are lawfully related by virtue of the law of phase propagation but they are not causally related.<sup>9</sup>
7. Finally, the teleological explanation of an action is not causal in our sense either: one typically explains an action by citing the reasons for which it has been undertaken. One indicates the action's goal, which consists typically in events happening or states holding in the future. Such future events or states cannot be said to be among the causes of the action.<sup>10</sup>

It is significant to note that our conclusion, that not all explanations are causal, does not depend on the adoption of any particular model of scientific explanation. It rests particularly valid in the framework of the realist account of explanation proposed by Railton<sup>11</sup> which construes the explanatory relation as depending only on objective facts and laws but not on our knowledge of them.

As we have noted, the reduction of causation to the transmission of conserved quantities presupposes the validity of conservation laws. The existence of such laws is necessary for the existence of causal relations. However, the fact that an event falls under a law because it possesses the antecedent property of a law is not sufficient for its being the cause of another event that possesses the consequent property. Indeed, in the explanations of the types of (2) through (7) listed above, one fact is explained in terms of another fact by virtue of a law without the first being a cause of the second.

But we have not yet mentioned a crucial reason for rejecting the reduction of causation to the instantiation of a law, and in particular Carnap's version of this doctrine. This reason stems from the reflection on the conditions of applicability we have analysed above (Chapter 3). There we have seen that one cannot deduce the fact that the consequent of a law is exemplified from the fact that its antecedent is. Such a deduction is invalid insofar as it bears upon the application of a law to a particular situation which is subject to provisos.



Furthermore, Carnap's analysis entails that the existence of a causal relation between *A* and *B* depends on the possibility of deducing (a description of) *B* from an *exhaustive* description of *A*. It makes no sense to say of an event, in the sense we give to this word, that it can be "deduced". What may be deduced and hence become an explanandum belongs to the ontological category of facts. However, Carnap's criterion is expressed in terms of events. The difference is important because it is in general not necessary to explain *all* facts concerning event *B*, nor even to explain all facts concerning its intrinsic properties, because a request for explanation is usually satisfied by the explanation of just one fact. Hence it is not necessary to require an "exhaustive description of event *A*", in other words, the conjunction of all facts about *A*. This restriction in what is required of the explanandum and the explanans entails an analogous limitation of the number of laws on which the causal relation should depend. Only laws bearing on the properties that are part of the causally related facts come into play. If the causal relation between events *A* and *B* necessarily brought into play all facts about *A* ("an exhaustive description of *A*"), as Carnap holds, it would be impossible to express it in language: the number of facts about *A* is indeterminate.

Our conclusion that it is impossible to reduce causation directly to a form of explanation is compatible with Davidson's (1967) distinction between statements directly designating causal relations and statements providing causal explanations. Davidson notes that some causal statements cannot be analysed within the purely extensional model, which he takes to be appropriate for statements designating causal relations. The following statements, "The collapse was caused, not by the fact that the bolt gave way, but by the fact that it gave way so suddenly and unexpectedly" and "The fact that the dam did not hold caused the flood" (Davidson 1967, p. 161), contain more than a reference to a causal relation between events. According to Davidson, such statements express "rudimentary causal explanations" (Davidson 1967, p. 161). He argues that the word "cause" does not, in statements of this type, make reference to a causal relation but must be interpreted as meaning "causally explains" (Davidson 1967, p. 162). Indeed, there are according to Davidson two sorts of causal statements: statements making reference to a causal relation between events and rudimentary causal explanations. However, the terms of the latter are neither events nor any other natural entities: "Explanations typically relate statements, not events" (Davidson 1967, p. 161).

The idea that will guide us through the rest of this book is that there is a logical gap between causal explanations and causal relations between events. For a "rudimentary causal explanation" to be true, it does not suffice that it designates two causally related events. What determines the truth value of a causal explanation is more than the existence of a causal relation between events. However, Davidson does not make any suggestion as to the nature of the condition that must be added to the truth conditions of a

causal explanation, beyond the existence of such a causal relation between events.

Our suggestion is that such a rudimentary causal explanation is true if and only if 1) the explanans and the explanandum are true, in other words state facts, and 2) the fact expressed by the explanans is *causally responsible* for the fact expressed by the explanandum. The relation of causal responsibility of the fact that  $c$  is  $F$  for the fact that  $e$  is  $G$  can be analysed into two parts that are individually necessary and jointly sufficient:  $c$  and  $e$  must be events (or temporal parts of objects, as explained above in Chapter 1) causally related by transference and there must be a law applying to the situation that subordinates property  $G$  to property  $F$ , so that the possession of  $F$  by the cause brings about the possession of  $G$  by the effect. This law need not be a conservation law, and it does not need to be strict.

Let us designate the causal relation between the events  $c$  and  $e$  in terms of transference by  $C_E(c, e)$ , and the relation of causal responsibility of the fact  $Fc$  for the fact  $Ge$  by  $C_F(Fc, Ge)$ . If “ $Ge$  because  $Fc$ ” is a rudimentary causal explanation,  $C_F(Fc, Ge)$  designates the relation between the facts  $Fc$  and  $Ge$  that makes the explanation true, in other words its truth-maker.

This analysis allows us to integrate an important truth contained in the nomological theory. Indeed, our concept of causal responsibility is close to Carnap’s (1966) notion of causation. It is a relation between two facts  $Fc$  and  $Ge$ , such that these facts bear upon events  $c$  and  $e$  causally related by transference, so that  $C_E(c, e)$ , and such that the properties  $F$  and  $G$  attributed to these events are related by a law. However, causal responsibility differs from Carnap’s causation in two important respects. First, the fact that  $Fc$  is causally responsible for  $Ge$  *presupposes* that the events  $c$  and  $e$  they bear on are causally related. Our examples (2) through (7) of non-causal explanations show that the conditions imposed by Carnap are not *sufficient* for the existence of a causal relation. In all these cases, no causal relation between events  $c$  and  $e$  exists to make true  $C_E(c, e)$ . It is the fact that Carnap’s analysis does not mention this condition that makes it inappropriate as a reduction of causation.

Second, Carnap interprets causal statements as stating that it is possible to deduce a description of the effect from a description of the cause and a law statement. According to our analysis (Chapter 3), such a deduction is valid only on condition that the law is applicable to the situation in which the causal relation holds. If the situation is exceptional with respect to the relevant law, the satisfaction of Carnap’s conditions is not *necessary*:  $c$ , having property  $F$ , causes  $e$ , having property  $H$  but not property  $G$  although there is a law subordinating  $G$  to  $F$  and although there is no law subordinating  $H$  to  $F$ ; in other words,  $e$  is  $H$  for reasons independent of  $c$ ’s being  $F$ .

Davidson (1967; 1970) introduces the following qualification in his version of the nomological theory. For two events to be causally related, it is not in general sufficient that they can instantiate an arbitrary law of nature: only strict laws can ground a causal link between events. This requirement can be understood by considering the fact that Davidson follows the

Humean tradition in conceiving causation in terms of regularity. According to the regularity view, “*A* causes *B*” is true only if all events of type *A* are followed by an event of type *B*. The causal relation is reduced to an instantiation of a law, and laws are reduced to regularities. Insofar as it conceives of a law as being nothing else than a universal regularity, this approach makes laws with exceptions inconceivable. If there are exceptions, the regularity is not strictly universal and is therefore not a law after all. Following this line of reasoning, Davidson concludes that only the instantiation of a strict law is apt to ground a causal relation.<sup>12</sup>

We have found two reasons to reject Davidson’s analysis: First, it is not because conservation laws are strict that their instantiation is necessary for causation. Rather, it is necessary because they guarantee the persistence of the individual amount of a conserved quantity that is transferred between cause and effect. Second, the instantiation of a conservation law is not sufficient for causation. Even if conservation laws are in fact strict, they are only one component in a more complex set of conditions sufficient for causation in terms of the transference of an amount of a conserved quantity.

Davidson’s analysis also suffers from a circularity due to his criterion of the identity of the events that are the terms of causal relations. Davidson offers this “clearly acceptable criterion for the identity of events [ . . . ]: events are identical if and only if they have exactly the same causes and effects” (Davidson 1969, p. 179). However, a theory that uses the concept of event in its analysis of the causal relation, becomes circular if it uses the notion of causation in the conditions of identity of those events.

This critique should not obliterate the debt that my account of causal responsibility, and in particular the postulate of a second aspect of causation beyond the relation between events, owes to Davidson’s analysis as it is developed in his “Causal Relations” (Davidson 1967). We shall pursue our comparison with Davidson’s account after having developed the concept of causal responsibility in a little more detail.

## **A second aspect of the concept of causation**

The causal statements that are used in common explanations contain more than the mere reference to two events and to the fact that they are related as cause and effect. The concept expressed in common causal statements has a more complex logical structure than the simple causal relation between two events. In most cases the expressions “causes”, “brings about”, “provokes” or simply “because” must be interpreted as expressing more than the technical predicate “. . . is causally related to . . . ” or “. . . is the cause of . . .”.<sup>13</sup> If we take the latter to express an extensional relation<sup>14</sup> between events, they are not synonymous with the ordinary predicate “cause” and the other causal idioms we have mentioned.

To elaborate the semantic distinction between two types of causal predicates, and between the aspects of the concept of causation they express,

we shall make use of the ontology of events introduced earlier in the book (Chapter 1, pages 44–7), constructed on the basis of the analogy between events and objects: an event is with respect to space–time what an object is with respect to space. Just as objects can be identified with what occupies a limited portion of space, an event can be conceived as what occupies a limited portion of space–time. This concept of event allows us first, to attribute different aspects or properties to one event and second, to distinguish the direct reference<sup>15</sup> to an event from the reference to an event by virtue of one of its aspects.

Before we develop, with the help of some examples, the reasons for postulating two fundamental aspects of the concept of causation, we should recall that our analysis is grounded on the presupposition of causal realism. We share with commonsense the realist interpretation of causal explanations, according to which the existence of events as much as their causal relations to other events are independent of our way of considering and describing them. Doesn't this make the truth of a causal statement independent of the particular descriptions that we use respectively for the cause and the effect? The intuition that this must be so seems to be contradicted by an observation one can make upon examination of the following pairs of statements.

- (1) S has caused an embarrassment by saying “Hi” to the president during a solemn ceremony.
- (2) S has caused an embarrassment by saying “Hi” to an old friend.<sup>16</sup>
- (3) The electron has a stable position with respect to the nucleus because it has negative charge.
- (4) The electron has a stable position with respect to the nucleus because it has a mass equivalent to 511 keV.<sup>17</sup>

We have to explain why there may be situations in which (1) is true yet (2) false although these statements make reference to the same causal relation between the same events, in the sense of “event” introduced above. With respect to the designation of the cause, comparing (1) and (2) shows that causal statements of this type do not only mention a property of each event but that these properties play a role in the semantic evaluation of the statement. We cannot interpret the word “cause” in (1) and (2) as expressing a simple relation between two events, for this would not allow us to explain why they differ in truth value. (1) but not (2) designates, over and above the cause, the *property* that is *efficacious* in bringing about the effect. (1) but not (2) designates the fact *causally responsible* for the embarrassment, i.e. *the fact that S said “Hi” to the president during a solemn ceremony*.

Similarly, although statement (4) designates by hypothesis the same events as statement (3), it nevertheless does not pick out the fact respon-

sible for the stable state. The property referred to by the expression denoting the cause in (3), i.e. the property of *bearing negative charge*, plays a central role in the causal relation expressed, whereas the property of *having a mass equivalent to 511 keV* is not lawfully linked to the stable position of the electron. This second example of diverging truth value of two statements making reference to the same causal relation between the same events is in a way more radical than the first: the two properties of the cause designated respectively by (3) and (4) share their extension in the actual world; it is therefore impossible to explain at a purely extensional level why (3) is true but (4) false. Exactly the same pairs of events satisfy the descriptions occupying the positions of cause and effect in these two statements.

To explain why they nevertheless do not share their truth value, we can make the hypothesis that statements (1) through (4) do not have the logical structure  $C_E(c,e)$  of a binary predicate whose arguments refer to events  $c$  and  $e$ , in other words that these statements do not express a causal relation between two events. Rather they have the more complex logical form of statements expressing causal responsibility between two facts, one fact,  $Fc$ , bearing upon the cause  $c$  and the other,  $Ge$ , bearing upon the effect  $e$ . According to this interpretation, the logical form of the arguments of the causal predicate in these statements is propositional, and these arguments designate facts, whereas the logical form of the arguments of a causal statement interpreted on the model of  $C_E(c,e)$  is “eventive”: this means that its arguments make reference to events. In order to make reference to a causal relation in the most complete sense of the concept, a statement must not only specify the events related but also those of the properties on which the relation depends. The additional structure of factive arguments, compared to arguments designating events or objects, opens up the logical space allowing us to account for the differences of truth value we have observed. Even if the reference to events is held fixed, two causal statements can differ in truth value insofar as they differ with respect to the expression making reference to the efficacious property of the cause, or to the property of the effect affected by the cause. (3) is true and (4) is false because they differ with respect to the property of the cause they pick out as efficacious although these expressions make reference to the same events.

(C) is the general form of a statement expressing a causal relation between facts. It indicates the properties,  $F$  and  $G$ , of the cause and the effect that are involved in the causal relation referred to, by virtue of a law subordinating  $G$  to  $F$ .

(C)  $e$  is  $G$  because  $c$  is  $F$ .

The logical form of a predicate with propositional (or, as we shall also say, factive) arguments gives parallel roles to expressions referring to objects or events  $c$  and  $e$  and to expressions referring to properties  $F$  and  $G$ . Let us represent the logical form of statement (C) by

(O)  $C_F(Fc, Ge)$ .

$C_F$  represents a causal predicate with factive arguments, expressing respectively the fact that the cause has property  $F$  and that the effect has property  $G$ .  $C_F$  expresses the concept of *causal responsibility*, which is the concept of causation most immediately relevant in contexts of explanation. Statements expressing causal responsibility have a richer content than eventive causal statements. This can be seen from the fact that one can infer from (O) a causal statement with eventive arguments but that the converse inference is not valid. The possibility of such an inference can be understood on the basis of the following analysis: (O) is equivalent to the following conjunction:

$$(O') C_E(c, e) \wedge Fc \wedge Ge \wedge L(F, G).$$

According to this analysis, factive causal statements of type (O) contain the following elements. First, events  $c$  and  $e$  are causally related. This is expressed by the predicate  $C_E$  taking these events as its arguments. Second, the cause  $c$  possesses property  $F$  and the effect  $e$  possesses property  $G$ ; finally properties  $F$  and  $G$  are linked by a law of nature  $L(F, G)$  in such a way that the possession of  $F$  by  $c$  nomically determines the possession of  $G$  by  $e$ .

In Quine's terminology (cf. Chapter 2), a statement containing the predicate  $C_F$  is *opaque* insofar as its semantic evaluation depends essentially on the reference to properties  $F$  and  $G$ . A statement having the logical form (O) does not always preserve its truth value when the predicate  $F$  is replaced by a predicate  $H$ , designating a property with the same extension, insofar as  $F$  but not  $H$  is causally responsible for the fact that the effect has the property  $G$ . What makes one property  $F$  efficacious in bringing about another property  $G$  is the instantiation of a law linking these properties. The intensional aspect of the relation of causal responsibility and the opaque aspect of the causal statements expressing it, stem from the dependence of causal responsibility on laws of nature. Law statements contain opaque contexts by virtue of the intensional character of the relations between properties they express (cf. Chapter 2). The instantiation of laws gives the relation of causal responsibility and the causal statements expressing these relations their intensionality or opacity. According to analysis (O'), the opacity of (O) stems from the presence of the nomic operator  $L(F, G)$ , which expresses the fact that a law of nature relating  $F$  to  $G$  applies to the concrete situation in which these properties are exemplified by  $c$  and  $e$ .

Now we can explain the difference in truth value between (1) and (2) and between (3) and (4). (1) and (3) are true and (2) and (4) are false although they all contain an element of truth corresponding to  $C_E(c, e)$ . However, the nomic element  $L(F, G)$  is true only in (1) and (3) but not in (2) and (4), because there is a law relating the properties designated by the relevant expressions in (1) and (3) and applying to the situation, but no such law relates the properties designated by (2) and (4).

The analysis also shows that the relation expressed by a statement of form  $C_F(Fc, Ge)$  contains an extensional element  $C_E(c, e)$ . This is why factive statements  $C_F(Fc, Ge)$  entail eventive statements of form  $C_E(c, e)$ . From the conjunction (O'), one may validly infer its first term, i.e.

$$(E) C_E(c, e).$$

(E) says only that  $c$  causes  $e$ , in other words it states the existence of an indeterminate causal link between events  $c$  and  $e$ . (E) does not contain more than the minimal information about a causal relation, consisting in the mere statement of its *existence*.

To infer a statement of the form  $C_E(c, e)$  from a statement with the more complex form  $C_F(Fc, Ge)$ , one must substitute for the propositional expressions  $Fc$  and  $Ge$  simpler expressions that refer *directly* to events  $c$  and  $e$ . The former expressions also refer to these events, but they do this by mentioning those of their properties that are involved in the causal relation. The argument places of statements of the form  $C_E(c, e)$  must be filled by eventive expressions. The truth of a statement of form  $C_E(c, e)$  is independent of the type of causal process linking  $c$  to  $e$ , as well as of the properties involved in that process. The concept expressed by the predicate  $C_E(c, e)$  is directly reducible to the physical concept of causation analysed in the first chapter of this book. The events  $c$  and  $e$  are related as cause and effect if and only if an amount  $Q$  of some physical conserved quantity is transferred between them.

We can reach the same result in the following more indirect way. Quantifying existentially over the predicates  $F$  and  $G$  in (O) yields:

$$(E') (\exists F)(\exists G) C_F(Fc, Ge),$$

which is equivalent to:

$$(E'') (\exists F)(\exists G) (C_E(c, e) \wedge Fc \wedge Ge \wedge L(F, G)).$$

(E') and (E'') express existential generalizations over the properties  $F$  and  $G$  involved in the causal relation. They say only that *there is* a property (or several properties)  $F$  of the cause and a property (or several properties)  $G$  of the effect, such that the causal process from  $c$  to  $e$  depends on the exemplification of  $F$  by  $c$  and of  $G$  by  $e$ . The content of such a statement has two components. The first corresponds to the *existence* of a causal relation between the events  $c$  and  $e$ , independently of the properties involved. This corresponds to the content of the transparent statement (E). The second component corresponds to the general statement that every causal relation is due to the exemplification of properties related by a law of nature. However, this statement does not contain any specific information about these properties. The existential quantification removes all information about the properties involved in the specific causal relation between  $c$  and  $e$ .

On the basis of our hypothesis that there are two types of predicates, represented by  $C_E$  and  $C_F$ , corresponding to two aspects of causation, we can state two theses about ordinary language causal statements, which raise an important problem: each of them is intuitively plausible, but taken together they seem to be inconsistent.

First, in analogy with Quine's terminology introduced earlier, statements (1) and (3) contain expressions placed in a *semi-opaque* context. This type of context is characteristic of the causal predicates of natural language. As with opaque contexts, the substitution of expressions contained in such a context by coextensional (or coreferential) expressions does not always preserve the truth value of the statement. This is exactly what is illustrated by statements (2) and (4), which result from such a substitution in statements (1) and (3) respectively. The reason why causal contexts are only semi-opaque rather than fully opaque is that such contexts allow for existential quantification *salva veritate*.<sup>18</sup> Causal statements such as (1) and (3), whose logical form is  $C_F(Fc, Ge)$ , entail statements of the form

$$(\exists x) C_F(Fx, Ge),$$

and, as we have seen,

$$(\exists F) C_F(Fc, Ge).$$

In other words: if a statement such as (1) or (3) is true, there is an event  $x$  that is the cause of the event  $e$  and there is a property  $F$  causally responsible for the effect  $e$ 's being  $G$ . If (1) is true, there is an event causing the embarrassing situation; if (3) is true, there is (at each moment) an event of causing the stable position of the electron with respect to the atomic nucleus. This observation can be summed up by the following thesis:

**Thesis I:** Ordinary language statements possessing the structure " $Fc$  causes  $Ge$ " have the logical structure (O) equivalent to (O'). They are semi-opaque insofar as they do not allow the substitution *salva veritate* of  $F$  and  $G$  by coreferential expressions  $F'$  and  $G'$ , but do allow existential generalization *salva veritate* over events  $c$  and  $e$  and over properties  $F$  and  $G$ .

The following thesis seems both intuitively plausible and incompatible with Thesis I.

**Thesis II:** Ordinary language statements possessing the structure " $Fc$  causes  $Ge$ " are extensional.

However, we shall see that they can after all be reconciled. According to this thesis, the causal predicates of ordinary language express a concept of causa-



tion that corresponds to a simple two-place relation: their arguments do not occupy opaque positions. Equivalently, Thesis II says that causal statements, insofar as they express *relations*, are transparent. This thesis has in particular been defended by Davidson<sup>19</sup> who argues that objective causal relations are expressed by an extensional causal predicate, which we have represented by  $C_E(c, e)$ . Those terms of the conjunction (O') that bring specific properties into play must be considered as elements of causal *explanations*, which introduce information exceeding the content of a causal statement. As Rosenberg and Martin (1979) show, the thesis of the extensionality of causal statements follows indeed from the adoption of causal realism. However, to reach Davidson's conclusion that causal statements do not only contain an extensional element but are totally extensional on the model of (E) or (E'), one must add the premise that the properties *F* and *G* and their nomic relation are not objective in the same sense as the events possessing these properties.

Theses I and II are clearly incompatible if they are understood as bearing on the whole content of a causal statement *C* in ordinary language. However, they can be reconciled as soon as Thesis II is understood to bear only on an *element* of the content of *C*. If one accepts that (O) and (O') express the logical structure of an ordinary causal statement, it appears that Thesis I is true of the statement because it contains opaque positions. In schema (O') these positions appear within the scope of the nomological operator *L*. Thesis II can be reconciled with Thesis I if it is taken as bearing on a *part* of the content of *C*, corresponding to the term of the conjunction (O') equivalent to the transparent statement schema (E). Thesis II is justified to the extent that it can be read as stating only the *possibility* of an extensional interpretation. However, such an interpretation does not exhaust the full content of statement *C* and does justice only to one of its components. Thesis II is true of a causal statement *C'* weaker than *C* and entailed by it. The weaker statement *C'* has the structure  $C_E(c, e)$ , whereas the full causal statement *C* has the structure  $C_F(Fc, Ge)$ .

Before we consider some objections to an account that acknowledges as terms of causal relations, both facts and events, it is useful to examine the inferential relations between statements of the two sorts. If causal predicates can take as arguments both eventive and factive expressions, the analysis of those inferences will help us understand the logical relations between eventive and factive causal statements.

## Inferences between factive and eventive causal statements

### *When does a factive causal statement entail an eventive causal statement?*

Given the logical forms (O') and (E), it is clear that (O') entails (E) but that the converse does not hold. However, the inferential relations between factive and eventive causal statements turn out to be more complex, as soon as one takes their explicit formulations in ordinary language into account. We

shall now try to find out under which conditions inferences of the following type are valid:

- (I1) FACT is the cause of ... (is causally responsible for ... )  
Therefore, EV is the cause of ...

where “FACT” represents a factive expression and “EV” an eventive expression.<sup>20</sup> Factive expressions, in the sense we give to this locution, have the logical structure of propositions: they attribute a property to an object. By contrast, eventive expressions designate events. It will be easier to bring out the difference between a fact in our sense from an event which is a *particular* entity, by considering an example.

The inference from (5) to (6) provides an illustration of inference type (I1):

- (5) The fact that Mary sang the song surprised me.  
(6) Mary’s singing of the song surprised me.

Intuitively, this inference is valid to because the relation between facts expressed by (5) contains more information than (6) about the property of the cause that has brought about the effect. (5) does not only say that Mary’s performance possesses some property or other, by virtue of which it has caused my surprise, but does explicitly indicate *what that property is*: it is the property of being the singing of the song by Mary. (6) contains less information: it only tells us that the particular *event* of Mary’s singing the song possesses *a* property, without telling us which property, by virtue of which it has produced my surprise. For example, (6) may be true if Mary habitually sings the song in her own bathroom but is too shy to perform on stage. In that case, a particular property of the performance to which (6) makes reference may be responsible for the surprise: her singing it on stage. But (6) neither contains the information that the song was performed in public nor the information that this was its surprising aspect. In such circumstances, (5) is false. (5) designates as the cause of the surprise *the fact that* Mary sings the song, which excludes the possibility that what is responsible for the surprise is a particular aspect of the event of her singing it. Indeed, in this situation, the fact that she sings the song is not surprising at all. This shows that (5) contains more information, or is informationally richer than (6): if (5) were true, i.e. if the mere fact that she sings the song were already surprising, it would entail (6) – in that case, the particular singing would cause surprise. The situation sketched, in which (5) is false although (6) is true, shows that the inference from the weaker (6) to the stronger (5) is not valid.

It is important to note that the inference from (5) to (6) is not valid by virtue of the mere logical form of the premise. It can only be drawn to the extent that the statement “FACT is causally responsible for ... ” can

be analysed explicitly into components corresponding to the conjuncts of (O'):  $C_E(c,e) \wedge Fc \wedge Ge \wedge L(F,G)$ .<sup>21</sup> On that condition, one can, on the basis of its logical form, explicitly infer from it the eventive statement  $C_E(c,e)$ . The difficult question is under which conditions the expression FACT – designating the causally responsible *fact* – contains the information required for constructing the expression EV making reference to the cause *event*.

It turns out that it is not always possible to construct an eventive statement entailed by the corresponding factive causal statement, without further information exceeding that contained in the factive statement. If this is not possible one can only make an inference of the following type.

- (I2) FACT is the cause of ... (is causally responsible for ... )  
 Therefore,  $(\exists e)$  ( $e$  is an event and  $e$  is the cause of ... ).

In such cases, one can not derive from the expression FACT any explicit expression designating the event underlying the causally responsible fact. However, two sorts of restrictions considerably reduce the domain of events which may replace the variable  $e$ . First, in a causal statement, the expression FACT attributes a property to a particular object or event;<sup>22</sup> in the former case, the cause-event  $e$  must spatially coincide with that object, or at least contain it. Second, the causal statement “FACT is causally responsible for ... ” is true at a given moment, i.e. at the moment at which the causal interaction takes place.<sup>23</sup> The temporal position of the event  $e$  is constrained by the fact that it must at least contain that instant or else coincide with it.

Consider the following statement:

- (7) The fact that Rome is south of Paris is causally responsible for its warmer climate.

The factive expression

- (7a) The fact that Rome is south of Paris

designates a propositional entity whose constitutive predicate is *relational*, more precisely a predicate expressing an *external* relation. A relation is external if its terms can exist without standing in the relation.<sup>24</sup> One can infer from (7) the existence of an event, more precisely of a state<sup>25</sup> – which I take to be a continuous series of events – which is the cause of the warmer climate of Rome. However, (7a) does not contain the information necessary for constructing an eventive expression designating explicitly the efficacious state – or series of events – underlying the fact.

From (7) one can only infer, by virtue of (I2), the following existential generalization:

(8)  $(\exists s)$  ( $s$  is a state and  $s$  is the cause of the warmer climate of Rome).

The climate of Rome can only be due to events taking place at Rome and to causal processes reaching Rome. However, the relational statement (7) contains no reference to such events or to the state consisting in a continuous series of such events.<sup>26</sup>

The situation is similar with causal statements containing *negated* expressions: the causal statement (9), in which the cause is designated by a negative factive expression, entails (by virtue of I2) the statement (10):

(9) The fact that the train was not on time caused some confusion in the station.

(10)  $(\exists e)$  ( $e$  is an event and  $e$  caused some confusion in the station.).

However, it is not possible to extract from

(9a) The fact that the train was not on time

an eventive expression that could be substituted for the variable  $e$  in (10). (9a) does not contain enough information to construct such an expression. The absence of the expected event, i.e. the arrival of the train, does not contain any indication of what has positively caused the confusion. An event that does not take place cannot cause anything. (9) does not tell us whether the confusion has been caused by the travellers' perception of the announcement that the train was expected to be an hour late, or whether it was caused by some travellers' comparison between the expected time of arrival and the time on their watch, or by any other event of that sort. The confusion taken as an event can only have been caused by another event. But (9) does not contain any information about the nature of that event.

What are the general conditions under which the inference

(I1) FACT is the cause of ... (is causally responsible for ... )  
Therefore, EV is the cause of ...

is possible, in the sense that the expression EV can be formally derived from the expression FACT? We have seen that the inference can be drawn in the case of statement (5) but not in the cases of statements (7) and (9). The possibility to draw inference (I1) seems to depend on the nature of the *predicate* contained in the expression FACT. From a statement attributing causal responsibility to a fact it is always possible to infer, with (I2), the *existence* of an event or a state involved in the corresponding eventive causal relation. From a factive causal statement one can always deduce a statement of type (8) or (10), in which the eventive variable  $e$  is in the scope of an existential quantification. However, one can only construct an explicit eventive expression

that can occupy the place of the variable  $e$ , if the constitutive predicate of the expression referring to the causally responsible fact expresses an *intrinsic* property<sup>27</sup> of the cause. Our hypothesis is that the inference scheme (I1) cannot be applied if the constitutive predicate of the expression referring to the causally responsible fact belongs to one of two categories: if the predicate expresses a relational property – as it does in statement (7) – or has a negative form – as in statement (9) – only inference schema (I2) is applicable.

***There is no valid inference from an eventive causal statement to a factive causal statement.***

The converse inference is never logically valid. There is no valid inference from a causal statement in which an eventive expression designates the cause, to a statement in which a factive expression designates the cause. To the extent that a factive statement contains more information about a causal relation than the corresponding eventive statement, the former cannot be deduced from the latter.

The only statement of factive form that can be validly derived is an existential generalization over the causally responsible fact  $f$ :

(I3) EV is the cause of ...

Therefore,  $(\exists f)$  ( $f$  is a fact and  $f$  is causally responsible for ...)

Let us examine the constraints that the content of the expression EV imposes on the possible arguments of the variable  $f$ . This will be easier if we indicate the structure of the fact represented by  $f$ . The general form of a fact,  $Pe$ , is the attribution of a property  $P$  to an event (or object)  $e$ . The eventive statement entails a propositional statement in which the predicate of the expression designating the causally responsible fact is bound by an existential quantifier.

(I3\*) EV is the cause of ...

Therefore,  $(\exists P)$  ( $Pe$ , where  $e$  is the event designated by “EV”, and  $Pe$  is causally responsible for ...)

For example,

(6) Mary’s singing of the song surprised me.

entails

(11) Something about Mary’s singing of the song surprised me.

In a similar way, from the statement:

(12) His awkward theft of the bicycle displeased me.

it can be inferred:

(13) Something about his awkward theft of the bicycle displeased me.

One can infer that there is a fact  $f (= Pe)$  concerning the event designated by “EV”, but the information contained in the eventive expression “EV” is not sufficient to construct an explicit factive expression that can replace the variable  $f$ . This is due to the impossibility to extract the predicate  $P$  from that eventive expression.

With (I3), (12) entails:

(14)  $(\exists f)$  ( $f$  is a fact and  $f$  is causally responsible for the fact that I am dissatisfied),

or with (I3\*),

(15)  $(\exists P)$  ( $P$  is a property of his awkward theft of the bicycle and the fact that his theft is  $P$  is causally responsible for the fact that I am dissatisfied.)

According to a reasoning analogous to the case of Mary’s performance, (12) can be true in a situation in which neither the fact that he steals nor the fact that he steals awkwardly is the aspect of the theft responsible for my dissatisfaction. Suppose that the crucial aspect is the fact that he did not steal a mountain bike as I told him but a street bicycle. Then the only expression of propositional type that can truthfully take the place of the variable  $f$ , is the following:

(16) The fact that he did not steal a mountain bike is causally responsible for the fact that I am dissatisfied.

However,

(16a) The fact that he did not steal a mountain bike

cannot be extracted from

(12a) His awkward theft of the bicycle.

### ***Conclusion***

According to the inference schema (I1), it is possible to derive an eventive causal statement from a factive causal statement, except if the expression designating the causally responsible fact attributes to the cause a relational or

a negative property. In those cases one can only, with inference schema (I2), infer the statement that a cause event *exists*. In a similar way, from an eventive causal statement one can only, with inference schema (I3), infer the factive causal statement that a causally responsible fact *exists*, or, with inference schema (I3\*), the statement that an efficacious property of the cause exists.

## Objections

Davidson (1966; 1967) and Parsons (1990) hold that a fact, being abstract, can be neither cause nor effect. For Parsons, “the idea that what is caused is a proposition, as required in the symbolization [proposed by Dowty], instead of an event, seems counterintuitive” (Parsons 1990, p. 108).

There is certainly a sense in which it is true that facts cannot strictly speaking be considered as causes or effects. But causal predicates can express one or the other of the two aspects of causation that we have distinguished. A fact cannot be a term of the causal relation when it is understood as an extensional relation between particulars. However, a fact can be cause or effect if causation is understood according to our model of causal responsibility. We have seen before that factive causal statements always entail eventive causal statements, though in some cases with an existential quantification. In other words, if a true statement designates a fact as being the cause of something, in the sense of being causally responsible for it, this implies that there is an underlying concrete event which is the cause in the first sense of the word “cause”. However, the factive causal statement contains a richer information about the cause: it does not only designate the particular event that is the cause (this is the content of the eventive statement it entails according to (I1) or (I2)), but it tells us also which of the properties of the cause is efficacious in bringing about the effect. In this sense, although the causally responsible fact itself may be abstract, it contains more information about the “concrete” cause in the sense of the cause-event than an eventive causal statement.

One of the main reasons for which Davidson (1966; 1967) denies that facts can be causes and effects is a famous argument often referred to as the “slingshot”. If the identity of facts is determined according to what the propositions expressing them refer to, then all facts are identical, insofar as one follows Frege (1892b) in considering that the reference of all true propositions is the truth value TRUE. According to this criterion of the identity of facts, there is only one fact. Indeed, if only reference to particulars and truth values are taken into account in the determination of the identity of a fact, “two true sentences express the same [ . . . ] fact if and only if they are interderivable either (i) a priori or (ii) by the replacement of one name by another with the same referent” (Bennett 1988, p. 37),<sup>28</sup> because in this case such an inference or substitution does not change truth value. The substitution can concern a proper name, a pronoun or a definite description. Here is how Davidson (1966, p. 117f.; 1967, p. 152f.)<sup>29</sup> shows that this

conception of facts yields the conclusion that there is only one fact. Let  $S_1$  and  $S_2$  be any arbitrary true statements. Then

The fact that  $[S_1]$

designates, according to condition (i) of the identity criterion for facts just quoted, the same fact as

The fact that: the class of objects identical to themselves is the class of objects such that (they are identical to themselves and  $[S_1]$ )

which designates, according to the transparency condition (ii), the same fact as:

The fact that: the class of objects identical to themselves is the class of objects such that (they are identical to themselves and  $[S_2]$ ),

which designates, again by virtue of condition (i), the same fact as:

The fact that  $[S_2]$ .

In this argument,  $[S_1]$  and  $[S_2]$  stand for arbitrary true statements, so that it leads to the conclusion that "Paris lies on the Seine" designates the same fact as "Nero fiddled" (Davidson's example), in the case these statements are true. Using this conception of facts in the framework of the theory of causal responsibility would lead to the absurd result that all effects have the same cause and all causes the same effect.

However, our thesis that facts can be causes and effects can be rescued by choosing a more fine-grained criterion of the identity of facts, which allows us to distinguish them from each other.<sup>30</sup> It suffices to replace, in the above identity criterion for facts, the expression "name" by the expression "Russellian name" (Bennett 1988, p. 36f.), where a "Russellian" or logically pure name is an expression whose only semantic value is its reference but which lacks descriptive content. If one substitutes in a true statement a definite description containing it, by another, the substitution will alter the designated fact. The identity of the designated fact will remain unaltered only if one substitutes a "Russellian" name for another with the same reference. Horgan (1978) expresses the same idea by saying that, in the first case, the original statement and the statement obtained by substitution "are only materially equivalent and not logically equivalent" (Horgan 1978, p. 34). Two statements that differ only in materially equivalent expressions share their truth value but they may not be intensionally equivalent, because they contain predicates referring to different *properties*; therefore, an identity criterion for facts grounded on the principle that all materially equivalent expressions can be substituted *salva veritate*, yields a concept of fact



according to which only reference to particulars, but not to properties, determines a fact's identity. But this is not the only legitimate concept of fact. In our analysis of causation we shall use a concept of fact according to which the identity of a fact designated by an expression "the fact that *a* is *F*" or "*a*'s being *F*" depends both on the particulars and on the properties to which it makes reference. This allows us to distinguish facts from each other, although they all share their Fregean truth value.

***Propositions stating that an event has happened***

We have shown that there is no purely formal way of deducing, from a statement of the form "EV is the cause of ...", a statement in which the eventive expression EV is replaced by an explicit factive expression. It may be objected that, as Zucchi (1993, p. 19) has noted, one can always infer from statements in which the eventive expression EV occupies a transparent position, a proposition of the type: "EV has occurred", "EV has happened" or "EV has taken place". This is indeed a valid inference form leading from statements containing an eventive expression to factive statements, which can be drawn on formal grounds without taking the content of EV into account. This exceptional possibility is due to the fact "that 'occur' (and 'happen' and 'take place') are special in having no content of their own; they merely assert the culmination of an event identified by other means" (Parsons 1990, p. 136).

However, at a closer look, the inference mentioned by Zucchi turns out to be compatible with our analysis. In a causal context, it is not always possible to substitute a proposition of the type "EV has happened" for an eventive expression EV, without changing the truth value of the statement. For example, the content of

(17) The fact that the breakdown of the enemy army has happened caused great relief.

differs in content from

(18) The breakdown of the enemy army caused great relief.

(18) may be true in a situation in which, what is responsible for the relief is not the breakdown of the army as such but a particular aspect that is not explicitly named, for example that the breakdown was quick. However, (17) would be false in that situation. This means that the inference from (18) to (17) is not only impossible for stylistic reasons. It is invalid because its conclusion (17) would contain more information than its premise (18).

In a similar way, Ramsey (1927) has argued that the statement of a fact is equivalent to a statement of the existence of an event:

“That Caesar died” is really an existential proposition, asserting the existence of an event of a certain sort, thus resembling “Italy has a King”, which asserts the existence of a man of a certain sort. The event which is of that sort is called the death of Caesar, and should no more be confused with the fact that Caesar died than the King of Italy should be confused with the fact that Italy has a King.

(Ramsey 1927, p. 37)<sup>31</sup>

According to our analysis, what is generally valid is only an entailment but not an equivalence. Causal contexts may bring out the semantic difference between the two types of expression. If the fact  $Fc$  is designated as the cause, this means that the fact that the eventive cause  $c$  has property  $F$  is causally responsible. However, if it is only said that an event  $c$  has happened which is the cause, the causally responsible property is not specified. In other words, the causal statement “FACT is causally responsible for ...” contains more information than the existential statement one can infer from it by virtue of (I2). This difference can be semantically important for the interpretation of causal statements. The statement “the fact that Caesar was dead horrified his friends accompanying him” excludes the interpretation according to which Caesar’s friends were not horrified by the fact that he was dead in itself, but by the cruel execution of the assassination. However, the eventive statement “Caesar’s death horrified his friends accompanying him” *can* be interpreted in this way. Hence there is a possible situation in which the former is false and the latter is true. The fact that the factive and the eventive statements can diverge in truth value shows that Ramsey is wrong in thinking that the expressions designating the cause in one and the other statement are semantically equivalent.

We conclude from the analysis of these examples that one cannot in general substitute, in a causal context, the factive expression “EV has happened” for an eventive expression EV without changing the truth value of the statement.

## Conclusion

We have shown the nomological theory of causation, the traditional account of causation developed by logical empiricism, to be inadequate. It is impossible to directly reduce causation to the instantiation of a law. The nomological theory is motivated by another reductionist project, the project of reducing both nomic regularities and causal relations to what may figure in scientific *explanations*. The fundamental idea of the nomological theory is to assimilate causal explanation and nomological explanation, and the latter to scientific explanation in general. However, these identifications neglect some fundamental distinctions: first, an explanation may bear on a causal relation although its explanans does not designate a cause of which the explanandum designates the effect. The logical structure of a causal expla-

nation is independent of the asymmetry of the causal relation on which it bears. Second, many explanations are not causal at all. We have seen a certain number of types of explanations, for which neither their truth nor their relevance depends on the existence of a causal relation between the explanans and the explanandum.

However, we have found that the instantiation of a law of nature is part of the content of the complex concept of causal responsibility. According to our analysis of this concept, the fact  $Fc$  is causally responsible for the fact  $Ge$  if and only if 1) there is a causal relation  $C_E(c,e)$  between events  $c$  and  $e$ , in the sense of causation analysed in the first chapter, and if 2) there is a law of nature  $L(F, G)$  instantiated by  $c$  and  $e$ . This allows us to analyse the relation of causal responsibility by the following equivalence:

$$C_F(Fc, Ge) \text{ if and only if } C_E(c,e) \wedge Fc \wedge Ge \wedge L(F, G).$$

The relation of causal responsibility depends on laws of nature in two respects: first, a causal relation between events  $c$  and  $e$  presupposes that there is a conservation law applying to the conserved quantities some amount of which is transmitted between  $c$  and  $e$ . (Cf. Chapter 1.) Second, the fact  $Fc$  can only be causally responsible for the fact  $Ge$  if there is a law  $L(F,G)$  instantiated by  $c$  and  $e$ . This account of the dependence of causal responsibility on laws allows us to clarify the links between explanation, laws and causation. Instead of following the nomological theory and reducing these concepts directly to each other, we have found that an explanation is causal only if it bears on a causal relation. The existence of a causal relation is therefore necessary for the truth of a causal explanation. But the explanation can only provide substantial information if it not only designates the cause and the effect as events but also mentions those among their properties which are involved in the causal relation. This leads us to the conclusion that a causal explanation makes reference to a relation of causal responsibility between a fact concerning the cause and a fact concerning the effect.

The two aspects of the concept of causation cannot be reduced one to the other. Both are required for a full understanding of causation: according to the first, causation is a relation between events; according to the second, it is the relation of causal responsibility between facts. Facts and events are both structured entities: a fact is what makes true the attribution of a predicate to an object (or to an event or to several objects or events). An event has only spatio-temporal structure. As particular entities, events have properties. But when a statement makes reference to them as events, these properties do not enter into its semantic evaluation.

With the help of the distinction between positive and negative predication and between the predication of relational and intrinsic properties, we have established conditions under which a factive causal statement entails an eventive statement, without requiring any information not contained in the

factive statement, and the conditions in which only an existentially generalized statement can be validly inferred. In the reverse direction, the major informational content of factive causal statements with respect to the corresponding eventive statements makes it impossible to validly infer from an eventive statement a statement expressing explicitly the causally responsible fact.

Our account gains explanatory strength at the expense of parsimony compared to other accounts which take only facts or only events to be capable of playing the role of cause and effect. We have rejected a famous argument against the inadequacy of facts for the role of causes, by rejecting its premise according to which the identity of a fact is determined only in terms of the reference to particulars and truth values (pages 159–61). Then we have shown that one cannot in general derive a factive causal statement on the basis of an eventive causal statement, by substituting the expression “EV has happened” for the eventive expression EV. This substitution does not conserve truth value if EV occurs within a causal context, as it does in “EV is the cause of” (pages 161–2).

The next chapter takes up the task of clarifying the identity criteria for the properties playing in  $C_F(Fc, Ge)$  the role of the efficacious property  $F$  of the cause and the causally affected property  $G$  of the effect. Furthermore, we shall consider the traditional concept of a “causal law” and enquire whether the laws  $L(F, G)$  constitutive of the relation of causal responsibility should themselves be taken to be causal.

## 6 Efficacious properties and the instantiation of laws

The semantic analysis of causal statements presented in the last chapter has led us to postulate two relations, irreducible to one another – the relation of causal responsibility and the relation of causation between events: the former makes true factive causal statements whereas the latter makes true eventive causal statements. We have seen that the relation of causation between events is more fundamental in the sense of being an essential part of the former, whereas factive causal statements are often more informative and more directly relevant for causal explanations. Relations between events are what makes true eventive causal statements. An event is constituted by the set of all properties which are entirely determined within its spatio-temporal zone. Each event is involved in a network of singular causal processes. But to understand these processes, and to explain and manipulate them, it is necessary to know their nomologically relevant properties. Therefore the immediate object of a causal explanation is a relation between a fact about the cause and a fact about the event, where the former is said to be causally responsible for the latter.

In order to elaborate our notion of causal responsibility, it will be useful to compare our analysis with Davidson's (1967). Recall Davidson's solution to the problem discussed in the preceding chapter, of reconciling two requirements: insofar as they express a relation between particular entities, causal statements must be extensional. However, a certain number of causal statements do not seem to fit in this framework because they seem to violate one criterion of extensionality. Here is how Davidson resolves this conflict: on the one hand, he does justice to the first requirement by postulating that "the relation of causality between events can be expressed (no matter how 'strong' or 'weak' it is) by an ordinary two-place predicate in an ordinary, extensional first order language" (Davidson 1967, p. 161). On the other hand, he takes account of the observation that some causal statements are not extensional by saying that "such sentences tell, or suggest, a causal story. They are, in other words, rudimentary causal explanations" (Davidson 1967, p. 161). Here Davidson seems to join those who reduce causation to explainability or predictability. We have seen Carnap (1966) and Stegmüller (1969) defend this thesis.<sup>1</sup> However, the apparent similarity of their

views with Davidson's hides a radical difference: for Davidson, the aspect of causation that may be reduced to explainability is not identical to the fundamental concept of causation as a relation between singular events. "The 'caused' of the sample sentences in this paragraph is not the 'caused' of straightforward singular causal statements, but is best expressed by the words 'causally explains'" (Davidson 1967, p. 162). Davidson places much weight on a distinction that the reduction of causation to explainability tends to wipe out. He points out that "we must distinguish firmly between causes and the features we hit on for describing them, and hence between the question whether a statement says truly that one event caused another and the further question whether the events are characterized in such a way that we can deduce, or otherwise infer, from laws or other causal lore, that the relation was causal" (Davidson 1967, p. 155). This distinction contains the germ of our distinction between causation and causal responsibility. However, there are at least two fundamental differences between Davidson's and our analysis. First, our analysis of causation between events is grounded on an independent criterion of the identity of events, whereas Davidson's criterion of the identity of events depends in a circular way on the notion of causation. Second and more important, causal responsibility is an objective relation, whose terms are facts containing objectively instantiated properties, and which depends on the instantiation of a law by these properties. By contrast, Davidson's concept of a causal explanation is an epistemic concept: its terms are descriptions, in other words statements that exist only thanks to a cognitive agent seeking an explanation. This does not mean that our analysis is not perfectly compatible with the classical concept of causal explanation, which Popper expresses in these words: "To give a *causal explanation* of an event means to deduce a statement which describes it, using as premises of the deduction one or more *universal laws*, together with certain singular statements, the *initial conditions*" (Popper 1934, 31, transl. p. 59; italics Popper's). Our concept of causal responsibility is intended to capture the ontological ground of causal explanations, in the sense that such relations are what makes causal explanations true. The truth of a causal explanation does not only require the existence of a causal relation between events  $C_E(c, e)$ , but also a causal relation between facts  $C_F(Fc, Ge)$ . The existence of a causal relation between fact  $Fc$  and fact  $Ge$  provides the basis for the possibility to explain the latter in terms of the former. This factive causal link between  $Fc$  and  $Ge$  requires that the events  $c$  and  $e$  instantiate properties  $F$  and  $G$  and that these property instances are part of the instantiation of a law of nature.

## Properties

In this analysis, one point, in particular, seems to be in need of further clarification: we have not yet said which kinds of properties are capable of causal efficacy, in other words which kinds of property can make events interact with each other. To find this out, we need a general criterion for

the identity of properties. Our hypothesis is that only *intrinsic* properties, in other words properties completely determined within the zone occupied by the events possessing them, are immediate sources of efficacy.<sup>2</sup> By contrast, relational, historical and comparative properties are not efficacious although they play an important role in many scientific taxonomies.

Are there “natural” properties, i.e. properties existing in the same sense as the objects and events possessing them? It might seem more prudent to adopt a nominalist stance: to grant existence only to particular entities, and to take account of property talk in terms of different ways of classifying objects and events by means of *predicates*. Take our example of a copper wire. Should we say that its identity is determined by a particular combination of natural properties: being metallic, having reddish colour, being a good conductor of electricity and thermal flow etc.? Or should we rather refrain from postulating any such property instances in the wire and content ourselves with saying that it is appropriate to attribute to it certain predicates, “is metallic”, “is of reddish colour”, etc.? Is it appropriate to say, as we have done before, that the *natural property* of the wire to be a good conductor of electricity is causally efficacious in bringing about the fact that (a certain amount of) electric current is flowing in it? Or should we just say, thereby committing us to a more parsimonious ontology, that there is a nomological *statement* linking the *predicate* “*x* is a good conductor of electricity” to the *predicate* “if a tension of *U* is applied to *x*, then electric current of more than *I* flows through *x*”?

To answer this question, let us look at the way in which properties appear in the analysis of causal responsibility. The relation of causal responsibility between two facts can be analysed by the conjunction

$$(O') C_E(c,e) \wedge Fc \wedge Ge \wedge L(F, G).$$

The properties *F* and *G* participate in the causal relation  $C_F(Fc, Ge)$  as instances not as universals. What is efficacious in event *c*'s bringing about event *e* is an *instance* of *F*. A realist account of particular causal relations commits us only to realism with respect to property instances, not to realism with respect to universals. For our purposes, we may conceive of the particular copper wire as of the complex of all properties instantiated at the space it occupies. What acts as a cause is the complex of all properties instantiated at the spatio-temporal region the copper wire occupies at the moment of the causal relation at issue. Among those properties, we call “natural” those that participate in nomic relations to other properties. One of these properties, which this particular copper wire shares<sup>3</sup> with all other samples of copper, is its electron structure, more precisely the distribution of its electrons in energy bands in a way characteristic for copper. It is this property that determines many aspects of its behaviour, for example its capacity to conduct electric current. Other properties of the wire determine other aspects of its behaviour: the properties of the crystal lattice made up

by its atoms contribute, together with its electron structure, to the wire's thermal conductivity.

The distinction between "causal" and "non-causal"<sup>4</sup> properties is appropriate only with respect to relational properties, which can indeed have or not have causal efficacy.<sup>5</sup> However, there is no question whether an intrinsic, in other words non-relational property, which is entirely determined within the region of space occupied by the object possessing it or within the spatio-temporal zone of the event possessing it, is causal or not. All intrinsic properties are "causal" to the extent that the interactions which they give rise to are the only reason to postulate their existence. One might express this point by saying that all intrinsic properties are dispositional:<sup>6</sup> one attributes a non-relational property to an object only if there are circumstances in which this property is efficacious. To say that an object is water-soluble means that there are circumstances in which certain properties of the object produce phenomena that are characteristic of solubility: put in water, it dissolves. However, as Goodman has noted, many more predicates designate dispositional properties than just those which bear a grammatical mark of dispositionality, such as the suffixes "-ible" and "-able". "Indeed, almost every predicate commonly thought of as describing a lasting objective characteristic of a thing is as much a dispositional predicate as any other" (Goodman 1955, p. 41).

What we have said above suggests the following interpretation of Goodman's qualifier "almost": all intrinsic properties, i.e. all properties completely determined within the space occupied by the object to which they are attributed, are dispositional. This means that the way they determine the causal interactions of their possessor is what constitutes their identity, what is essential to them. The fact that their identity is determined in terms of their influence on causal interactions makes them *natural* properties, whose existence does not depend on language. Even if the primitive properties of physics are not explicitly defined in terms of dispositions, their nature consists in a set of features corresponding to their causal manifestations. In some cases the probability that the property is responsible for interactions may be very weak. The properties of some elementary particles, e.g. neutrinos, are such that they interact only very rarely with their environment. However, if they *never* interacted with anything, if the probability of the property's giving rise to (or modifying) interactions were strictly zero, we would have no reason to think it exists. A hypothetical object, none of whose properties has any power at all of bringing about or influencing interactions, would simply not exist; and if the probability is strictly zero that a hypothetical property *H* of an object is responsible for any interactions of the object, we have no reason to attribute property *H* to the object in the first place. The notion of a totally inert non-relational property lacking causal powers of any sort is incoherent.<sup>7</sup>

Against this conception of the link between a property and the causal relations it determines, Holt (1976) offers the following argument. He



begins by pointing out that it is equivalent to say that an object possesses a property and to say that this property is determined independently of the causal relations between the object and other objects. Then he considers several conceptions of the *link* between intrinsic properties and the causal relations they determine, ending up with a dilemma. According to one conception, the link is contingent. Holt rejects it because it is impossible to conceive of an object independently of its causal interactions.

“If our essential conception of matter does not involve causal properties, then we clearly should be able to conceive of a chair with all of these removed. [ . . . ] Suppose however that it were suggested to me that there is now a chair present before me, even though owing to it having no effects upon light I cannot see it, owing to its failure to affect material bodies I cannot feel it and can pass through it without resistance, and so on. My reaction clearly would not be that there might indeed be a chair present: rather, I should point out that what *counts* for me as a chair’s being present is the ways in which, so to speak, it makes its presence felt. I would feel that there is no sense in speaking of a chair through which I can walk, which I cannot see, and so on” (Holt 1976, p. 21).

He concludes that the link between intrinsic properties and their causal impact cannot be contingent after all. But he finds it equally inappropriate to conceive of the link between an intrinsic property and the causal relations it gives rise to as necessary. Holt argues against the idea that intrinsic properties are *essentially* properties that manifest themselves in certain effects, because it does not do justice to the substantial nature of objects. It makes us, as he says, “lose the substance of the world” (Holt 1976, p. 23). The reason is this: “A substance is surely something which has its own being – a being independent of the existence of other things – and this does not appear to be true of something whose nature is essentially causal, since the very notion of a cause involves implicit reference to other things” (Holt 1976, p. 25). Another way to make this point is that it leads into an infinite regress (or a vicious circle) to conceive of objects merely as of “bundles of causes” (Holt 1976, p. 23), for such a conception has the unacceptable consequence that “every object seems to be dependent for what it is on other things which are themselves equally insubstantial” (Holt 1976, p. 23).

Two considerations may weaken Holt’s objection. First, we may characterize more precisely the link between intrinsic properties and their causal efficacy: they must be *capable* of bringing about or influencing causal relations by virtue of their nomic links to other properties. This is compatible with the possibility that, for a given property of a given object, this capacity never becomes actualized. A red piece of glass has the causal capacity to “transform” white light passing through it into red light by filtering out certain wavelengths. If such a piece of glass is produced and then immediately destroyed before ever having been exposed to light, its capacity to transform white light into red is never actualized. However, according to

our conception of properties, this does not mean that the glass doesn't have the property. The mere capacity to exercise its power is sufficient for its existence.

The second consideration is more important. Holt's objection is that a causal conception of intrinsic properties is incompatible with the idea that things have a substance which makes their existence independent of other things. However our account suggests replacing the concept of substance by a causal conception of the persistence of objects through time. The point is not simply to have events, instead of objects, play the role of substances. Holt is right that a causal conception of properties is incompatible with a substantial conception of objects. Our proposal is to conceive events not as substances but as bundles of intrinsic properties. Part of these properties are conserved by law, and can so serve as a basis for the persistence of objects through time.

Saying that all intrinsic properties and only they, are capable of causal efficacy means taking a stand in the debate about scientific realism. Their capacity of being causally efficacious makes them natural properties, in the sense this term takes in that debate. We shall see how the concept of causal responsibility allows us to link the concept of natural property to that of law of nature. Since Goodman's *Fact, Fiction, and Forecast* (1955), it is well known that the issue of natural properties is part of a set of problems inextricably bound up with each other. There is no hope of solving the problem of justifying the existence of a class of privileged predicates designating natural properties in isolation: in particular, the problem of justifying the distinction between predicates designating natural properties such as "being blue" and predicates arbitrarily made up such as "being grue"<sup>8</sup> is intimately linked to the problem of induction. The possibility of making predictions depends on the "projectibility" of the predicates one employs. Only such inductive inferences about yet unobserved cases, in particular in the future, can be justified that are made with the use of projectible predicates. Moreover, these distinctions are inseparable from the distinction between nomological statements and statements expressing only accidental regularities. Are projectible those predicates that participate in laws, fundamental or not, and which designate natural properties.

One may try to get round this whole bunch of problems by taking into consideration only the sciences themselves; in particular, sciences such as fundamental physics that have reached the stage of axiomatization yield clear criteria for classifying predicates into those appearing among its axioms and theorems (law statements) and those not appearing there. Thanks to this strategy, one can replace the category of "natural properties" by a category of predicates that contains all predicates appearing in statements of scientific laws, according to the actual state of science. However, given that the actual state of science is bound to change, it seems ill advised to take actual science as a criterion for a distinction between properties, as they are in themselves.

In order to remedy the arbitrariness of choosing the present state of science for construing the category of “natural predicates” (by taking to designate natural properties, all and only predicates figuring in the scientific axioms and theorems presently accepted), one can use instead the idea of a hypothetical ideal or achieved science. This leads to the thesis that the predicates designating natural properties are those that appear in the axioms and fundamental laws of an ideally achieved science. The state of achievement can be judged by such criteria as information content and simplicity. In case there turn out to be several scientific axiomatic systems that fulfil the combined criterion of maximal informativity and simplicity to the same degree, the predicates designating natural properties are those that figure in the axioms and theorems of all these systems. In fact, this account uses a criterion for what it is to be a law in the context of the analysis of natural properties. To establish a link between these two problems, it suffices to make the plausible hypothesis – to which we shall come back later – that natural properties are designated by predicates figuring in true law statements.

The idea to identify laws of nature with the deductive consequences of the axioms common to all theories that may count as achieving science is due to John Stuart Mill (1843). It has been defended by Frank Ramsey in 1928 and more recently by David Lewis.<sup>9</sup> This is how Ramsey states the central idea of this account – at a time at which he has himself abandoned it: “I [...] put up a [...] theory by which causal laws were consequences of those propositions which we should take as axioms if we knew everything and organized it as simply as possible in a deductive system” (Ramsey 1929, p. 150). In Lewis’ terms, a law of nature is any contingent<sup>10</sup> generalization that “appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength” (Lewis 1973, p. 73). However, this account faces some important difficulties most of which have been spelled out by Armstrong (1983, p. 66ff.) and van Fraassen (1989, p. 48ff.). These difficulties are linked to the nominalist framework of the Ramsey-Lewis account, in which laws are taken to be certain *statements*. We have seen that the realist conception is on the contrary based on the distinction between the law itself and the statements expressing it.

The first difficulty is that it seems inevitable that a subjective element participates in establishing the balance between the two requirements of maximal simplicity and maximal information content. There is no objective justification for privileging either a rationalist point of view giving more weight to simplicity or an empiricist point of view giving more weight to information content. It follows that there may well not be any unique ideal theory. Drawing a consequence from an argument of Armstrong’s against the Ramsey-Lewis account of laws, if this turns out to be the case for ideal science, there would be no unique set of privileged “natural” predicates either.

The second difficulty for the Ramsey-Lewis approach is that it has no means to exclude “pseudo-uniformities” that appear as a consequence of

the introduction of artificial predicates such as “being grue”. Given that it is our aim to provide a criterion for the distinction between predicates designating natural kinds and pathological predicates such as “being grue”, we cannot use that distinction for selecting those regularities which an ideal theory must take into account and on the basis of which one measures its simplicity and information content. If it turns out that the introduction of certain predicates intuitively appearing as artificial such as “being grue”, allows us to simplify the theory while maintaining its information content, the Ramsey-Lewis approach leads to the conclusion that the pathological predicate designates a natural kind after all.<sup>11</sup>

This problem throws doubt on the alleged capacity of achieved science to select those predicates that designate natural properties and natural kinds. In the framework of the Ramsey-Lewis approach, “we can have *no* reason to expect that science will tend to develop such a ‘correct’ language” (van Fraassen 1989, p. 54; emphasis van Fraassen’s), i.e. a language such that all and only those of its predicates that figure in axioms and theorems designate natural properties. In the framework of the “ideal science approach”, as we may call the Ramsey-Lewis account of laws, the only criterion for natural predicates is to figure in the axioms and theorems of the ultimate theory. However, we may doubt that criteria evaluating a theory can establish that its predicates make reference to properties and kinds that are natural in an ontological sense. Why should nature itself obey our standards for what it is to be a good *theory* of nature?

The third difficulty is that “the best deductive system is a mere *de facto* systematization” (Armstrong 1983, p. 69; italics Armstrong’s). Why should laws, conceived as consequences of the axioms of an ideal theory, determine the behaviour of *possible* objects, which are not part of the domain which has been the inductive basis for constructing the theory? Why should such laws allow us to determine the truth value of *counterfactual* conditionals? Why should laws established on the basis of the axiomatization of *actual* regularities be valid in a possible world containing objects that have properties not existing in the actual world, e.g. because it contains species of living systems, or combinations of properties not existing in the actual world?<sup>12</sup>

The fourth problem arises for any theory that grounds truth on criteria in terms of coherence. The Ramsey-Lewis approach excludes the possibility that science ends up with several incompatible axiomatizations obeying the norms of simplicity and information content to the same extent. According to Lewis, a law must “appear [ . . . ] as a theorem (or axiom) . . . in *each* of the true deductive systems that achieves a best combination of simplicity and strength” (Lewis 1973, 73; my emphasis). In the situation just sketched, one would have to draw the conclusion that there are no laws of nature at all. However, the premise doesn’t seem to provide good grounds for this conclusion. A fact about the mutual incompatibility of theories that may count equally as ideal does not seem to justify the conclusion that there are no laws in nature.

The fifth problem is that the Ramsey-Lewis approach excludes a priori the existence of a law (1) whose statement cannot be deduced from other axioms of the ideal theory and therefore must itself be considered as an axiom but (2) has very few instances in the actual world. By virtue of (1), the statement of such a law diminishes the simplicity of the axiomatic system; by virtue of (2), it does not much enhance its information content. By virtue of its criterion of simplicity, the Ramsey-Lewis approach yields the conclusion that the statement does not express a law. However, a priori considerations of this sort should not suffice to exclude the possibility that there may be such a law, because our knowledge of laws is essentially empirical. The inverse case poses an analogous problem. The criteria of simplicity and maximal information content can lead to wrongly classifying a large-scale coincidence as a law: "If the simplest, strongest system is not very simple and strong, it is likely to lack the resources to eliminate such accidental uniformities from the privileged set" (Armstrong 1983, p. 72) of theorems considered to express laws. Once again the Ramsey-Lewis approach excludes for a priori reasons the possibility that it is a coincidence that ravens are black. However, this seems once again to be an empirical matter: it depends on whether the colour of actual ravens is determined by the genome of the species or by a contingent aspect of their environment.

Finally, it is impossible to evaluate the information content of a theory independently of other theories, for the only way to measure the information content of one theory is by comparing it with other theories.<sup>13</sup> The problem is that one will never detect the lack of information content of a theory with respect to certain facts if there is *no theory* at all that accounts for them. This throws doubt on the idea presupposed by the Ramsey-Lewis account that empirical facts determine the information content of a theory in a unique way.

By virtue of the fundamental distinction between a law and its statement, these difficulties do not arise at an ontological level within realist accounts of laws of nature. Within a realist framework, analogous problems only arise when one tries to answer the *epistemological* question how we can come to *know* the laws. In our conception, the issues of laws and natural properties are very closely linked: A property is real (or, as we shall often say, is a natural property) if and only if it participates in a law of nature. Thus arguments for realism with respect to laws are also arguments for realism with respect to natural properties. Emeralds are green and not grue, or contain instances of green and none of grue, because their green but not their grue colour is a nomic consequence of the properties of the emeralds' crystal structure.

This conception offers no target for Quine's two objections to the reality of properties.<sup>14</sup> His first objection is based on the principle of parsimony. It is legitimate to enrich the ontology by a new kind of entity only in case scientific explanation requires making reference to entities of that kind. This criterion favours realism to the extent that it can escape problems that arise for the best alternative accounts.

Quine's second objection is that the criteria for the identity of properties are not clear enough to justify postulating their existence. However, the realist criterion for the identity of properties in terms of laws is clear. Two predicates,  $P$  and  $P'$  make reference to the same property if and only if their referents are embedded in the same nomic relations to other properties. This is a necessary condition: if  $P$  is nomically linked in a certain way to  $Q$  but not to  $Q'$ , and  $P'$  is nomically linked in a certain way to  $Q'$  but not to  $Q$ , then  $P$  and  $P'$  are different properties. The property to have a mass equivalent to 511 keV, e.g., is nomically linked to gravitational attraction but not to electrostatic attraction. However, the property to have elementary electric charge is nomically linked to electrostatic attraction but not to gravitational attraction. These properties are different because they participate in different laws.

The condition of identity of properties is also sufficient: at the beginning of this chapter, we have noted that it is unjustified to postulate absolutely inert properties, in other words properties that are not nomically linked to any property at all. Such an absolutely inert property would not only never actually be efficacious but it would be incapable of being efficacious by virtue of the laws of nature. But if all properties for which one has grounds to think they exist are integrated in the network of nomic relations, one can use their position in that network as a criterion of their identity. What makes the property of bearing elementary electric charge the property it is are its nomic relations to the electric field it creates, to the properties of attracting or repelling positive or negative electric charges, the law according to which it resists decomposition into smaller electric charges and the set of all other laws – some of which may be unknown – in which it participates.

Before we close this section, we should draw attention to a mistake that can easily be made as a consequence of a superficial interpretation of laws of nature that are expressed as numerical identities, such as the law of the classical pendulum:

L(CP) The period  $T$  of a classical pendulum of length  $L$  is  

$$T = 2\pi\sqrt{L/g}$$

It is crucial to distinguish between the determinable property of a pendulum to have a length and the determinate property of a particular pendulum to have a particular length  $L_0$ , to which measurement attributes a numerical value. (Presupposing a fixed scale of measurement,  $L_0$  can also be taken to designate this numerical value.) L(CP) states the existence of a lawful relation between the numerical values taken by these determinate properties of each particular pendulum. This numerical identity does of course not mean that either the determinable or the determinate properties of having a length and having a period are themselves identical. In other words one must not confuse the idea that the identity of properties is determined by its nomic relations to other properties, with the idea that it suffices for two properties

to be linked by a law of nature to be identical to each other.<sup>15</sup> The law *links properties by a relation between their numerical values without making them identical*.

Realism with respect to laws provides a strong basis for realism with respect to properties: in this framework, the distinction between a predicate and the property it designates is as fundamental as the distinction between a nomic statement and the law it expresses, in the framework of realism with respect to laws. Our criterion of the identity of natural properties makes use of the fact that all natural properties are part of a network of nomic relations. According to this “nomological” criterion of the identity of properties, two properties are identical if and only if they share all their nomic relations to other properties. Having distinguished between a predicate and the property it designates, it is clear that we must also distinguish between a property and the extension of the predicate expressing it.

First, different predicates can designate the same property. For example the colours designated by the following expressions:

- (1) The colour of the double line D of sodium.
- (2) The colour of light of wavelength 589 nm.

are identical although the expressions (1) and (2) have different senses. Not only are the properties designated by (1) and (2) exemplified by the same objects or events, but there is no law that brings about “one” without bringing about the “other”; therefore, no *possible* object or event could possess “one” without possessing the “other”. The predicates share their extension on nomological grounds. They have it in common in all nomologically possible worlds, i.e. in all worlds sharing their laws with the actual world. According to our criterion, this is equivalent to saying that (1) makes reference to *the same* property as (2).

Second, there are predicates sharing their extension although the properties they refer to are not identical. The predicates “is a material object of mass equivalent to 511 keV” and “bears elementary negative electric charge” may serve as an example.

With respect to the causal role of the properties of events, we have made the hypothesis that only *intrinsic* properties of an event are capable of causal efficacy, in other words properties entirely determined within the spatio-temporal zone occupied by the event. We have also restricted the set of properties capable of modifying causal interactions by requiring that such properties must be natural, which means that they must participate in at least one law of nature.

It is important not to confuse the categories of “intrinsic” and “natural” properties with that of essential properties. The first and second, but not the third category are absolute in the sense of being independent of theoretical context and pragmatic context of explanation. I use the first two categories

in the context of an absolute classification of properties, and take “essential” to be a category relative to the context of explanation.<sup>16</sup>

### **Are laws of evolution causal laws?**

When one states that one fact is causally responsible for another fact, the truth of such a statement  $C_F(Fc, Ge)$  requires that a law of nature links the instance of  $F$  to the instance of  $G$ . However, this law need not be strict. True, strict laws of conservation are required for the existence of a causal relation between the events  $c$  and  $e$ . But the law  $L(F, G)$  linking the properties  $F$  and  $G$  constitutive of the facts  $Fc$  and  $Ge$  is in general not a conservation law and need not be strict. It is sufficient that it is instantiated in the circumstances of  $c$  and  $e$ . What makes the relation  $C_F(Fc, Ge)$  *causal* is not the instantiation of the law  $L(F, G)$  but the transference of an amount of a conserved quantity between the events  $c$  and  $e$ . The law  $L(F, G)$  determines only the *type* of the causal relation, transmission being responsible for its being causal. Hence, we do not need the hypothesis that there are two types of laws, causal laws and non-causal laws.<sup>17</sup> This allows us to avoid a certain number of problems encountered by accounts trying to reduce causation to a particular type of causal laws.

Before we turn to showing this, it may be useful to note that the analogous distinctions between “causal” and “non-causal facts”, “causal” and “non-causal states of affairs”<sup>18</sup> have a derived status, because they presuppose the more fundamental distinction between causal and non-causal laws.<sup>19</sup> The state of affairs expressed by a true proposition is a fact. It is only if the proposition is true and contains a relational predicate that it makes sense to ask whether the relation it designates is causal or not. For every causal relation, there are “causal facts” in the sense that there are propositions whose truth depends on this causal relation. If a statement of the form  $C_F(Fc, Ge)$  is true, it expresses a causal fact. Such a causal fact presupposes that the relation between the underlying events  $c$  and  $e$  be causal. However, it makes no sense to distinguish between causal and non-causal with respect to facts expressed by propositions whose predicates are not relational, in other words which have only one argument.

We must distinguish the thesis that there are two types of laws, causal and non-causal, from the thesis that there are two types of singular world lines, which may be governed by those laws: one type containing the causal processes and the other containing those non-causal world lines called “pseudo-processes”.<sup>20</sup> It is at the level of singular relations between events that the distinction between causal and non-causal has its fundamental sense. But we have shown that one can explain the origin of the difference at that level without making the hypothesis that there are two sorts of *laws*. Rather than explaining where the difference between causal and non-causal relations comes from, introducing a distinction between two types of laws



only pushes the distinction between causal and non-causal one step further, without making it any clearer.

One might try to justify the distinction between causal and non-causal laws by assimilating it to the distinction between laws of evolution and laws of association.<sup>21</sup> But the distinction between laws of association and laws of evolution is in fact independent of the distinction between causal and non-causal at the level of the situations in which these laws are instantiated. To show this, it is sufficient to show that a law of association can play the role of the law  $L(F,G)$  linking the properties appearing in two facts related by causal responsibility, and that laws of evolution do not guarantee that all processes instantiating them are causal. We have already presented an example of the second type in which a pseudo-process, in other words a non-causal process, is nevertheless subject to a law of evolution: the case of the propagation of the phase of a wave (cf. Chapter 1, pp. 60–1).

The so-called ideal gas law

$$L(IG) PV = nRT$$

is a paradigmatic law of association, bearing on the relation between the numerical values taken by three determinate properties of a gas: the pressure  $P$ , the volume  $V$  and the temperature  $T$ ,  $n$  representing the quantity of gas in moles, and  $R$  being a constant factor.

Insofar as  $L(IG)$  is a law of association, the hypothesis we are about to criticize, which identifies causal laws with laws of evolution, makes us expect that it can ground non-causal explanations such as the following:

- (5)  $P(g,t)$  because  $T(g,t)$ ,  
where  $t$  is an instant of time,  
 $g$  is a sample of a gas that can be considered as “ideal”,  
 $P(g,t)$  is the pressure of sample  $g$  at time  $t$ , and  
 $T(g,t)$  is the temperature of sample  $g$  at time  $t$ .

As we have seen before (in Chapter 1, pages 48–50), the explanans in (5) does not express a fact that is causally responsible for the fact expressed by the explanandum. Otherwise, the transitivity of the causal relation and the symmetry of the law  $L(IG)$  with respect to  $P$  and  $T$  would lead to the conclusion that  $P(g,t)$ , or for that matter  $T(g,t)$ , is causally responsible for itself.  $L(IG)$  grounds the explanation according to which, at time  $t$ ,  $T(g,t)$  because  $P(g,t)$ , as much as it grounds (5); therefore, if “because” here makes reference to the relation of causal responsibility, one gets the unacceptable result that  $T(g,t)$  is causally responsible for  $T(g,t)$ .

However, a law of association such as the ideal gas law can also cover relations of causal responsibility between facts bearing on different events. This shows that *it is not the law that is causal or non-causal but the parti-*

*cular relation instantiating it.* The observations used to test experimentally whether the law is instantiated in a particular sample  $g$  of gas are a case in point. To check whether  $g$  may be considered as an ideal gas, one may proceed in the following way. One holds fixed one of the three variables appearing in the law, manipulates the second and observes the value taken by the third. One fixes, e.g. the volume  $V$  with the help of a rigid container, tight with respect to matter but permeable for heat, and sets it at a certain temperature  $T$  by providing an appropriate quantity of heat. In this situation the following factual causal statement will be true:

- (6) The fact that one provides the sample  $g$  of gas with energy in the form of a quantity  $\Delta Q$  of heat sufficient to raise its temperature by  $\Delta T$  is causally responsible for the increase in its pressure by  $\Delta P = \Delta T \times nR/V$ .

The truth of statement (6) depends, among other things, on the instantiation of the law L(IG), and thus on sample  $g$ 's being sufficiently ideal. To say that  $g$  instantiates L(IG) at  $t$  means that

- (7)  $P(g,t) \times V(g,t) = nRT(g,t)$ .

The numerical values of the relevant properties of  $g$  are so related because the law L(IG) is valid and because  $g$  satisfies, at  $t$ , its conditions of applicability. The law makes it the case that these properties are coexemplified; however, statement (7) itself is not causal, simply because it does not make reference to two events but only to one.

(6) describes a process extending from the cause, the flow of heat through the walls of the container, to the effect, the new state of equilibrium of the gas once the new temperature  $T$  has been reached. The following statement shows that one can make reference to the same causal process without explicitly mentioning the ideal gas law.

- (8) The fact that one provides the gas with a quantity  $\Delta Q$  of energy in the form of heat is causally responsible for the fact that its temperature raises by  $\Delta T$ .

In fact, (6) is equivalent to the conjunction of (8) and (7), which states that the law L(IG) applies to  $g$ . That a law  $L$  is applicable to a particular sample  $g$  may be implicit in a causal statement such as (6) or stated explicitly in the form (7). It follows that the dependence of the truth of any of the statements (7), (6) and (8) with respect to the validity and applicability of the law L(IG) is independent of the difference between causal statements and non-causal statements. On one hand, (6) and (8) are causal statements, but (7) is not; on the other hand, (6) and (7) but not (8) depend on L(IG). Therefore, the dependence with respect to the validity and applicability of this law cannot provide a criterion of causality. (8) is causal without depending on

the instantiation of L(IG) and (7) depends on the instantiation of L(IG) without being causal. Furthermore, although the dependence of the truth of a statement on the instantiation of a law does not by itself make it causal (as (7) shows), the fact that the truth of a statement depends on a non-causal law such as L(IG) neither prevents it from being causal: Indeed, the relation stated by (6) is causal although the law L(IG) on which the truth of (6) depends, is a law of association.

According to the criterion we have developed in the first chapter (in particular on pages 26–57), causally related events must be spatially and temporally separate and effect events must be located in the future with respect to cause events. This criterion allows us to exclude (7) from the category of causal statements for the simple reason that the expressions it contains make reference to a single event, whereas it leaves (6) and (8) as candidates for the status of causal statements. That they are really causal may be established in terms of our reduction of causation: a quantity of heat  $\Delta Q$ , which is a form of energy, is transferred between cause and effect.

It may be objected that (6) does not make reference to events exemplifying properties contained in L(IG), insofar as the crucial property of the cause is the arrival, in the gas, of a certain quantity of energy  $\Delta Q$ , which is a property not mentioned in L(IG). This objection gives us one more occasion to show the superiority of a theory in terms of transference over the nomological theory of causation. What is decisive for the existence of a causal relation in this situation is the fact that the quantity of energy  $\Delta Q$  is transferred to the gas. Once it has reached the gas it changes its temperature  $T$ , which is one of the properties contained in L(IG), whereas the effect referred to by (6) involves the gas's pressure  $P$ , another property contained in L(IG). The effect event to which both (6) and (8) make reference, is the gas in its new equilibrium state after the arrival of  $\Delta Q$ . (6) and (8) make reference to this effect with the help of two of its properties that are related by L(IG). This difference is important but it is independent of the fact that both statements make reference to a causal relation between the same two events. The underlying eventive causal relation is independent of the law L(IG) and depends only on the law of the conservation of energy. (6) and (8) express two relations of causal responsibility involving the same causal relation at the level of events. In (6), the law L(IG) establishes a link between a property of the cause and a property of the effect, whereas the law relevant for (8) links the supply of heat to the increase in temperature.

By virtue of the fact that L(IG) is clearly a law of association, we have found a situation in which a law of association plays the role of the law grounding a relation of causal responsibility. As we have already shown that there are laws of evolution covering non-causal pseudo-processes, such as the propagation of the phase of a wave, it follows that the distinction between laws of evolution and laws of association cannot ground the distinction between causal and non-causal relations between facts. This distinction

cannot therefore ground the distinction between causal and non-causal laws either.

### **May causation be reduced to a nomic relation of increase in probability?**

Michael Tooley (1987) has made another proposal for grounding the distinction between causal relations and non-causal nomic relations on a distinction between two types of laws. He links this idea to another important hypothesis according to which causation is reducible to a relation in which the occurrence of an event of type *F* enhances the probability of occurrence of another event of type *G*. Tooley's justification of the idea that there are causal laws is worth being closely examined, if only because it gives us the occasion to evaluate the widespread idea that causation can be reduced to a relation of increase in probability.

Tooley begins by distinguishing two causal concepts one of which operates at the level of laws and the other at the level of singular causal relations. He introduces both a "relation of direct causal necessitation" (Tooley 1987, p. 272) (symbolized by " $\rightarrow$ "), which is a second-order relation between universals (properties), and a "special, first-order relation" (Tooley 1987, p. 274) (symbolized by "*C*") between states of affairs consisting of objects exemplifying properties.<sup>22</sup> The former category is particularly relevant for us because Tooley identifies it with the category of "causal laws".

The account he constructs on the basis of the nomic relation represented by " $\rightarrow$ " seems to be inappropriate for two reasons. The first is that Tooley's thesis that causation can be reduced to "transmission of probabilities"<sup>23</sup> confuses one of the observational indicators,<sup>24</sup> which allow an observer to gain knowledge of causation, with causation itself. In short, it confuses an ontological with an epistemological issue. My second objection is that Tooley's theory does not succeed in explaining the asymmetry of causation, which distinguishes it from lawful relations in general. If these objections are justified, Tooley's implicit definition of the concept of causation, which he calls "direct causal necessitation", by the method of Ramsey-Lewis, does not give this concept a well-determined meaning.

As I have said, Tooley conceives of causation in terms of "transmission of probabilities". More precisely, his account contains this theorem (Tooley 1987, p. 277):

Direct Causal Necessitation is Related to Increase in Probability.

The justification of this theorem depends on the fallacy of drawing ontological consequences from the epistemic use of the observation of frequencies (conditional probabilities), *identifying* causation with the distribution of conditional probabilities it brings about in ordinary circumstances. Insofar as Tooley identifies the indicator with the phenomenon it usually indicates,

his theory is vulnerable to counterexamples in which the indicator, i.e. the distribution of frequencies, is present without being brought about by what it usually indicates, i.e. a causal relation.

Before considering such counterexamples, we must look for an adequate interpretation of the probabilities that play a central role in Tooley's account. Although he does not explicitly state what he means by "probability", the most faithful interpretation of Tooley's analysis appears to be in terms of the epistemic, or subjective, interpretation of probability. Thus, he takes the expression " $Prob(Px, E) = k$ " to mean: "The logical probability that  $x$  has property  $P$ , given only evidence  $E$ , is equal to  $k$ " (Tooley 1987, p. 256). If a probability is said to depend on evidence and to be logical, we may safely conclude that it is an epistemic probability. It is attributed on the basis of reasoning from general and particular premises about the situation in which one considers the objects  $x$ . The particular premises consist in evidence obtained by observation. Tooley's probability must be interpreted as epistemic insofar as it is defined in terms of rational judgment and indirectly in terms of a subject's knowledge. This interpretation becomes even more plausible when Tooley takes the unconditional probability that an object  $x$  has property  $P$  to be an "a priori probability" (Tooley 1987, p. 257). Such a subjective interpretation of probability is incompatible with the project we share with Tooley, of constructing a realist theory of causation. However, I shall not pursue this objection any further because his subjective interpretation of probability does not seem to be essential to Tooley's account. I shall therefore from now on adopt a charitable reading and interpret his probabilities in terms of frequencies.

The example Tooley uses to motivate his account of causation in terms of probabilities may help us understand his reasoning. Consider a fair coin and a fair die, such that, when they are thrown, the probability that the coin falls on heads is  $1/2$  and that the die falls on 6 is  $1/6$ . Tooley now proposes a thought-experiment. He imagines a very particular place in which one can observe a modification of these probabilities if the two objects are thrown at the same time and come to rest at the same time, so that, "in that sort of room, [ . . . ] the coin comes up heads *when and only when* the die comes up six" (Tooley 1987, p. 251; emphasis Tooley's). If this is all the information one possesses, there is, as Tooley observes, no reason to favour the hypothesis that it is the die, together with the particular properties of the place, that causes the behaviour of the coin, over the hypothesis that it is the coin that exerts a causal influence on the die. But Tooley overlooks that the lack of information of an *observer* of the phenomenon does not entail that there is no objective fact of the matter as to which of the two objects really exercises causal influence on the other.

Later, Tooley describes a way of finding out which of two causally related events exerts causal influence on the other, without going beyond the observation of frequencies. If it is not only the case that the events "heads" and "six" systematically occur together, but if these events occur in exactly

one out of two trials, it is reasonable to suppose that the cause lies in the coin, involving a property exemplified when the coin falls on heads, and that the behaviour of the die is the effect. If one observes on the contrary that the correlated events occur in exactly one out of *six* trials, it is reasonable to suppose that the cause lies in the die, involving a property exemplified when the die falls on six. In this case, the observation of these particular frequencies gives the observer indeed a good reason to conclude which of the two objects possesses the causally efficacious property influencing the other. But this is perfectly compatible with the further hypothesis that the deviation of the observed frequencies from the ordinary frequencies for each event ( $1/2$  for “heads” and  $1/6$  for “six”) is only an indicator whose reliability should not make us confuse it with what it indicates. On the contrary, it is reasonable to suppose that the correlation, which would otherwise be mysterious, is in fact due to an underlying causal process that is not manifest to the naked eye. In this case, it is not really the probabilities themselves that are “transmitted”<sup>25</sup> from one type of event to the other. Rather, the causal influence is transmitted by a real though unknown process flowing from one object to the other. Rather than accepting that the dependency of the probability of occurrence of an event with respect to the probability of occurrence of another event is primitive and inexplicable, it seems more reasonable to make the hypothesis that this dependency is the consequence of a process that propagates locally from one event to the other. But this means that the dependency of one probability on the other is not identical to the causal relation between the events; rather, it is something that *indicates* the existence of such a causal relation. Indeed, when Tooley presents this thought-experiment, he gives the observed probabilities a purely epistemic role: their distribution gives the observer reasons to believe that two series of events are causally related and that one contains the causes and the other the effects.<sup>26</sup> Even if we interpret Tooley’s concept of probability in terms of frequencies rather than subjectively, in terms of a subject’s degree of confidence in the truth of a proposition, his reasoning suffers from a confusion between ontology and epistemology. From the fact that the observed probabilities may be good *indicators* of causal influences, and in particular of their directions, he draws the illegitimate ontological conclusion that these distributions of frequencies *are* the causal relations.<sup>27</sup>

Still worse for the probabilistic account, nothing guarantees that a given causal relation manifests itself by augmenting the probability of its effect. Indeed, judgments of increase in probability presuppose a comparison between the actual world in which there is a causal law  $P \rightarrow Q$  with possible worlds in which there is no such law. The fundamental idea of the probabilistic account is that the probability of the effect is higher when the causal law exists, i.e. in the actual world, than it is in other possible worlds where the law does not exist. Tooley allows that another law  $R \rightarrow Q$  may coexist in the actual world alongside with the law  $P \rightarrow Q$ ; it is even essential

for his account that such laws may exist, insofar as the asymmetry between cause and effect that is implicit in the postulates of Tooley's theory entails that  $Q$  can have other causes than  $P$ . Indeed, in the framework of this theory, to assume that other types of events may cause  $Q$  is equivalent to assuming the existence of causal laws  $R \rightarrow Q$ . But then nothing allows us to exclude the possibility that the closest possible world from ours, in which there is no law  $P \rightarrow Q$ , is a world containing another law  $R \rightarrow Q$ , such that this law makes  $Q$  *more* probable than it is in the actual world, whereas the probability of (the occurrence of events of type)  $P$  is the same in both worlds. When one compares actual probabilities with the probabilities of certain events in a possible world defined as being "the closest to our world among those worlds that do not contain a given law", the constraints are insufficient to exclude that the probabilities are inverted: this is the case if the probability that the "effect" property be instantiated is in fact *greater without* the law in question, compared to its probability in the actual world where the law is present. After all, Tooley cannot exclude the possibility that the presence of a law  $P \rightarrow Q$  *diminishes* the probability of events of type  $Q$ , compared to the probability of  $Q$  in a world without the law  $P \rightarrow Q$ . However, such a possibility is incompatible with a theory such as Tooley's, which identifies causation with the *increase in* the probability of the effect, given the occurrence of the cause.

My second objection against Tooley's account is that it does not account for a fundamental difference between causal relations and lawful relations in general: it does not account for the fact that the relation between cause and effect is asymmetrical whereas many laws of nature, in particular in physics, have the form of equations stating perfectly symmetrical relations of dependency between the numerical values taken by certain measurable quantities. The explanation of the origin of the asymmetry of the cause-effect relation, on the background of the symmetry of laws of nature, is a central task for the theory of causation. As I shall try to show, Tooley's attempt to ground the asymmetry of causation on the logical structure of causal laws is ultimately unsuccessful. The problem does not lie in the attribution of an asymmetrical logical form to laws in general, which Tooley expresses in this way: "Its being the case that the property-universals  $P$  and  $Q$  stand in the relation of nomic necessitation [ . . . ] logically entails its being the case that for all  $x$ , if  $x$  has property  $P$ , then  $x$  has property  $Q$ " (Tooley 1987, p. 79).<sup>28</sup> The main difficulty lies in justifying the existence of a class of *antisymmetrical* laws, which are supposed to make it impossible that, if  $P$  nomically implies  $Q$ ,  $Q$  also nomically implies  $P$ . In Tooley's account, a law can be considered as causal only if it is antisymmetrical: the fact that  $P$  is subordinated to  $Q$  in a causal law can only make  $P$  the cause of  $Q$  if that law is incompatible with the existence of a converse law by which  $Q$  is subordinated to  $P$ .

The "direct causal necessitation  $P \rightarrow Q$ ", which is Tooley's model for causal laws, obeys to stronger constraints than nomic necessitation as such;

indeed the former entails the latter. This is expressed by postulate (M<sub>6</sub>), which says that causal laws are laws (Tooley 1987, p. 260):

(M<sub>6</sub>): If  $P \rightarrow Q$ , then it is a law that, for all  $x$ , if  $Px$ , then  $Qx$ .

Let us examine how Tooley tries to establish that causal laws are anti-symmetrical. This thesis is the content of his “Theorem 9” (Tooley 1987, p. 278):

Theorem 9: Direct Causal Necessitation is Necessarily Antisymmetric.

There are two problems with Tooley’s argument for Theorem 9. First, one of the premises of its derivation is incompatible with other postulates of the theory. According to this premise, it is possible to find situations  $S$  in which “ $\text{Prob}(Qx, Px \ \& \ S) \neq 1$ ” (Tooley 1987, p. 279). However, acknowledging the possibility that  $\text{Prob}(Qx, Px \ \& \ S)$  be smaller than one is incompatible with the fact that  $P \rightarrow Q$  is a law (postulate M<sub>6</sub>), for this entails, by virtue of Tooley’s definition of a law, the truth of the universal implication ( $x$ ) ( $Px \rightarrow Qx$ ). This conflict can only be resolved by weakening the strength of the universal implication. Admitting – what Tooley refuses to do – that the situation  $S$  may contain laws involving properties  $P$  and  $Q$  that are apt to bring about *exceptions* to the law  $P \rightarrow Q$ , the law  $P \rightarrow Q$  entails only that “outside exceptional situations, ( $x$ ) ( $Px \rightarrow Qx$ )”. In this case, it would indeed be possible that  $\text{Prob}(Qx, Px \ \& \ S)$  be smaller than one. But the possibility of exceptions is incompatible with Tooley’s theory.

The second problem with Tooley’s proof of Theorem 9 is the following. To obtain a contradiction allowing him to conclude that it is impossible that the law  $P \rightarrow Q$  coexists with the law  $Q \rightarrow P$  – which is equivalent to establishing the antisymmetry of the law – Tooley presents the following inference as valid. He says (Tooley 1987, p. 246) that

(M<sub>2</sub>):  $\text{Prob}(Px, P \rightarrow P_1 \ \& \ P_1 \rightarrow P_2 \ \& \ \dots \ \& \ P_n \rightarrow Q \ \& \ S) = \text{Prob}(Px \ \& \ S)$

entails that (cf. Tooley 1987, p. 279)<sup>29</sup>

(a)  $\text{Prob}(Qx, Q \rightarrow P \ \& \ P \rightarrow Q \ \& \ S) = \text{Prob}(Qx, S)$ .

But in fact, (M<sub>2</sub>) entails only:

(b)  $\text{Prob}(Qx, Q \rightarrow P \ \& \ P \rightarrow Q \ \& \ S) = \text{Prob}(Qx, P \rightarrow Q \ \& \ S)$ .

The essential idea behind (M<sub>2</sub>) is that the existence of causal laws by virtue of which one property  $P$  causes another property  $Q$  does not alter the probability of instantiation of property  $P$ , which plays the role of the cause. However, it is crucial – this is the asymmetry expressed in Tooley’s postulates –



that the existence of causal laws by virtue of which one property  $P$  causes another property  $Q$  alters the probability of instantiation of the property  $Q$ , which plays the role of the effect. This means that  $(M_2)$  does not justify what is expressed in (a), i.e. that the probability of  $Qx$  and  $S$  is not affected by the law  $P \rightarrow Q$ , for it is essential to Tooley's account that this law raises the probability of  $Qx$ .

Thus Tooley does not succeed in deriving (a). This invalidates his demonstration, insofar as he derives the following identity from the conjunction of (a) and (b) (Tooley 1987, p. 279):

$$(4) \text{Prob}(Qx, P \rightarrow Q \ \& \ S) = \text{Prob}(Qx, S)$$

This identity is then used as one of the terms of the contradiction required for his argument, which has the form of a reductio and whose other term is (Tooley 1987, p. 279):

$$(3) \text{Prob}(Qx, P \rightarrow Q \ \& \ S) \text{ is greater than } \text{Prob}(Qx, S).$$

Let us sum up. The argument supposed to establish the impossibility that two causal laws  $P \rightarrow Q$  and  $Q \rightarrow P$  coexist has the form of a reductio, which begins by assuming the opposite, i.e. that these laws are both valid, and aims at deriving from this hypothesis the contradictory statements (3) and (4). But at least (4) cannot be derived from Tooley's premises. Furthermore, the statements that are derivable do not contain any contradiction. The argument for the anti-symmetry of causal laws fails.

The importance of the incorrect inferences in the proof of Theorem 9 expressing the antisymmetry of causal laws lies in the fact that these errors are the consequence of some mistaken fundamental choices in Tooley's account. On the one hand, Tooley identifies conditional probabilities with causation itself, instead of interpreting them as indicators which it is reasonable to use for finding out about causal relations, and in particular for finding out which of two causally related events is the cause and which, the effect. On the other hand, Tooley considers causation as a property of a certain category of *laws*, but as he cannot derive the required antisymmetry of the supposed causal laws from his general premises, the postulate of this antisymmetry appears to be ad hoc. This confirms our thesis that the specific asymmetry of the causal relation does not have its origin at the level of laws. Indeed, causation is not a property of a certain class of laws; causation and its asymmetric character exist only at the level of singular relations between particular events.<sup>30</sup>

## Conclusion

This chapter has allowed us to further investigate some important aspects of our analysis of the relation between laws of nature and causality. I have

argued that only *intrinsic* properties of events can be causally efficacious, in other words, make an event causally influence other events. A property is intrinsic if it is entirely determined within the spatio-temporal zone occupied by the event possessing it. This is above all supposed to exclude purely relational properties such as “being at 800 kilometers from the Eiffel Tower” from the class of causally efficacious properties. Another requirement for a property to play a role in causal responsibility, is that it participates in laws of nature, which makes it a “natural” property. According to our “nomological” criterion, the identity of a property is determined by its nomic relations to other properties. Such a criterion fits well to a realist account according to which theories aim at discovering which properties are natural and which exist only by virtue of the introduction of predicates, such as “being grue”. The latter properties are not natural. All natural properties participate in the network of nomic relations: there is no nomically isolated property; and its position in that network determines a property’s identity.

We have shown earlier that the distinction between causal and non-causal processes and relations can be analysed in terms of transference; here we have argued against the thesis that this distinction can be grounded on an analogous distinction at the level of laws. We have seen that the transference of an amount of a conserved quantity depends on a conservation law holding for that quantity. But there is no reason to consider such conservation laws or indeed any other laws as “causal”. This is true in particular for laws of evolution. I have argued that the distinction between laws of association and laws of evolution is independent of the causal or non-causal character of the singular relations in which these laws are instantiated. This has been shown by considering situations where a law of evolution covers a non-causal pseudo-process and situations where a law of association covers a causal process.

Then we have examined a little more closely one particular account which situates the origin of causation at the level of laws: Michael Tooley (1987) defends the idea that it is causal laws that make causes increase the probability of their effects. We have first objected to this idea that it confuses an indicator allowing us to discover causal relations by a heuristic though not infallible strategy, with the causal relations it indicates. The underlying causal relation can well be present even if the typical indicators are absent, and conversely, the typical indicators can be present even if the causal relation that normally brings them about is not. Second, we have objected that Tooley does not succeed in justifying the antisymmetry of causal laws that is necessary to use them as a reduction basis of causation: if the laws supposed to make a relation between events causal are symmetrical then, given the transitivity of the relation of nomic implication, a particular cause can cause itself, which is absurd.

Our own proposal avoids this difficulty: the asymmetry of causation appears at the level of the instantiation of laws in particular events. We have

shown in the first chapter that one can explain the origin of causal asymmetry at the level of events. The logical structure of the laws covering causal relations is independent of the latter's asymmetry. The same ideal gas law covers the increase in pressure produced by an action that increases temperature by adding heat and causal relations in which the roles of cause and effect are inverted, because it is an action that increases pressure (e.g., by reducing volume) that causes the increase in temperature.

One must not confuse the thesis that the laws of nature are not causal with a view that is as old as modern philosophy, and famously defended by Russell in the beginning of the twentieth century: that the notion of causation should be banned from a scientific worldview, and replaced by the notion of law. The fact that the laws described by physical theories are not themselves causal does not prevent their instantiation in particular situations from being part of what makes causal explanations true. The explanation of particular events is a fundamental aim of science; what is required to make such explanations true cannot be excluded from a scientific worldview.

## 7 Causal responsibility and its applications

### Emphasized expressions in causal statements

The analysis of causal statements developed in Chapter 5 can be used to make progress in the debate about causal statements containing emphasized or highlighted expressions like (1) and (2):

- (1) The cause of Socrates' death was that he *drank hemlock* at dawn.
- (2) The cause of Socrates' death was that he drank hemlock *at dawn*.

In written language we may represent such an emphasis by italicizing the emphasized expression. It may well be that the interpretation of this special kind of causal statements is not in itself of paramount importance.<sup>1</sup> Even if that is so, we may take causal statements containing highlighted expressions as a test case for our general hypothesis about the logical structure of ordinary causal statements: our analysis of causal statements will be strengthened if it can be shown to provide us with the means of explaining why a change in emphasis, such as the change from (1) to (2), may affect the truth value of such a statement.<sup>2</sup>

Emphasizing part of the expression describing the cause gives information about the *aspect*<sup>3</sup> of the cause event that has been efficacious in bringing about the effect. (1) is true insofar as the causally efficacious element that has brought about Socrates' death is the chemical composition of the drink he ingested. By contrast, (2) emphasizes the fact that it was at dawn that Socrates drank the hemlock; this makes the statement false because there is no nomic relation between this aspect of the event and Socrates' death. We can represent the structure of (1) by using the concept of causal responsibility introduced in Chapter 5 and represented by the predicate  $C_F$ .

- (1')  $C_F$  (drinks hemlock (Socrates at  $t$ ), dies (Socrates at  $t + \Delta t$ ))  $\wedge$  drinks hemlock at dawn (Socrates at  $t$ ).

More informally, (1') interprets (1) as saying that the fact that Socrates drank hemlock at  $t$  was causally responsible for the fact that Socrates died at  $t + \Delta t$ , and that Socrates' drinking hemlock at  $t$  took place at dawn.

By contrast, (2) has the structure:

(2')  $C_F(\text{drinks at dawn (Socrates at } t), \text{ dies (Socrates at } t + \Delta t)) \wedge \text{drinks hemlock at dawn (Socrates at } t)$ .<sup>4</sup>

More informally, (2') interprets (2) as saying that the fact that at  $t$ , Socrates drank something at dawn was causally responsible for the fact that Socrates died at  $t + \Delta t$ , and that Socrates' drinking hemlock at  $t$  took place at dawn.

(2) is false whereas (1) is true. Where does the difference in truth value come from? There seems to be no other way to explain it than by the difference in emphasis. The arguments  $Fc$  and  $Ge$  of the predicate  $C_F$  are facts. The only difference between (1) and (2) lies in the different value attributed to the variable  $F$  that, in a statement  $C_F(Fc, Ge)$  expressing causal responsibility, represents the efficacious property of the cause.

When (1) is interpreted as (1') and taken to have the structure  $C_F(Fc, Ge)$ , it can be further analysed according to our schema:

(O')  $C_E(c, e) \wedge Fc \wedge Ge \wedge L(F, G)$ .

This brings out the fact that (1) entails that there is a nomic relation  $L(F, G)$  between the properties of drinking hemlock and dying. The nomic element in its content makes (1) an opaque statement.<sup>5</sup> (2) makes reference to the same events as (1)<sup>6</sup> but does not attribute the same causal role to the properties of these events. Both (1) and (2) attribute to the cause (a temporal part, or "time slice", of Socrates including time  $t$ ) the properties of drinking hemlock and of it taking place at dawn. The crucial difference between them lies in the nomic role they attribute to one and the other of these properties with respect to the effect, Socrates' death.

Not all statements containing emphasized expressions are like (1) and (2). Indeed it turns out to be necessary to distinguish between two kinds of usage of emphasis:<sup>7</sup> semantic and non-semantic. Only in the former type of usage can a change in emphasis alter the truth value of a statement. In cases of non-semantic usage, changes in emphasis do not make any difference for truth value.<sup>8</sup> Imagine a situation in which the auditor already knows that Socrates did *something* at dawn but not *what*. In this case the speaker may wish to draw attention to what is the new piece of information in his statement, by emphasizing it:

(3) Socrates *drank hemlock* at dawn.

Nevertheless, he thereby makes reference to the same event as when, speaking to someone else, he puts the emphasis on different words:

(4) Socrates drank hemlock *at dawn*.

In the latter situation, the new piece of information for the auditors is the *time of the day* at which Socrates' drinking hemlock took place. It depends on the audience, more precisely on the knowledge that the speaker supposes is already acquired by the audience, where it is appropriate for him to put the emphasis: on what he supposes to be the new piece of information provided. However this choice of emphasis, as well as switching the emphasis, does not affect the truth value of the statement, for it does not change the propositional content of the statement.

In another type of case, the expressions used to describe a cause are ambiguous in the grammatical context of the statement. In such a case, it is necessary to use emphasis in order to "disambiguate"<sup>9</sup> the statement and determine its truth value. In statements (1) and (2) emphasis is used semantically. Without any emphasis put on part of the expression designating the cause, such statements may be ambiguous, in case the context does not pick out one interpretation. Consider the following situation in which an utterance (1a) without any emphasis would be ambiguous:

(1a) The fact that Socrates drank hemlock at dawn caused his death.

Suppose (1a) is uttered during a discussion about whether the effect of a poison such as hemlock on the human metabolism differs according to the time of the day at which it is ingested. In this case it seems plausible that (1a) has no well determined truth value as long as the speaker has not made clear whether it should be interpreted like (1), in which case it is true, or like (2), in which case it is false. In such cases, the use of emphasis has a *semantic* function. Where it is put makes a difference to the truth value of the statement.<sup>10</sup>

The case of causal statements containing an emphasized expression can be used as a test case to confirm the capacity of the following three theses to account for the semantics of causal statements:

Thesis I: Ordinary language statements possessing the structure "*Fc* is the cause of *Ge*" have the logical form (O), or equivalently (O'). They are semi-opaque to the extent that they do not allow the substitution of coreferential expressions *F'* for *F* and/or *G'* for *G*, *salva veritate*, but they do allow existential generalization.

Thesis II: Ordinary language statements possessing the structure "*Fc* is the cause of *Ge*" contain an extensional element.

Thesis III: A particular causal relation between events *c* and *e* can be reported by statements that differ in truth value only because they differ in the expressions occupying the places of the predicative variables *F* and *G*.

I shall first show that Thesis III provides a framework that allows us to explain both why Thesis I is true for ordinary language causal statements and why the same statements contain an extensional element, as stated by Thesis II. Both Thesis I and Thesis III are consequences of our earlier analysis according to which ordinary causal predicates have the logical form  $C_F$ , whose arguments are facts. This analysis articulates the idea that the expressions designating cause and effect contain two semantic ingredients that can be separated by analysis. One of these elements refers directly to the cause or effect event, and it is indifferent whether the properties it uses to describe these events are involved in the causal process or not. A second element within each of these expressions has the function of referring to the property of the cause that is efficacious for bringing about the effect and to the property of the effect that is produced by the cause. The presence, in the semantic structure, of an element making direct reference to the cause or effect event makes it possible that, as stated by Thesis III, two different statements report the same causal relation at the level of events, although they make reference to different properties of these events. This is possible a fortiori for statements differing only with respect to emphasis.

For an object, it is not controversial that its various properties offer various ways to refer to it. I can speak of an apple by mentioning different properties of it, calling it once “this red thing” and once “that round thing”. Thesis III says that the same is true for the expressions describing the cause and effect events in causal statements.

Thesis III allows us to do justice to the intuition that the expressions designating the cause in (1) and (2) have a common element making reference to the same event, independently of their descriptive content or the position of the emphasis. This event is a temporal part of Socrates containing the moment at which he drinks the hemlock. This explains the intuition that a statement such as (2), in which the wrong expression is emphasized, contains nevertheless a true element. If the causal statement

- (1) The cause of Socrates’ death was that he *drank hemlock* at dawn.

is analysed according to schema (O), one obtains the following result:

(1-O)  $c$ , designated by the properties of being a drinking of hemlock by Socrates and of taking place at dawn, is cause of  $e$ , designated by the property of being Socrates’ death; the causal process by which  $c$  brings about  $e$  depends on  $c$ ’s property of being a drinking of hemlock by Socrates and affects  $e$ ’s property of being Socrates’ death.

In the explicitly opaque paraphrase (1-O) of (1), two structural elements are separated: an extensional element consisting of the predicate  $C_E$  taking the eventive arguments  $c$  and  $e$ , stating that these events are causally related. The descriptive content of the expressions used to designate  $c$  and  $e$  is not part of the content of this structural element of the statement, their only role being to identify those events, something that could in principle also be

accomplished by simply indicating the spatio-temporal zone they occupy. But (1-O) also contains a second structural element, bearing explicitly on those of the cause's and effect's properties that are involved in the causal process linking  $c$  to  $e$ . In the case of our statement (1), which identifies one of the events, the cause, by mentioning more than one of its properties, emphasis can be used to determine which of these properties is intended to be involved in the second, nomological and intensional, structural element of the content of (1). The use of emphasis allows the user of the statement to make clear which predicates he intends to play the role of the predicate variables  $F$  and  $G$  in a statement of causal responsibility  $C_F(Fc, Ge)$ . No ambiguity can arise with respect to the effect, for the predicate "is Socrates' death" plays both the role of designating the event  $e$  and of specifying  $e$ 's property  $G$ , which is said to be causally affected by the cause  $c$ . By contrast, given that the predicate "is a drinking of hemlock at dawn" can be analysed as a conjunction of two predicates "is a drinking of hemlock" and "occurs at dawn", some contexts of utterance may contain no clues allowing an interpreter of (1a) to understand which of them is intended as designating the efficacious property  $F$  of the cause; emphasis is one means by which a user of (1) can resolve this indeterminacy.<sup>11</sup>

(1) is true because it identifies, beyond the events causally linked, the property of the cause that is efficacious in bringing about the effect. It does not only correctly describe the cause event with the help of the description "is a drinking of hemlock by Socrates at dawn", but it also indicates the property responsible for the effect, i.e. Socrates' death, by emphasizing the predicate "drinks hemlock". The truth of the statement depends in part on the correct identification of the property or aspect of the cause that is responsible for the fact that it brings about  $e$ .

Switching emphasis modifies the value attributed to one of the elements of the logical structure  $C_F(Fc, Ge)$ : it modifies the property selected to play the role of  $F$ . We may isolate the element of falsity in statement (2) resulting from such a switch in emphasis:

(2) The cause of Socrates' death was that he drank hemlock *at dawn*.

by using schema (O):

(2-O)  $c$ , designated by the properties of being a drinking of hemlock by Socrates and of taking place at dawn, is cause of  $e$ , designated by the property of being Socrates' death; the causal process by which  $c$  brings about  $e$  depends on  $c$ 's property of being a drinking by Socrates, occurring at dawn and affects  $e$ 's property of being Socrates' death.<sup>12</sup>

(2-O) shows that what makes (2) false is its mistaken determination of the predicate  $F$  in the factive expression that constitutes the first argument of



$C_F(Fc, Ge)$ . (2) cannot be used as a correct explanation of Socrates' death, because it designates the wrong property of the cause as the property efficacious in bringing about that death. A correct causal explanation of Socrates' death must provide information about which property of the cause is responsible for it. To do this, it must correctly specify all elements of the full concept of causal responsibility, because knowledge of the efficacious property is required for explaining *why* he died. (2) gives the wrong answer to this question whereas (1) provides the correct one. The element of (2)'s content bearing on the efficacious property makes it inappropriate as an explanation, whereas (1) can be used as a correct explanation because it correctly identifies this efficacious property.

To justify the intuition that (2) nevertheless contains an element of truth, it is enough to show that it contains an extensional element  $C_E(c, e)$ . Even though the arguments making reference to the properties involved in a causal process are crucial for the purposes of causal explanation, the concept of causation contains an element that does not bring them into play. When one abstracts from the information a causal statement provides about those properties involved in the causal process, one is left with an informational residue, bearing on the *existence* of a causal relation between the events  $c$  and  $e$ . This residue is expressed with the help of the extensional causal predicate  $C_E(c, e)$ . Interpreting a causal statement according to schema (E) means to take into account only its extensional element, represented by the first conjunct of (O').

In this sense, statement (1) can be given a purely extensional interpretation:

(1-E)  $c$ , designated by the properties of being a drinking of hemlock by Socrates and of it taking place at dawn, is the cause of  $e$ , designated by the property of being Socrates' death.

It has the same informational content as the existentially quantified statement:

(1-E') There is (at least) a property  $F$  of  $c$  and (at least) a property  $G$  of  $e$ , such that  $c$  is the cause of  $e$  and the causal process by which  $c$  brings about  $e$ , where  $c$  has the properties of being Socrates' drinking hemlock and of occurring at dawn and where  $e$  has the property of being the death of Socrates, nomically depends on  $c$ 's having the property  $F$  and on  $e$ 's having the property  $G$ .

Statement (1-E), which is entailed by statement (1), allows us to establish the relation between our Theses II and III. Recall that II says that causal statements have an extensional element and that III says that causal statements differing in truth value may nevertheless make reference to the same events. The content of (1-E) corresponds to an extensional proposition

$C_E(c, e)$ , which constitutes the common element of (1) and (2). The expression  $c$  in (1-E) is the referential element common to the factive expressions  $Fc$  and  $F'e$  appearing respectively in (1) and (2). This referential element  $c$  is the first argument of  $C_E(c, e)$ , which is itself a common element of (1) and (2).  $C_E(c, e)$  expresses the fact that there is an extensional relation of causal type between events  $c$  and  $e$ . For ordinary explanatory purposes, this is however the least important part of the content of the causal statement (1). The extensional element  $C_E(c, e)$ , though a fundamental part of the content of every causal statement, is not in general the most relevant part of the content causal statements used in ordinary and even scientific contexts. However, its presence can explain two facts:

- a) one is the intuition that a statement such as (2), although it is false, contains nevertheless a true element, corresponding to  $C_E(c, e)$ . Statement (2) is false because it attributes a wrong value to the variable  $F$ , but it *entails* the true statement (2-E), which is identical to (1-E) entailed by (1). This is due to the fact that (2) assigns the correct value to the arguments making reference to the events  $c$  and  $e$ . The predicate of statement (1-E) stating only the existence of a causal relation has only eventive arguments  $c$  and  $e$ , which are common to (1) and (2). This is the part of (1) and (2) that bears only on the events but not on the properties involved in the causal process.
- b) The presence of an extensional element in causal statements allows us to explain how it is possible that statement such as (1) and (2) differ in truth value although they make reference to the same causally related events. This is possible because what makes one false and the other true depends only on the value assigned to the predicative arguments  $F$  and  $G$ , but not on the eventive arguments  $c$  and  $e$ . The true statement (1) and the false statement (2) agree with respect to the latter.

### **Føllesdal's argument for the non-extensionality of the causal context**

As we have just seen, the factive concept of causation represented by  $C_F(Fc, Ge)$  provides a solution to the problem of interpreting causal statements containing emphasized expressions. This concept can also be used to evaluate Føllesdal's (1965) thesis that the causal context is not extensional.<sup>13</sup> He uses the following three statements as an example:

- (5) "It is causally necessary that the man who drank from that well got poisoned."
- (6) "The man who drank from that well = the man who was born in  $p$  at  $t$ ."

(7) “It is causally necessary that the man who was born in  $p$  at  $t$  got poisoned.”

(Føllesdal 1965, p. 264)

His argument is this. It is possible to imagine a situation in which (5) and (6) are true although (7) is false. Therefore, the context created by the expression “it is causally necessary that” – which Føllesdal identifies with the causal context in general – is not extensional: (7) results from the substitution, in (5), of one expression for another of which (6) says that they are coreferential. However, the substitution does not preserve truth value, for in the imagined situation, (7) is false although (5) is true.

Our analysis allows us to explain these facts without abandoning, as Føllesdal recommends, the thesis that causal statements allow for an extensional interpretation. This explanation is possible thanks to our distinction between the causal relation between events and the relation of causal responsibility of one fact for another. Insofar as one interprets the expression “it is causally necessary that” as Føllesdal does, i.e. as expressing the mere existence of a causal relation, we should deny his intuition that (7) is false. Causal relations between events link particular entities. Here the cause is an event during which the person drinks water from the well, and the effect is an event during which the person undergoes the consequences of the poison contained in the water. In order to state the mere fact that these two events are causally related, it is irrelevant which *descriptions* one uses to make reference to the person involved in these events. In particular, the truth value of (5) does not change when one replaces the expression “the man who drank from that well” by the expression “the man who was born in  $p$  at  $t$ ”. To the extent that, according to hypothesis (6), these expressions make reference to the same individual, (5) and (7) designate the same causal relation at the level of events.

What lies behind the impression that (7) is false, is the tendency to interpret the expression “it is causally necessary that” as if it expressed the concept of *causal responsibility*. In other words, one may easily interpret (7) as if it designated, over and above the causal relation itself, the property of the cause that is efficacious in bringing about the effect. Then one interprets (7) as meaning: “The fact that this man was born in  $p$  at  $t$  is *causally responsible* for the fact that he got poisoned.” This is obviously false in the situation described by Føllesdal, whereas the analogous interpretation of (5) yields the correct result: “The fact that this man drank from that well is *causally responsible* for the fact that he got poisoned.” The tendency to interpret causal statements as having the factive meaning of causal responsibility stems from their use in causal *explanations*. What counts most in a causal explanation is not the identification of the causally related events as such, but the identification of the *properties* involved in the process leading from cause to effect. It is in general not relevant to provide information about properties that are causally inert in the process one speaks of, such as the

date and place of birth of the poisoned person. When one designates the cause, the auditors expect that it be characterized by those of its properties that are responsible for its bringing about the effect. For this reason, fully satisfying causal explanations can only be given with factive causal statements expressing the concept of causal responsibility.<sup>14</sup>

### **Causal responsibility and negative and relational facts<sup>15</sup>**

Many causal judgments bear on facts. However, factive causal statements cannot be directly translated into eventive causal statements. The causal link expressed by such statements attributing causal responsibility of one fact for another fact is more complex than the relation between events we have analysed in terms of transference. In fact, as we have seen, a causal link between events is part of (and presupposed by) the relation of causal responsibility between facts concerning these events. However, some cases of factive causation seem to be incompatible with our claim that a relation of causation between events is always part of what makes a statement of causal responsibility between facts true. We have already come across (in Chapter 1, pages 38–9) an example of a class of judgments raising this problem: the statement “the ice cube cools the surrounding water”. This is a causal judgment where the effect is a *negative* fact. Such statements attribute causal responsibility for the fact that the effect *does not have* a certain property. We have argued earlier that it is equivalent to the statement that “the addition of the ice cube to the water at  $t$  is responsible for the fact that at  $t+\Delta t$  the surrounding water *lacks* part of the thermal energy it possessed at  $t+\Delta t$ ”.

The statement about the ice cube cooling down the water raises a problem for our account because the relation it expresses between cause and effect goes in the reverse direction with respect to the direction of the transference of energy: the statement designates the fact that the ice cube is added to the water as the cause, and the fact that the surrounding water gets cooler, as the effect. However, energy is not transferred from this cause to this effect, from the ice cube to the surrounding water, but rather in the reverse direction, from the surrounding water to the ice cube. We have seen that cases like this lead Fair (1979, p. 242–243) to the conclusion that the direction of causation, at least in such cases, is objectively indeterminate. This conclusion cannot be accepted within a reductionist theory of causation, because it entails that neither of the two commonsense causal judgments about the situation is objectively correct: that the addition of the ice cube causes the cooling of the water and that the surrounding water causes the ice cube’s warming up. Rather, it entails that there is no objective asymmetry making one event the cause and the other the effect, as long as one has not arbitrarily chosen one of the possible descriptions of the two causally related phenomena. Fair takes the direction of causation to be determined by subjective explanatory interest, which is unacceptable for a theory aiming at reducing causation to an objective relation.

We cannot either be satisfied by Aronson's (1971a) strategy to explain the coexistence of these two apparently contradictory causal statements by introducing a new terminology that allows him to avoid an explicit contradiction. Aronson says of the cooling of the water by the ice cube: "As a *result* – but not an effect – of this process, the water loses its heat and, thus, is at a cooler temperature than before. In other words, the correct description of this case is that in the process of causing the ice to melt, the water gave up some of its heat, i.e. became cooler" (Aronson 1971a, p. 425; emphasis Aronson's). Introducing a distinction between *effect* and *result* may be the first step towards an explanation, but it does not by itself explain how these two statements can simultaneously be true.<sup>16</sup>

Our own analysis uses the fundamental distinction between factive causal statements and eventive causal statements; it yields the result that we have here two factive statements with different objective truth conditions, which attribute the roles of cause and effect to different facts:

(8) The fact that the surrounding water emits heat is the cause of the fact that the ice cube melts.

(9) The fact that the ice cube absorbs heat is the cause of the fact that the surrounding water cools down.

Only the former statement entails an eventive statement according to our inference schema (I1).<sup>17</sup> Statement (8) expresses a relation whose objective direction corresponds to the direction of the causal process from cause to effect. The physical reduction of the asymmetry between cause and effect concerns the relation between events: in this case, the physical ground of the asymmetry lies in the intrinsically irreversible character of the causal process of the transmission of heat. The direct way to determine which of the facts referred to by a factive causal statement corresponds to the cause and which to the effect, is by studying the underlying relation between events that is part of what makes the statement true.<sup>18</sup>

The difficulty in doing this stems from the fact that it is not always possible to apply inference schema (I1) and thereby to construct an eventive statement on the basis of the mere linguistic form of the factive statement. As we have seen in Chapter 5, this is not possible if the predicate of the factive expression designating cause or effect has a negative sense. The difficulty is that there is no formal criterion for when a predicate has a "negative" sense.<sup>19</sup> However, there is an empirical fact of the matter about which is cause and which effect. Let us try to transform both statement (8) and (9) into eventive form:

(10) The emission of heat by the water is the cause of the melting of the ice cube.

(11) The absorption of heat by the ice cube is the cause of the water's cooling down.

According to our reductive account of causation between events, only (10) is literally true but not (11). The events that (10) identifies as cause and effect correspond to the direction of the transference of energy in the form of heat between the water and the ice cube. (11) expresses this relation in the reverse order which makes it strictly speaking false. Contrary to (8), (9) only refers to two nomically related properties, and to the existence of a causal link between two events possessing these properties. However, one cannot extract from (9) explicit expressions designating the cause and effect events of that causal relation. Its resemblance to (8), where this is possible, is therefore superficial.

A similar account can be given for another problematic case reported by Fair (1979): "The recoil ball causes the incident ball to *stop* because stopping, the cessation of motion, is a *loss* of translational energy which the incident ball loses to the recoil ball" (Fair, 1979, p. 243; emphasis Fair's). Fair is wrong to think that the truth of such statements forces us to abandon the conviction that there is an objective difference between cause and effect. Once again, there are two factive statements attributing the roles of cause and effect to different facts. The first says:

(12) The fact that the incident ball stops is causally responsible for the fact that the recoil ball starts moving.

The second says on the contrary:

(13) The fact that the recoil ball starts moving is causally responsible for the fact that the incident ball stops.

As soon as we transform these statements into an eventive form, we get a similar result as with the preceding example:

(14) The incidental ball's stopping is the cause of the recoil ball's starting to move

is true whereas:

(15) The recoil ball's starting to move is the cause of the incident ball's stopping

is false. Once more, our reduction of causation to transference allows us to explain the difference in truth value. (14), but not (15) designates as the cause the event that emits an amount of a conserved quantity – in the present case, momentum – and as the effect, the event receiving that amount. (15) is false because it designates as the effect the event that really plays the

role of the cause and as the cause, the event that really plays the role of the effect.

In an important paper, Mellor (1987, p. 207ff.) argues first that events are not the fundamental terms of the causal relation and second, that some causal statements are not made true by any relation between particular entities. He bases his case on two types of factive causal statements, where the expressions referring to cause and effect contain respectively negative and relational predicates. Let us first consider a case in which cause and effect are negative facts. Here is Mellor's example (Mellor 1987, p. 207; emphasis Mellor's):

(16) Don "did *not* die, because he did *not* fall".

If we transform (16) into a statement in which cause and effect are represented by nominalized factive expressions, we get:

(17) Don's not falling was causally responsible for his not dying.

Expressions (18) and (19) have a negative sense:

(18) Don's not falling. (FACT<sub>1</sub>)

(19) Don's not dying. (FACT<sub>2</sub>)

According to our results from Chapter 5 (on pages 153–7) this means that we cannot apply inference schema (I1). However, we can formally apply inference schema (I2), once for the cause argument and once for the effect argument. Statement (17) has the form:

FACT<sub>1</sub> is causally responsible for FACT<sub>2</sub>.

According to (I2), applied twice, we can infer from (17) that

(20)  $(\exists e_1)(\exists e_2)$  ( $e_1$  and  $e_2$  are events and  $e_1$  is the cause of  $e_2$ ).

However, the content of the factive expressions FACT<sub>1</sub> and FACT<sub>2</sub> imposes constraints on the events  $e_1$  and  $e_2$  corresponding to them. In particular, the individual Don satisfying the negative predicates "does not fall" and "does not die", must participate in the events  $e_1$  and  $e_2$ ; in other words, the spatial extension of  $e_1$  and  $e_2$  must include the spatial extension of Don's body. But the crucial consideration concerns a temporal constraint. The facts (18) and (19) exist at every moment at which Don satisfies their respective predicates. However, in (17) these expressions are embedded in a causal context. Thus, the truth of (17) entails first that both facts (18) and (19) exist; but second, given that causal relations are localized both in space and time, the predicate "is the cause of" imposes on its arguments a precise temporal condition: If

these arguments designate facts, such as (18) and (19), they must be interpreted so that the individual Don satisfies their respective predicates at the moment at which the facts (18) and (19) satisfy the predicate “is causally responsible for”. True, (18) and (19) express a fact at every moment at which Don satisfies the predicate they contain. However, as soon as they are embedded in a causal context as they are in (17), the truth of (17) guarantees that (18) and (19) are facts *at the moment* of the causal interaction. More precisely, if (18) is causally responsible for (19), then the event on which (18) bears, happens earlier than the event on which (19) bears. In the present case, these events are temporal parts of an individual object, Don; spatially, the events are localized where Don is.

As we have seen in Chapter 5, if the predicates constitutive of (18) (in the terminology of Chapter 5, the expression  $FACT_1$ ) and (19) ( $FACT_2$ ) are negative, it is impossible to extract from the mere content of the expressions  $FACT_1$  and  $FACT_2$  explicit eventive expressions designating the events  $e_1$  and  $e_2$  on which these facts bear. However, this is no reason to conclude, as Mellor does, that these events do not exist or are not causally related. The fact that no expressions explicitly designating  $e_1$  and  $e_2$  can be extracted in this case from the expressions  $FACT_1$  and  $FACT_2$  does not show that such expressions do not exist. Mellor himself considers the expressions:

(21a) Don’s “hanging [ . . . ] on”

and (Mellor 1987, p. 208):

(21b) his “surviving”.

(21a) and (21b) have a grammatical form that allows to interpret them as referring to events, in other words as eventive expressions. What is Mellor’s reason for rejecting the idea that such expressions can refer to events, and that eventive expressions can occupy the argument places of causal statements? Let us consider “survival” which is unambiguously an eventive expression. Why does Mellor think that

(21) Don’s hanging on caused his survival.

does not express a causal relation? Two confusions must be avoided with respect to expressions (21a) and (21b). First, these expressions are general or classificatory, in the sense that they do not designate a unique event. On the contrary, a general expression such as “his survival” or “his surviving” may designate innumerable events: Don’s whole life consists in a series of events falling in the category “Don’s survival”. Of course, the more a category of events encompasses, the fewer are the occasions in which the category may be useful, in the sense of providing relevant information in a given context.



This is in no way specific for events. The same is true for objects. One can refer to any concrete object with the expression “material object” but there are very few situations (most of them occurring in metaphysics seminars) in which the statement containing this expression may provide relevant information. Consider the statement:

(22) Yesterday I bought a material object.

What is important in the present context, is that (22), although it is practically useless and insufficiently informative in most situations, may nevertheless be true. More precisely, it may be true by virtue of the fact that the expression “a material object” may designate in the context of utterance a particular material object. (22) may well violate the rules of relevance, according to which an act of communication should provide as much information as possible and as required, but this alone does not prevent it from being true.

The same is true for statement (21) and the expressions (21a) and (21b) it contains. “Don’s survival” can be used to make reference to all temporal parts of Don in which he is alive, and “Don’s hanging on” can be used to make reference to all temporal parts of Don in which he hangs on. Furthermore, statement (21) can express innumerable causal relations: every event falling under the category “Don’s hanging on” is the cause of an event falling under the category “Don’s survival”. This makes (21) less informative than other causal statements that make reference to a unique causal relation. (21)’s capacity to be used to refer to a particular causal relation is analogous to (21b)’s capacity to be used to refer to one particular event falling under the category “Don’s survival”. (21) may be uninteresting from a pragmatic point of view, but this does not prevent it from being capable, from a semantic point of view, of referring to a particular causal relation between particular events.

These considerations lead us to the second reason that may stand behind Mellor’s refusal to admit that expressions (21a) and (21b) can designate events. Mellor says that the pseudo-entities these expressions refer to “are really negative events, existing only by definition”, and that “negative events cannot be particulars, because there *are* no negative particulars” (Mellor 1987, p. 208; emphasis Mellor’s). Nevertheless, he acknowledges that it is not necessarily apparent from the logical form of an expression whether it is “negative”. The predicates “survives” and “hangs on” do not contain any explicit negation. It would seem that the reason for which Mellor denies that expressions (21a) and (21b) can designate particulars, lies in his conviction that all events are changes. “Events, when extended in time, have temporal parts, and things do not [ . . . ]. That is what lets events be changes in things” (Mellor 1987, p. 203).

True causal statements such as (21) seem to show the contrary. However, Mellor will not be persuaded even by a large number of causal statements in which neither the cause nor the effect are changes, insofar as he does not accept causal statements from ordinary language as relevant data for this

debate. “Fundamental to grammar it may be, but not to ontology” (Mellor 1987, p. 203).<sup>20</sup> “Grammar”, more precisely the existence in natural language of eventive expressions – constructed with gerundive forms of predicates expressing states – that designate non-changes seems to confirm our thesis and disconfirm Mellor’s. Mellor retorts that grammar is not decisive for questions of ontology. What other reasons can be brought to bear on the ontological question whether there are particular events that are not changes, reasons that do not depend on the semantic analysis of predicates? One cannot even justify the existence of a well-determined class of such events, without using a dichotomy between predicates expressing changes and predicates expressing stable states. Non-changes are precisely those events that can only be designated with the help of expressions containing predicates that do not express change. In this sense, it would seem that the whole distinction is only grounded on a semantic distinction between predicates.

The impossibility to ground the distinction between changes and non-changes independently of the predicates designating them pleads against Mellor’s thesis, insofar as he denies the relevance of language for ontological questions. Mellor cannot ground his distinction on the existence of two classes of predicates. But at the physical level, it is hard to find any criteria for distinguishing events implying change from others. The distinction is clear with respect to a given phenomenological aspect of objects: a chameleon can change colour but it can also maintain its colour state during a certain time. Mellor’s thesis is that there are states that do not contain any particulars which can causally interact. However, we can retort that there is, from a physical point of view, no moment in the existence of an object in which it is not taking part in causal interactions. From the point of view of causal relations, there is no difference between events of light being reflected from the chameleon’s skin while it changes colour and while it does not change colour. At every moment of its existence an object is both cause and effect. By distinguishing between moments in which a given property of an object changes and other moments at which it does not change, one can classify the events that are the temporal parts of an object in two classes. But this does not prevent both classes from being entirely composed of events, and hence of particulars. As long as Mellor denies the relevance of language, he owes us a physical argument to the effect that a state, such as Don’s state while he hangs on, is not composed of particulars capable of interacting causally.

The idea that the *absence* of a property can be a cause is as old as Aristotle’s *Physics* which says that “the same thing is the cause of opposites. That which, by being present, is the cause of so and so, is sometimes held responsible by its absence for the opposite; thus the loss of a ship is set down to the absence of the steersman, whose presence would have been the cause of its being saved” (Aristotle, *Physics* II, 3, 195a, 11–14; transl. p. 29).<sup>21</sup> We may account for this example in the following way: designating the cause by a property it does not have, the statement that the absence of

the steersman is the cause of the ship's sinking allows us to infer, according to the schema (I2), that there exists an event that is the cause of the sinking and that there is a positive property  $P$  of this event that is causally responsible for the disaster. It is only that the statement does not tell us what this property  $P$  is: it might be the property of the ship's hull to be pierced by a riff.

It remains to be shown that one can, in an analogous way, systematically infer the existence of a particular event when a true causal statement designates as the cause a fact whose constitutive predicate expresses a *relational* property. Mellor gives the following example (Mellor 1987, p. 211):

(23) "Don's rope being the weakest caused his fall to be the first."

In this statement the nominalized expression

(23a) Don's rope being the weakest

designates the fact that Don's rope was the weakest. (23) has the structure:

FACT<sub>1</sub> is causally responsible for FACT<sub>2</sub>.

Once again we may infer by virtue of schema (I2):

(20)  $(\exists e_1)(\exists e_2)$  ( $e_1$  and  $e_2$  are events and  $e_1$  is the cause of  $e_2$ ).

Event  $e_1$  must spatially include Don's rope and  $e_2$  must include Don. Here it is the relational meaning of the predicates "is the weakest" and "is the first" that makes it impossible to extract from (23a), at a purely formal level, a nominalized expression of eventive type. Constructing such an expression requires knowledge of the causally efficacious property of the rope with respect to the effect. Surely the rope has such an intrinsic, i.e. non-relational, property, because the event consisting in a temporal part of the rope can only act as a cause by virtue of its intrinsic properties.<sup>22</sup> (23) does not deliver any information about the property by which the rope brings about Don's fall, the effect described in (23). But this does not show that it does not *have* such a property. It is not the rope that lacks a causally efficacious property; rather, the hearer of (23) lacks information about it, which would enable him to construct an explicit eventive expression EV designating  $e_1$ . The same is true for the effect  $e_2$ .

What is Mellor's reason for claiming that the relation referred to by (23) cannot take particulars as its terms? It is true that the entities directly designated by (23) are factive entities and therefore not particulars. But Mellor holds that furthermore, one "cannot get from [ . . . ] [statements like (23)] a relevant truth of type" (Mellor 1987, p. 211) "' $c$  causes  $e$ ', which represents cause and effect as events" (Mellor 1987, p. 205). This can only be justified by the idea that a statement to the effect that two events  $e_1$  and

$e_2$  are causally related, cannot be a “relevant truth” if one is not in a position to designate the events by explicit eventive expressions. But, as we have seen, (23) contains a lot of information about the spatial and temporal locations of  $e_1$  and  $e_2$ . This is enough to show that it can be informative, and therefore a relevant truth, to state that  $e_1$ , which obeys those constraints, is the cause of  $e_2$ , which obeys other constraints. In short, the fact that one cannot – on a purely formal basis – infer from (23) a statement in which the factive expressions are replaced by explicitly eventive expressions, neither shows that there are no particular events underlying the facts mentioned in (23) nor that there are no eventive expressions referring to them.

According to Mellor (1987, p. 203), the causal statement (23) “depends for its truth on how these particulars are referred to” (Mellor 1987, p. 211–12), i.e. the particular object referred to by “Don’s rope” and the particular event designated by “Don’s fall”. True, the identity of the entities designated as cause and effect in (23) is determined by the predicates of the respective factive expressions. But (23) entails (20) – with the constraints on  $e_1$  and  $e_2$  we have indicated – and the truth of a causal statement relating particulars does not depend on the way in which these particulars are designated.<sup>23</sup>

### **Interruption and triggering**

We have seen that it is not possible to infer an explicitly eventive statement from a factive causal statement in which the predicates of the factive expressions designating cause or effect have a negative or comparative sense.

These are not the only sorts of causal statements raising a challenge for our account: Statements reporting an interruption or a triggering are two other types of causal statements that do not directly report a relation between events reducible to transmission, nor a relation between facts from which one could infer a statement reporting a relation between events that is reducible to a transmission. Consider first a statement of interruption (Ehring 1986, p. 251; 1997, p. 9):

(24) A switch is turned off causing the light to go out.

My touching the light switch can be said to be the cause of the light’s going out. However, as Ehring (1986) points out, what such statements designate as a cause does not transmit anything to the effect, but rather interrupts such a transmission, and what they designate as an effect does not receive anything transmitted from the cause, because it is rather the interruption of such a reception.<sup>24</sup> Ehring takes statements such as (24) to refute theories reducing causation to transmission because “there is no transference of a quantity from the switch to the light” (Ehring 1986, p. 251). This seems to be a bad way of putting the objection for if the effect is taken to be the light’s going out, one should not blame the transference account for the fact that there is no transference to *the light*, which is very different from the

event of the light's *going out*. But let us restate the objection: (24) refutes the reduction of causation to transference because the turning of the switch transfers nothing to the event of the light's going out, and nothing to the subsequent state of the light's being out, a state negatively characterized by the absence of light.

Once more,<sup>25</sup> Aronson (1971a) solves this problem for the transference account by introducing a terminological distinction. Aronson proposes to call the triggering factor an "occasion" (Aronson 1971a, p. 425) rather than a cause. But such a terminological choice is only the first step of an analysis of such cases. The crucial fact which makes our account compatible with statements such as (24) is this. Statements of interruption deliver a complex piece of information which is causal only in part. (24) makes reference to a causal relation *t*, but it explicitly designates only its cause event *c* which is the turning of the switch.

Figure 7.1 sketches the situation reported by (24). The turning of the switch (*c*) causes (*t*) the opening of the electric circuit (*e*) that sustained the emission of light from the bulb. However, (24) designates neither the effect event *e* which is the final point of *t* nor a relevant fact bearing on this event. Instead, it designates a state *f*, the light's being out, which is negatively characterized as a consequence of *t*. Rather than to *e*, (24) refers to *f*, which is a nomic, but not causal, consequence of *e*, and thus of *t*. *f*'s nomic dependence on *e* can be expressed by saying: state *f* is present because *e* has occurred. In this statement, the word "because" does not have a causal sense: it points to a relation of *nomic dependence* which *explains* *f*, but not in a causal sense by mentioning a cause of *f*. The lawful dependence of *f* on *e* might be expressed by a (non-causal) conditional: if the circuit linking the light source to the source of electric energy is interrupted, no current flows and no light is emitted.

What makes the situation reported by (24) particularly complex from the point of view of causation is the fact that *f* is designated in a purely negative

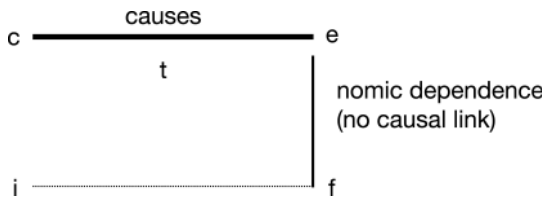


Figure 7.1

*c*: turning of the switch

*e*: opening of the electrical circuit

*t*: causal relation based on transference (in particular, of electrical charge)

*i*: initial state – light on

*f*: final state – light out

way. (24) says that  $t$  makes the light “go out”, which means that it provokes a transition from an initial state  $i$  where the light is on, to  $f$ . The difficulty in giving a causal analysis of statement (24) lies in the fact that  $t$ , the causal relation expressed, is not a relation linking  $i$  to  $f$ . It is just that the result  $e$  of  $t$  is such as to make impossible the causal relations underlying the state  $i$ , namely the flow of electricity through the light bulb. Now someone might be prepared to accept the idea that the relation between  $e$  and  $f$  is not causal because nothing can cause the absence of something, but still ask whether  $i$  might not somehow be implied in the causal process designated by (24). The answer is again, no:  $i$  is a stable state consisting of the permanent flow of electrons through a wire and a light source.  $i$  is in itself a causal process but it is not the cause of any of the events constituting  $t$  which is referred to by (24). (24) does not say, nor is it the case in standard situations in which (24) would be true, that the flow of electrons constituting  $i$  intervenes either as the cause or as the effect in the causal relation  $t$  resulting in the interruption  $e$ . Maybe the flipping of the switch  $c$  uses some electrical current, but this need not be the same current which is the basis of state  $i$ .

Finally, Ehring’s objection can be answered in the following way. (24) reports a complex situation involving both a causal process and a relation of nomic but non-causal dependence. (24) directly mentions  $c$ , the turning of the switch, which is the cause of a process  $t$  whose end point  $e$ , the interruption of the electrical circuit, is not explicitly mentioned. However, the truth of (24) requires the occurrence of  $e$  or some similar event. Instead of mentioning the effect  $e$  of  $c$ , (24) mentions the state  $f$  which is a nomic but non-causal consequence of  $e$ . An aspect of statement (24) which is potentially misleading for its causal analysis, is the fact that  $f$  is described by contrasting it with a previously obtained state  $i$  which is not causally involved, neither as a cause nor as an effect, in  $t$ .

Similarly, factors that release or trigger a causal process are often treated as causes, whether or not they contribute anything to what is transmitted to the effect. Consider statement (25) (Ehring 1986, p. 252):

(25) A catch is released “causing” a spring to pull a weight some distance.

Given that the removal of the catch does not contribute any part of the energy the spring deploys, should we consider (25) as a causal statement?<sup>26</sup>

Figure 7.2 represents the structure of the situation making statement (25) true. The crucial point is that the relation between the removal of the catch,  $e$ , and  $f$ , the spring’s expanding is a relation of nomic dependence, not a causal relation. For practical purposes it is often of little significance whether one obtains a certain result by causing it or by causing something else that has the desired result as a nomic consequence. This explains why (25) may seem plausible although the relation between  $e$  and  $f$  it refers to is not causal.

A potentially misleading aspect of the situation reported by (25) is the fact that  $f$  itself consists in a causal process, but, in general, the triggering

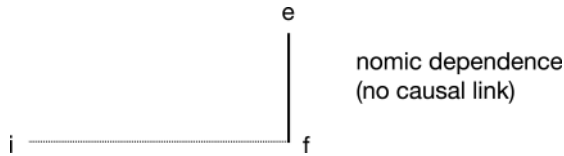


Figure 7.2. Releasing/triggering

- e*: the catch being released  
*i*: initial state – spring compressed  
*f*: final state (process) – spring expanding

event *e* is not even a part of *its* cause.<sup>27</sup> The relation between the removal of the catch *e* and the causal process *f* is one of non-causal nomic dependence. The absence of the catch is a negative condition which makes it possible for the cause of *f* to act. But the absence of the catch is not even part of the cause because an absence cannot be a cause any more than it can be an effect. (The latter fact is relevant in cases of interruption which have been discussed before.) The removal of the catch is a causal process, but no part of it contributes *causally* to the process of expansion of the spring. It only makes it nomically possible.

Both cases of interruption and of triggering may illustrate the fact that factive causal statements contain in general more information than eventive causal statements.<sup>28</sup> Statement (24) contains not only information about *c*, *e* and *f*, but entails the existence of a state *i* preceding *f*, in which electrical current passes through the switch and makes the lamp emit light. Similarly, statement (25) entails the existence of a state *i* preceding *f* in which the spring was compressed.

### Counterfactual information and moral responsibility

Fair (1979) describes another situation that raises a problem for our reduction of causation to a physical relation of transmission. Three devices *A*, *B* and *C*, positioned around a lake, are capable of producing waves of frequency *f*. Their movements are so coordinated that the waves produced by *A* and *B* are always in phase but shifted by 180° (or  $\pi$ , in radial units) with respect to the waves produced by *C*. At a fixed position on the surface of the lake, the waves emitted by *A* and *B* interfere constructively, so that if both *A* and *B* are active, the summit of the waves at that point are twice as high as if only one of *A* and *B* was active. A float at that point rises twice as high when *A* and *B* both operate as when *A* or *B* operates alone. By virtue of the phase shift, the waves produced by *C* interfere destructively with the waves produced by *A* and *B*. This means that if *A* and *C*, or *B* and *C*, are operative but not the third device, the float remains immobile. When the

third device is switched on in this situation, the float will move exactly in the same way as if this third device were operating alone.

Fair presents this situation as a case “where there is seemingly no matter of fact about the origins of a particular quantity of energy” (Fair 1979, p. 238). However, it seems incorrect to conclude with Fair that such cases “exhibit a parallel lack of matter of fact about causation”. It may help our reasoning about this case to give the circumstances a more dramatic twist. Ehring (1986) imagines that the float is connected to the trigger of a gun: when the float reaches the level that corresponds to the amplitude of the waves produced by one of the devices *A*, *B* or *C*, it pulls the trigger. To dramatize the situation, imagine that Smith finds *C* in operation and switches on device *B* but not *A*. Jones who intends to kill Fred arrives at time  $t_1$ , when both *C* and *B* are in operation. At time  $t_2$ , when Fred stands exactly in the aim and range of the gun, Jones switches on *A* and Fred is killed. In this case, says Ehring, “although there are admitted indeterminacies about the source of the energy which is manifested in the gun’s firing, Jones [by mistake, the text says “Smith”] causes Fred’s death by causing *A* to operate at  $t_2$ ” (Ehring 1986, p. 255).

According to Ehring (1986, p. 254–255) and Dowe (1992a, p. 203), such cases refute all transference theories of causation, for it seems that, in such a situation, the effect, Fred’s death, is well determined but the source of energy causally responsible for it is not. Reducing causation to transmission seems to lead to the erroneous result that the cause of Fred’s death is indeterminate, whereas it seems clear that Jones’s switching on *A* at  $t_2$  is the cause. The difficulty here is not the fact that the case involves triggering, where the energy reaching the effect does not come from the triggering factor (cf. preceding section). Rather, the problem is that it does not seem to be possible to locate the origin of the kinetic energy that reaches a particular point of the surface of the water, the float, and finally the trigger of the gun, in a case in which this energy is transmitted through superposed waves.

The distinction between eventive and factive causation allows us to account for cases of this type without abandoning the reduction of eventive causation to transference. Speaking in terms of events,

(26) Fred’s death was caused by an event that includes, spatially, the three devices in activity, and that precedes the triggering of the gun by the temporal interval necessary to bridge the distance between the devices and the float with the group velocity  $v_{gr}$  at which amounts of energy are transmitted by the waves.

If the distances between the float and the three devices *A*, *B* and *C* are not equal, the temporal limits of the cause event must be broad enough to include all instants at which amounts of energy reaching the float at the moment of triggering leave one of the three devices. However, to find out what is causally – and morally – responsible for the fact that Fred dies, we



must switch to facts. Considering the instant after Jones's switching on the third device, one may attribute causal responsibility to the following fact:

(27) The fact that the three wave producing devices, all having the same amplitude  $H$ , two of them interfering constructively and one destructively, are operating simultaneously is causally responsible for the fact that the float reaches height  $H$ , and thereby for the fact that the gun is triggered and finally for the fact that Fred dies.

From the mere point of view of causation, Jones has no more responsibility for Fred's death than Smith. Each of them has switched on one of the three devices. Of course they do not have the same *moral* responsibility. But determining moral responsibility requires taking into consideration factors that go beyond actual causation: it requires taking *possible* causal relations into account. Establishing Jones' moral responsibility requires taking into account several events, in the temporal order in which they happen in the story. Most importantly, it requires the evaluation of counterfactual conditionals. Factive causal statements can contain information that goes beyond the facts concerning the actual causal relation reported. This is the case with the more specific statement:

(28) The fact that Jones switches on device  $A$  at  $t_2$  is causally responsible for the fact that Fred dies.

(28) differs from (27) by making reference to the instant  $t_2$ : Jones' action is decisive first because the other components of the total cause, named in statements (26) and (27), are already present at  $t_1$  which precedes  $t_2$ , and second because, *if* those other components had been operative alone at  $t_2$ , the float *would not have* reached height  $H$  and the gun *would not have been* triggered. The factive causal statement (28) presupposes these counterfactual conditionals; this is why it contains more information than statements (26) and (27), which only report the actual causal relation leading to Fred's death. The evaluation of such counterfactuals often plays a crucial role in contexts of the attribution of moral responsibility. What makes them true goes beyond actual causation but can be conveyed by factive statements.

Bennett mentions another case in which counterfactual information is a crucial part of the content of a causal statement. One can easily imagine a situation in which the statement (Bennett 1988, p. 139)

(29) The massage was a cause of his death

seems to be false because, on the contrary, the massage has retarded the patient's death. The factive statement providing the complete information may be (Bennett 1988, p. 139):

(30) The massage was a cause [ . . . ] of his dying as late as he did.

The patient's death is really due to the interaction of several factors whose efficacy with respect to death goes in opposite directions: the massage is a property of the cause that retards the patient's death, whereas other factors not explicitly mentioned in (29) or (30) are positively efficacious in bringing it about or accelerating it. Humphreys (1989, p. 286) calls factors of the first kind "counteracting causes" and factors of the second kind "contributing causes". Only factive expressions are apt for carrying this complex information in a single causal statement: (30) contains, as well as information on actual events and actual causal relations between them, counterfactual information about possible causal relations, which would have existed if the circumstances had been different. It contains the implicit information that, in the absence of the massage, the positive cause that is not explicitly mentioned would have acted more rapidly.

However, (29) is not literally false. It is only unacceptable because it *suggests* another factive causal statement that is literally false. The structure of (29) is

(31)  $EV_1$  is the cause of  $EV_2$ .

Applying our inference schema (I3\*) twice, we can infer:

(32)  $(\exists P_1)(\exists P_2)(P_1$  is a property of  $e_1$ ,  $P_2$  is a property of  $e_2$ , and the fact  $P_1e_1$  is causally responsible for the fact  $P_2e_2)$ ,

where  $e_1$  and  $e_2$  are either the events designated respectively by  $EV_1$  and  $EV_2$  or objects contained in these events.

The content of (29) alone does not give us information about the properties  $P_1$  and  $P_2$ : it does not tell us which property of the massage was causally responsible for which property of the patient's death. (29) seems false although there is really a causal relation between the events designated by "the massage" and "the patient's death", because it suggests, by virtue of a pragmatic rule of communication, that the descriptive terms used to designate these events also contain information on the properties involved in the relation of causal responsibility. However, this is not the case with (29). More precisely, (29) suggests – but only suggests by virtue of a pragmatic implicature, rather than entailing it semantically – that the properties  $P_1$  and  $P_2$  in the underlying structure of causal responsibility (32) are those of being a massage and of being the patient's death. In other words, the descriptive content of the eventive expressions in (29) suggests that the corresponding factive statement with structure (32) is:

(33) The fact that the patient was massaged was causally responsible for the fact that he is dead.

(33) is false in the circumstances. But this does not show that (29), which suggests it by virtue of pragmatic rules, is itself false.

## Conclusion

We have used the concept of causal responsibility to shed light on a number of controversial issues. Our use of a concept with such a relatively complex logical structure is justified by its explanatory fruitfulness. As we have shown in the first section of this chapter, causal statements containing an emphasis on a predicate (within the expression designating the cause) that does not designate the efficacious property of the cause are false although they contain a true element. The complex structure of the relation of causal responsibility helps us understand these apparently incompatible semantic intuitions. It is false to say that “the cause of Socrates’ death was that he drank hemlock *at dawn*” because the emphasis lets the hearer misidentify the efficacious property of the cause. Nevertheless the statement entails a weaker statement which is true and which is also entailed by the statement with the correct emphasis: “the cause of Socrates’ death was that he *drank hemlock* at dawn.” The weaker statement entailed by both bears on the underlying relation between events: “the event at which Socrates drinks hemlock and which happens at dawn is the cause of Socrates’ death.”

We have looked at a certain number of causal statements that have been taken by many authors to show that causes are not events but rather facts, or that, to use Humphreys’ words, “events cannot be identified with spatio-temporal regions because any given spatio-temporal region ordinarily contains many properties and changes, some of them causally relevant to a given effect, many of them not” (Humphreys 1989, p. 288). Humphreys expresses the conclusion that causes are facts in a different terminology from ours. However, to say that events are particulars and cannot therefore be causes because a cause contains only certain properties of a particular and not all of them, is equivalent to saying that “events” *are not* particulars. Humphreys adopts the terminological convention, which goes back to Kim (1969; 1973; 1976) and which calls “events” what we have called “facts”. We have shown that our distinction between the cause-event and its efficacious properties allows us to account for all cases that make Mellor and Humphreys, among others, believe that only facts can be causes.

But we have also seen that statements which designate the cause by a property that is not efficacious in bringing about the effect or which mention as a cause a negative or relational fact, an event of interruption or triggering, and statements containing counterfactual information about properties that would have been efficacious if the cause had possessed them, are all factive causal statements. These bear only indirectly on a causal relation between events or on a relation of causal responsibility between facts. We have in each case distinguished between these relations and the way the statement refers to them. Recall our analysis of the substitution of

the expression “the man who was born in  $p$  at  $t$ ” for the expression “the man who drank from that well”, on pages 194–6. This substitution does not change the truth value of the statement reporting the cause of the man’s being poisoned, insofar as it is evaluated as an eventive statement. Insofar as the expressions refer to the same events, our analysis accounts for the intuition that these two statements have a true element in common. They designate the same causal relation at the level of events although only one of them designates the efficacious property of the cause.

# Conclusion

Our aim was to reconstruct the concept of causation while meeting the requirement that it be general enough to adequately deal with both elementary physical processes and the macroscopic level of phenomena we encounter in everyday life. We sought for a unique type of relation that operates at all levels and is integral to what makes causal statements true in all contexts. Even if its analysis makes use of conceptual tools from physics, the causation we are looking for should be an element of the content of all causal statements. Nevertheless, we want causation to be specific enough to exclude some relations, in order to avoid making it an empty concept. In particular, we have argued that it would be wrong to count as causal all relations that one can come to know deductively using scientific laws. Causation is a useful concept only to the extent that it is clearly distinguished from other scientifically grounded but non-causal relations. Our requirement is that the concept of causation be applicable both in commonsense judgments and in applied science. Our analysis should provide a scientifically respectable concept of causation that can in principle be applied in all domains where a commonsense notion of causation is applied. However, it turns out that, in order to preserve the overall coherence in our practice of making causal statements, some commonsense intuitions with respect to this or that causal statement must be corrected.

This requirement seems incompatible with the point of view reflected in the traditional conceptions we have presented in the first chapters of the book. Those who identify causation with determinism make the application of the concept impossible in a world in which certain fundamental laws are statistical. In such a conception, causation can only be applied to deterministic models, for example to certain simple models of classical mechanics. Thus, the more sophisticated a model becomes, and closer to real-world systems with their indeterministic processes, the less it is justified to apply the concept of causation to it. Such a conception is incompatible with our requirement that the concept of causation be applicable to the world of commonsense – which is not deterministic – as well as to phenomena whose scientific explanation requires the use of statistical laws.

The classical account of causation developed in the tradition of logical empiricism robs the concept of causation of any specific content by considering *every* scientific explanation as causal. According to this account, there is no cleavage between commonsense and a scientific account of the world insofar as the reasoning in both accounts is deductive, if often incomplete or elliptic. But, if every scientific explanation were a nomological deduction and if every nomological deduction were a causal explanation, it would be tautological to call “causal” any relation or process that can be scientifically explained. We have seen that there are many non-causal relations that can be scientifically explained. Furthermore we have given examples of world lines with respect to which the question whether they are processes and therefore causal or pseudo-processes and therefore non-causal is of real interest and not just a matter of terminological stipulation.

Finally, to hold, as Russell did in his famous paper “On the Notion of Cause”, that causation is incompatible with a scientific conception of the world, violates our requirement of continuity between commonsense and scientific conceptions of the world. Furthermore this “eliminativist” theory does not even allow for our requirement that it should be the same causation that operates at all levels of scientifically studied phenomena, insofar as it holds that a theory can count as scientific only insofar as it does not use the concept of causation.

Our own reconstruction of the concept of causation meets the requirement. Causal responsibility grounds causal explanations in the domain of phenomena whose evolution and interaction is subject to laws of nature. We have found that there are two aspects of causation, corresponding to two types of causal statements, between which there exist precise relations of entailment. The concept of causal responsibility allows us to take into account the difference between *stating that* event *c* is the cause of event *e* and *saying why* *c* is the cause of *e*. The causal statements figuring in scientific explanations do typically, not only say that event *c* is the cause of event *e*, but also why it is. A causal explanation consists in pointing out that the *fact that c* has property *F* is – by virtue of a law of nature – causally responsible for the fact that *e* has property *G*. The relation of causal responsibility between fact *Fc* and fact *Ge* is a complex relation with the following constituents: 1) There is a causal relation between the events *c* and *e*; 2) the cause event *c* has property *F* and the effect event *e* has property *G*; 3) the instance of *F* in *c* gives rise to the instance of *G* in *e* by virtue of a law of nature linking property *F* to property *G*.

This conception allows us to justify both that causal explanations are perfectly respectable scientific explanations and that they constitute a specific kind of explanation. This is just another way of saying that it allows us to acknowledge that there are also non-causal scientific explanations.

Showing that there is a causal relation in the sense of a relation of transference between events can in certain situations be explanatorily interesting; nevertheless, in the vast majority of explanatory tasks pursued in

science and everyday life, more information is required. In most cases a causal explanation does not only identify the cause and effect but indicates in addition by virtue of which property and by virtue of which law the cause brings about the effect.

A crucial element of the concept of causal responsibility is this: the fact that  $c$  is  $F$  can only be causally responsible for the fact that  $e$  is  $G$  if the property  $G$  constitutive of the effect is nomically dependent on the property  $F$  constitutive of the cause. Whether the phenomenon to be causally explained belongs to psychology, biology or microphysics, a condition for considering one fact as causally responsible for another is that their constitutive properties are linked by a law. This conception is neutral with respect to the issue of the possibility of *reducing* higher-level theories to lower-level theories: it treats microphysical properties in the same way as natural properties studied by other sciences than physics. Still, the concept of causal responsibility contains enough restrictions to put it to critical use. The properties that can participate in relations of causal responsibility must obey two conditions. The first is that they must participate in laws of nature; the second is that they must be non-relational or intrinsic. Even if some laws bear on relational properties, an event is never interacting with another by virtue of such a relational property. True, Coulomb's law links the property of being an electric charge, situated at a distance  $d$  from another electric charge, to the property of undergoing a certain force, attractive or repulsive, in direction of the other charge. The first of these properties is clearly relational; therefore, according to our account, it cannot be causally efficacious. We have solved this apparent paradox by suggesting that Coulomb's law accounts for the macroscopic manifestation of an underlying microscopic process that operates strictly locally: the electromagnetic force does not act at a distance but is the result of the propagation of a field, which can also be described in terms of a local exchange of photons.

The other constitutive element of causal responsibility is a relation between events that we have reduced to transmission:

(S) Two events  $c$  and  $e$  are related as cause and effect if and only if there is at least one physical quantity  $P$ , subject to a conservation law and exemplified in  $c$  and  $e$ , a determinate amount of which is transferred between  $c$  and  $e$ .

Interpreting this reduction statement requires specifying the nature of the events that are the terms of the causal relation: we have chosen a conception of events according to which they are what occupies a spatio-temporal zone or region. This makes events particulars in the same sense as ordinary objects: objects are what occupies a *spatial* region.

The next pressing question is: what exactly needs to be transmitted between such events in order for them to be causally related? The result of

our inquiry is the very general criterion (S): an amount of any conserved quantity suffices for the existence of a causal relation. Energy and momentum are the most familiar among those quantities, but a reflection on the concept of transference itself has led us to the conclusion that they are but particular instances of the general concept of a conserved quantity: what is conserved during the process of transference can be transferred. This is just as far as philosophical inquiry can pretend to go. It is up to physics to discover the explicit, and maybe one day exhaustive, list of conserved quantities which participate in grounding causation.

The situation sketched above in which two electrical charges attract each other can be correctly analysed in terms of causation only after having clearly identified the events taking place in this relation. If the two charges are separated by a finite distance but considered at the same moment then there can be no transference between the events consisting of the temporal parts (sometimes called “time-slices”) of the charges at these moments. Relativity theory excludes the possibility of instantaneous transmission. Given that the existence of a causal relation between events  $c$  and  $e$  is a necessary condition for the possibility that the fact that  $c$  is  $F$  is causally responsible for the fact that  $e$  is  $G$ , any causal responsibility between facts concerning these events is impossible. But if the two charges are taken at moments sufficiently distant with respect to their spatial separation, they may interact causally by exchanging photons, a transmission process grounding causation that is covered by a law of conservation. This shows how the requirement of transmission as grounding causation can be put to critical use: it excludes certain pairs of events, and facts bearing about them, from being causally related. However, in the large majority of causal explanations, the criterion of transference is trivially satisfied with respect to the pairs of events under consideration. In such cases, finding out whether the fact that  $c$  is  $F$  causally explains the fact that  $e$  is  $G$  requires finding out whether the constitutive properties  $F$  and  $G$  are intrinsic, i.e. non-relational, and whether they are related by a law of nature that is instantiated by  $c$  and  $e$ .

The absence of any consideration of transmission in most classical theories of causation may be explained by the fact that the transference criterion is trivially satisfied in many situations in which causal statements are evaluated. Indeed, transference is mentioned neither in the nomological theory of causation nor in theories reducing causation to the increase of the probability of the effect in the presence of the cause, compared to its probability in the absence of the cause. Both these theories reduce causation to a relation covered by a law of nature. We have seen that this makes them incapable of accounting for nomic relations that are not relations from cause to effect: in particular, they cannot distinguish them from relations between two effects of a common cause, from relations going from effect to cause, nor from nomic relations that are not causal at all, such as nomic dependencies between different properties an object possesses simultaneously. What is missing from these theories in order to distinguish causal



relations from other relations covered by laws of nature, is the idea that cause and effect are particular entities located in space and time, and the condition that they are causally related if and only if an amount of a conserved quantity is transmitted between them.

Many philosophers who have at least considered the idea of reducing causation to a relation of transference have abandoned the project because they took certain problems with the concept of transference to be insuperable. Quine's verdict has certainly had a considerable impact and discouraged efforts to further develop a transference account of causation. Quine's main objection is that there is no way to understand the concept of transference which does not make use of causal concepts. If this were true, transference theories would be condemned to circularity by the mutual dependence of the concepts of causation and transference.

However, it is possible to avoid the circle by understanding transference between two events as requiring simply the presence of an identical particular amount of a conserved quantity in these events. This interpretation makes both transference, and the concept of causation it helps to construct, fundamentally symmetrical. That causal relations in our actual world are asymmetrical, and that there is an objective difference between cause and effect, are empirical facts that cannot be discovered a priori, simply by analysing the concept of causation. Having clarified the relation between the asymmetry of causation in the actual world, and the symmetry of transference as we interpret it, we can use conservation laws in order to justify the idea that an identical particular amount of a conserved quantity can be present in two different events.

The issue of asymmetry turns out to be crucial for the defence of the realist account of causation against its main rival, the nomological account, for it may ground the tripartite distinction between non-causal explanations, causal explanations, and statements of the existence of a causal relation. The traditional nomological account conflates these categories and makes these distinctions impossible by categorizing all scientific explanations as causal. It reduces the cause-effect relation to the relation between initial conditions and explanandum in a deductive-nomological explanation. But many laws, such as laws expressing functional dependencies, are symmetrical, and the asymmetry between explanans and explanandum is only produced by the subjective focus of interest of those who seek and provide the explanation. If causation is construed in terms of transference, the asymmetry of causation can be taken to have an objective ground although functional laws are symmetrical: the transference account makes this possible because it does not simply identify causation with the instance of a law.

The analysis of causation in terms of transference between events and of causal responsibility between facts set the agenda for spelling out the details: the first task was to analyse the concept of a law of nature and its instances in concrete situations (Chapters 2 and 3). Then we had to clarify the concept of natural property and find out which properties could play

the role of the constitutive properties  $F$  and  $G$  of cause and effect, which play a key role in the analysis of causal responsibility (Chapters 5 and 6).

The best way to find out which properties can play the role of the causally efficacious  $F$  of the cause and the causally affected  $G$  of the effect, is to study the inferences that can validly be drawn between statements expressing causal responsibility between facts and statements expressing causation between events. We undertook this task by studying the form and meaning of such statements: factive causal statements express relations of causal responsibility between facts; their arguments are factive expressions of the form “ $c$ ’s being  $F$ ”, or “ $c$ ’s  $F$ -ing”, such as “Mary’s singing the song”. Such expressions make reference to facts; a fact, that  $c$  is  $F$ , is construed as what makes true a proposition, that  $c$  is  $F$ . A factive causal statement expressing a relation of causal responsibility between facts contains more information than an “eventive” statement that makes only reference to the underlying causal relation between events. Statements of the latter type only say of a particular event  $c$  that it is the cause of another particular event  $e$ . As an example, we have considered the statement: “Mary’s performance of the song caused my surprise”. “Mary’s performance” is an eventive expression designating a particular event, which has many properties: there are as many facts bearing upon this event as it has properties; I may say of each of these facts that it is causally responsible for the fact that I am surprised. However, among all those statements of causal responsibility, which bear all on the same causal relations between events, only some are true; the others are false although part of what they say is true. The additional information contained in a factive causal statement over and above the information contained in the corresponding eventive statement makes it impossible to deduce the former from the latter. From the statement that Mary’s performance – i.e., an event – caused my surprise, one cannot infer the truth of any factual statement, stating which fact about Mary’s performance is causally responsible for my being surprised. In general terms, from a statement of the form “ $EV_1$  is the cause of  $EV_2$ ”, where “ $EV_1$ ” and “ $EV_2$ ” represent eventive expressions, it is impossible to extract the information as to which of the properties of the event designated by “ $EV_1$ ” was efficacious in bringing about the event designated by “ $EV_2$ ”.

We have also found that the reverse deduction is not always possible either. Not all factive causal statements make explicit reference to the properties involved in the causal relation. This is the case, in particular, when the predicate of the factive expression designating the cause has a negative or relational meaning. In those cases, it is impossible to construct eventive expressions designating cause and effect on the basis of the form and content of the corresponding factive expressions. For example, from the statement “the fact that electric charge  $a$  is at 10 cm distance from electric charge  $b$  is causally responsible for the fact that  $a$  repels  $b$  with force  $F$ ”, one cannot infer an eventive statement explicitly designating the events involved in the underlying process of transference.

The concept of law of nature has borne considerable weight in all steps of our reconstruction. Notably, we have analysed causation in terms of transference with the help of the notion of a conservation law; and the concept of causal responsibility has been analysed with the help of the notion of the instance of a law in a concrete situation. It was, therefore, crucial for our undertaking to take a stand on the nature of laws. Two facts have guided us in our analysis of the concept of law. First, law statements are *opaque* whereas statements expressing accidental generalizations are *transparent*. In statements of these types, when one substitutes one predicate for another referring to a different concept with the same extension, law statements do not generally preserve their truth value, whereas accidental generalizations do. What is more controversial, and what we have tried to justify at some length (Chapter 3), is that typical laws can be valid and nevertheless have exceptions. This means that for most laws there are exceptional situations, in which they are not instantiated, or as one can also say, to which they are not applicable.

None of the traditional conceptions of laws succeeds in accounting for both of these facts. Our own hypothesis is that a law of nature is a universal relation between property instances. This conception allows us to account for both problematic aspects of laws by making the following semantic hypothesis. Insofar as the generalization

All mammals have red blood

is not understood as being nomic, in other words insofar as it is not understood as expressing a law of nature, it should be interpreted in an extensional way: the generalization bears on the concrete objects in the extensions of the concepts expressed by its predicates, in our example mammals and beings with red blood, rather than on the concepts themselves. In the case of generalizations expressing objectively accidental and non-nomic relations, one *must* stick to this extensional interpretation. With a statement such as

All coins in my pocket are of silver,

the extensional, or accidental and non-nomic, interpretation is the only appropriate one; it would be clearly wrong to interpret the relation between being a coin in my pocket and being of silver as lawful.

Frege's thesis that within opaque contexts, words do not have their ordinary reference, has suggested to Dretske the idea that if a universal statement is interpreted as nomological, its predicates occupy opaque positions. The nomological interpretation can be made explicit by prefixing the expression "it is a law that" to a general statement: it has the effect of putting the statement in an opaque context. Bringing these two ideas together leads to the following analysis of the meaning of law statements.

Interpreting a generalization as nomological means to assign an unusual reference to its predicate expressions: in a law statement, predicates refer to the properties themselves rather than to the concrete objects possessing those properties. This unusual reference of predicate expressions is due to their being embedded in an opaque context. This account explains why the substitution, in a nomological statement, of one predicate for another, where both express coextensive concepts, does generally not preserve the truth value of the statement (except if these concepts are identical). By the same token, it explains why substitution of coextensional predicates does *not* change the truth value of non-nomic generalizations: these bear on the objects in the extension of the concepts expressed by the predicates, and these objects remain unchanged by the substitution. This analysis of the difference between nomological and accidental universal generalizations explains why it is not valid to infer the existence of a law subordinating property *G* to property *F* from the universal generalization that all *F*s are *G*s.

This analysis of exceptions shows that the converse inference is not valid either, at least in its usual form: From the fact that *G* is subordinated to *F* by law, one cannot validly infer that all *F*s are *G*s. The key to understanding exceptional situations is the simple fact that the properties related by a law are never instantiated alone in any concrete situation. Other properties, or other instances of properties, can interfere with a given instance of the antecedent property of a law so as to prevent the consequent property from being instantiated. The law may, nevertheless, be valid. The tension between the validity of the law and the existence of situations in which its antecedent but not its consequent is instantiated, can be resolved by making the hypothesis that laws bear directly on pure property instances, and only indirectly on concrete situations. Our conception of laws as universal relations between property instances is realist, insofar as it implies that natural properties exist independently of their designation by predicates of human languages and more particularly of scientific theories. And yet it avoids a certain number of problems that plague conceptions of laws as relations between universals, such as those developed by Dretske, Armstrong and Tooley. For these authors, “all mammals have red blood” is a law if there is a certain relation between the universals *being a mammal* and *having red blood*. We have seen that those difficulties can be avoided by making the alternative hypothesis that the law consists of a universal relation between the *instances* of those properties.

Such a conception of laws has a number of interesting consequences. For one thing, it leads to a new evaluation of the status of the special sciences with respect to fundamental physics. Transference is spelled out in physical terms, but the properties taking part in the relation of causal responsibility may be natural properties belonging to the subject matter of any science. Physics suffices for establishing the very existence of a causal relation between two given events, but special sciences are needed for establishing the causal efficacy of the macroscopic chemical, biological, or psychological

properties of these events. To be capable of causal efficacy, it suffices that a property be natural and non-relational, where being natural means participating in laws of nature. Laws discovered by special sciences bear as much weight in this context as physical laws.

Traditionally, special science laws, and the properties between which they express dependencies, are taken to have no or at least less ontological weight, than the laws of fundamental physics and the properties they bear on; as a consequence, the properties that are the object of special science laws are often taken to play no role in causation. The reason for this conviction is that the laws of the special sciences are taken not to be *strict*, as the laws of (fundamental) physics are, and to hold only “*ceteris paribus*”, or under certain “*provisos*”. This is thought to make special science laws ontologically less reliable. If a law holds only “if all else is equal”, the law statement is at best vague: insofar as the content of the *ceteris paribus* condition cannot be explicitly stated, a “*ceteris paribus*”-law statement does not seem to have any definite content. At worst, its content is as tautological as this: the law applies to all situations except to those to which it does not, because in those, “not all else is equal”. However, the traditional conviction that only physical laws are strict and therefore apt to play a role in the reduction of causation, depends on two theses: first, that physical laws are not subject to *ceteris paribus* clauses, in other words that the validity of physical laws is not limited by conditions of applicability. Second, that the fact that a law has exceptions deprives it of ontological relevance: a law that can be stated only with a *ceteris paribus* clause does not justify considering the properties it links as natural and capable of causal efficacy.

I have spent considerable efforts (in Chapters 3 and 4) in refuting these two theses. If I have succeeded in establishing that there is no fundamental difference between special science laws and laws of fundamental physics insofar as their application to concrete situations is concerned, this shows that it is just as correct to consider special science properties as natural and causally efficacious as it is to dignify fundamental physical properties with those characteristics. Our conception of laws as universal relations between property instances has led us to a conception of the application of laws to concrete situations that undermines the traditional dichotomy between *ceteris paribus* and strict laws. We have found that there are exceptional situations not only for geological, biological or psychological laws, but also for many physical laws. Insofar as special sciences establish relations of nomic dependency between properties, there is as much reason to consider these properties to be real and capable of taking part in relations of causal responsibility, as there is with respect to fundamental physical properties. Finding those laws and thereby those natural properties is a task that must be left to empirical science.

No doubt, our conception of laws and of the situations to which they apply, of microphysical and macroscopic properties, of eventive causal relations reduced to transference and causal responsibility, raises many questions

we have not answered in this book. One of the tasks suggested by the present study is to analyse the relations between macroscopic properties and underlying microscopic properties, and more particularly the relations between the natural properties playing the role of the causally efficacious property  $F$  and the underlying amounts of conserved quantities playing a role in transmission. The former must certainly *supervene* on the latter: this is just to say that there cannot be any difference with respect to the macroscopic properties of two objects or events without there also being a difference between their microscopic properties. But it is often suggested that being a natural or scientifically respectable property also requires that it be *reducible* to underlying microscopic properties. Reducibility is certainly a stronger requirement than supervenience: In order for  $F$  to be reducible to a set of properties  $M_1, \dots, M_n$ , there must be a theory grounding the inference that a given object has  $F$  on the basis of a description of the object's microscopic parts in terms of their microscopic properties  $M_1, \dots, M_n$ . Without going any deeper into this topic, we may note that our conception of causally efficacious properties does neither require reducibility in this sense nor entail that reducibility is incompatible with a property's causal efficacy. Reducibility is not necessary for efficacy because it suffices for being efficacious that a property participate in nomic dependencies with respect to other properties belonging to the same level in the micro-macro hierarchy. But neither does theoretical reduction deprive a property of efficacy, by making it ontologically superfluous with respect to the underlying properties of the reduction base. On the contrary, reduction provides additional justification for taking a property to be natural by showing how its existence is grounded on laws.

All intrinsic, in other words non-relational, properties possessed by an event can be causally efficacious to the extent that they figure in laws. A property  $F$  is causally efficacious if and only if, some event's being  $F$  can be a causally responsible fact. Psychological properties are among those macroscopic properties whose causal efficacy is a major issue in contemporary philosophy. It is important for us to know whether the fact that we believe and desire certain things can causally contribute to our behaviour. There is a long-standing controversy about whether facts about human rational agents, attributed to them in a mental vocabulary, are relevant for causation in the same sense as facts attributed to humans in a neuroscientific vocabulary. Our analysis of the relations between events, laws, causation and causal responsibility prepares the stage for a fresh inquiry into the nature and causal role of the properties identified by psychology.

One important debate on this topic bears on whether commonsense psychology is capable of being refined by a scientific psychology that shares with it, its fundamental convictions, or whether its most fundamental assumptions will turn out to be illusory. Paramount among these assumptions is our conviction that our beliefs and desires bearing on the world and thus the field of influence of our actions, have a causal impact on these

actions. Finding out whether this conviction is well grounded means nothing else than finding out whether, and how, the mind influences causal interactions. Can mental states and events such as beliefs and desires be causes? More precisely, can they be causes by virtue of being mental states, in other words by virtue of the properties that make them specifically mental? This is the problem of mental causation. This problem is closely related to another problem bearing on the participation of mental properties in causation: Is it possible to explain the fact that our propositional attitudes, such as our beliefs and desires, or other intentional mental states, have a well determined content, on the basis of the causal interactions that link the subject to its environment? This is the problem of the origin of the content of mental representation.

The concept of causal responsibility is sufficiently general to cover all interactions due to nomological properties, be they physical, biological or psychological. However, before we can treat mental properties in the same way as properties studied by other special sciences, we must address the issue of whether mental properties, and in particular propositional attitudes, are intrinsic or whether they are irreducibly relational. To the extent that psychology discovers laws linking mental properties to each other, to sensory input and behavioural output, it justifies considering mental properties as real and causally efficacious. There is no other reason for taking as real a gas's property of having a certain pressure: if the gas is contained in a bottle, its pressure can be efficacious in making the bottle burst, and what is efficacious is real. However, there are strong arguments against taking intentional, i.e. contentful, mental properties to be causally efficacious. Our account gives special importance to the "externalist" thesis that content bearing mental properties are irreducibly relational. This would prevent them from being involved in causal responsibility in the same way as intrinsic properties.

An analogous inquiry could be developed with respect to causal explanations in biology. The conceptual framework of causal responsibility prepares us to ask relevant questions. One can argue that the phenotype of an animal is real because it is efficacious in the process of natural selection. But can it be literally true that the fitness of a given population is causally responsible for its reproductive success? The answer to this question depends crucially on whether fitness can be taken to be an intrinsic property of the animals in the population or whether fitness is fundamentally a relation between the animals and their environment.

Among the achievements of a philosophical inquiry it is not the least if it points to new and fruitful ways of addressing traditional questions. My hope is that the analysis developed in these pages may help clarify such pressing questions as those regarding the reality and causal efficacy of macroscopic physical properties as well as biological and psychological properties, in the context of the undoubted efficacy of the microphysical properties underlying them.

# Notes

## 1 What is a causal relation?

- 1 Neither (S) nor (O) and (O') should be considered as *definitions* of causation. Rather, they should be understood as a posteriori hypotheses that might turn out to be false on empirical grounds, independently of the possibility that they may have undesirable conceptual consequences.
- 2 Commenting on Hume, I shall stick to his terminology according to which causes and effects are “objects”, at least when there is no risk of confusion. Later, I shall argue that it is *events*, and in particular “temporal parts” of objects, rather than these objects themselves, that constitute the ontologically fundamental terms of the causal relation. However, we shall see later in the book that the crucial relation in the context of explanation is *causal responsibility*; its terms are *facts*.
- 3 What reason could there be to postulate an entity that is incapable of taking part in causal relations? We are here expressing the idea that the identity of objects is determined by their causal interactions (cf. Kistler 2002). The origin of this idea goes back at least to Plato: “I suggest that anything has real being that is so constituted as to possess any sort of power either to affect anything else or to be affected [ . . . ] I am proposing as a mark to distinguish real things that they are nothing but power” (Plato, *Sophist* 247d-e, 1961, p. 992).
- 4 In this sense, the scope of Russell’s and Carnap’s criticism of the concept of causation is not as large as the scope of Hume’s. Russell and Carnap insist on the fact that the concept of causation belongs to the conceptual scheme of commonsense, but has no place in the conceptual scheme elaborated by the most advanced sciences, and particularly by mathematical physics. However, and contrary to Hume, they do not contest, over and above that criticism, the objectivity of the concept of causation as it figures in the conceptual scheme of commonsense.
- 5 This argument will be criticized later in this chapter, on pages 40–1.
- 6 In English, there are two main syntactic constructions allowing to make reference to events. One is a gerundive construction as in “the leaf’s touching the cat”, the other is nominalization, such as in “the cat’s leap” (cf. Kistler 1999). Some languages, such as Italian, have more syntactic structures referring to events than English; others, such as French, have less. For a linguistic analysis of expressions referring to events in English and Italian, see Zucchi (1993).
- 7 See Ducasse (1926), Anscombe (1971), Gasking (1955), Aronson (1971a; 1971b).
- 8 Cf. Carnap (1931, p. 435). He elaborates it in Carnap (1934, p. 180ff. and 210ff.; English trans. p. 237ff. and 284ff.).
- 9 See Russell (1912) and Aronson (1971b) for a criticism of Hume’s identification of causation with a necessary relation.



- 10 The interpretation of natural necessity is of course controversial. Cf. page 23.
- 11 This argument is of course grounded on the premise – for which Ducasse offers no argument – that the notion of causation is not reducible to that of nomic relation. This is, however, the leading idea of one of the most important philosophical accounts of causation. Cf. Popper (1934), Carnap (1966), Hempel (1966), and Chapter 5 below. This account defends a thesis exactly opposite to Ducasse's: Popper, Carnap and Hempel reduce causation to the instantiation of a law, and the law to an objective universal regularity. They reduce the *causal* character of a general proposition to its *generality*. In other words, they consider the concept of causation to be redundant, beside the notion of nomic regularity. The nomological account of causation will be criticized later in this book (Chapter 5). This will lead us to agree with Cummins that we should “abandon the Humean doctrine that to be causally connected is simply to be an instance of a nomological generalisation” (Cummins 1983, p. 3).
- 12 Cf. Anscombe (1971, p. 81) and Heathcote and Armstrong (1991). Heathcote and Armstrong indirectly make reference to the analogous incapacity of accounts of laws that reduce them to regularities, to account for the possibility of laws that are instantiated only once. (In the context of the analysis of Hume's position, we can ignore the distinction between a law and a causal law. We shall come back to this distinction in Chapter 6 below.) They defend their “strong” interpretation of laws by arguing that it can explain, whereas the regularity theory cannot, “that a particular positive instantiation of a law might be an instantiation of that law despite the non-existence of any other instantiations” (Heathcote and Armstrong 1991, p. 64).
- 13 Tooley (1990b) has recently defended a singularist theory of causation; however, Tooley considers causation as a theoretical concept and not, as Ducasse, as an observable phenomenon.
- 14 For criticisms of this aspect of Ducasse's definition, cf. Sosa and Tooley (1993, p. 17–19), Fair (1979, p. 227) and Tooley (1990a, p. 287f.).
- 15 Ducasse's concept of event is compatible with ours: if an event is what fills a space–time zone, *S* is the spatial surface around the zone and *I* its temporal location.
- 16 Castaneda (1984, p. 20) proposes another counterexample. He imagines a mechanism moving a ball *A* at intervals whose length is random or determined by external factors completely unrelated to the movement of a second ball *B*. However, in a given particular situation, *B* happens to arrive at a position contiguous with *A*, just at the moment at which the mechanism moves *A*. The mechanism is constructed in such a way as to prevent *B* from exercising any causal influence on *A*. By virtue of the construction of this thought experiment, the mechanism is clearly the only cause of *A*'s movement. Nevertheless, Ducasse's analysis delivers the wrong result that is on the contrary *B*'s arrival at the position contiguous to *A* that causes its movement.
- 17 This is not the place to dwell upon the historical origins of this aspect of the concept of causation. Since Aristotle, causes have been understood as what allows for *explanation* of events and facts. Now, in ordinary circumstances, one feels no need to explain permanent states involving no change. More precisely, in the framework of an Aristotelian conception of causation and explanation, it is not necessary to look for causes in cases where the absence of change is *natural*. In Aristotelian terminology, we may say that philosophers like Ducasse consider the absence of change to be natural, and therefore to require no explanation in causal terms.

It is however interesting to note that Ducasse begins by considering both changes and the absence of a change as events (Ducasse 1926, p. 126). I shall presently mention a reason for which Ducasse cannot take advantage of this wide

notion of an event, which would allow him to conceive of stable states as causes. With Mill, Carnap is one of the few authors whose philosophical notion of cause does not require it to be a change. Speaking of a table, Carnap says: "This event, the table existing without change, is also a process. It is a *static* process, one in which the relevant magnitudes remain constant in time. If processes or events are spoken of as involved in cause-effect relations, it must be recognised that these terms include static processes" (Carnap 1966, p. 190; italics in original).

- 18 I shall later defend the following position with respect to this question: The temporal precedence of the cause (and the precedence of the effect in hypothetical cases of "backwards causation") is a contingent property of causal relations. It is not a priori determined on conceptual grounds. Whether there are, in the actual world, backwards causal relations, in which the effect precedes the cause, is a controversial empirical question. However, I shall argue that there is no simultaneous causation: if two spatially distant events are causally related, they are never strictly simultaneous. If two things happen at the same place, at the same time there can only be lawful relations between them, but no causal relation. Contrary to the answer to the former question, my answer to the question whether there are simultaneous causal relations is justified on *conceptual* grounds.
- 19 I shall return in Chapters 5 and 7 of this book to the issue of the ontological status of the terms of the causal relation, and in particular to the causal roles played respectively by events and facts.
- 20 Gasking holds that this asymmetry is not grounded in any conceptual necessity, but is simply a fact about the actual world that we discover by empirical means. To argue for this thesis, he describes a conceptually possible world in which the causal order between the change of temperature and the change of luminosity is inverted. Tooley (1990a, 1990b) and Sosa and Tooley (1993) argue in a similar way that most "analyses" of the concept of causation are not conceptually adequate because one can conceive of possible worlds in which causation does not fit them. However, such analyses can be adequate if they are understood as a posteriori hypotheses about what causation is like in the actual world. I shall later deny that the relation between the changes of temperature and colour is causal. It is a case of nomic but non-causal dependence. The appearance of causality stems from the fact that the way to bring about both changes is by transferring heat to the piece of iron. This transference is the cause of both changes which are nomically interdependent.
- 21 It is undeniable that our epistemic access to causation is closely tied to our experience of ourselves as agents. However, acknowledging this fact does not prejudice in any way the ontological question of the nature of causation.
- 22 Hausman criticizes anthropocentric or, in his terms, "manipulability theories" of causation because they are "in need of manifold qualifications, narrowly circular and unduly anthropomorphic" (Hausman 1986, p. 150).
- 23 Cartwright (1983); Price (1991, 1992); Menzies and Price (1993).
- 24 See also von Wright (1971, pp. 73–74).
- 25 On the relation between logical and natural modality, see also Carnap (1936) and von Wright (1984).
- 26 Hacking (1983) and Cartwright (1983; 1989) have recently made a strong case for the importance of experimental intervention for the acquisition of scientific knowledge.
- 27 Cf. Aronson (1971a, 1971b, 1982), Quine (1973), Fair (1979), Dowe (1992a, 1992b).
- 28 This can be learnt from Tooley's (1990) and Sosa and Tooley's (1993) general argument against all theories of causation that reduce causation to some particular type of physical relation. They show that transference theory cannot pre-

tend to accomplish a perfectly general conceptual reduction of the concept of causation. Indeed, (S) has the status of a hypothesis about the conditions of application of this concept in the actual world. It says that all causal relations in the actual world are cases of transference. In other words, (S) is a hypothesis about the real essence of causation, conceived of as a natural kind of relation. We shall examine Tooley's arguments a little later (on pages 35–7).

- 29 See pages 42–3.
- 30 Special relativity imposes the constraint of acting locally on all forces of interaction. It can be proven that any hypothetical degree of freedom of a system, which would be determined non-locally, would be in contradiction with Lorentz's principle of invariance, which is the fundamental postulate of special relativity. See Kerner (1972).
- 31 This example, due to Bell (1976), illustrates the difference between the quantum correlation studied by EPR from correlations at a distance that are due to perfectly local processes of propagation.
- 32 "So to speak", because all acquisition of knowledge is, in a sense, "simultaneous" with its intentional object, i.e. the knowledge acquired.
- 33 Bunge (1959) has put forward the idea that causation is one particular form of determination among many others. Whether EPR type phenomena constitute cases of causation at a distance is hotly debated in the literature. See Skyrms (1984), Scheibe (1991), Chang and Cartwright (1993).
- 34 Among the important explanations of the origin of the asymmetry of causation that have been proposed, I should mention Reichenbach's (1956) and Salmon's (1984) thesis that it is grounded on the intrinsic asymmetry of "open forks". Other authors who try to reduce causal asymmetry to the asymmetry of conditional probabilities, are Suppes (1970), Cartwright (1979), Eells and Sober (1983), Tooley (1987), Carroll (1991). I shall briefly examine Tooley's proposal later (cf. pages 180–5). According to Popper (1956–7), the directions of time and causation are determined by the direction of the "propagation of order". It seems to me that this suggestion is ultimately grounded on the fact that the entropy of isolated systems can only increase, and is therefore compatible with the account put forward in this book. The increase of the entropy of isolated systems is indeed the most important kind of process (though not the only one) grounding the asymmetry of the causal relations of the actual world. We shall come back to this topic later in this section. However, those analyses according to which the asymmetry of causation stems from the asymmetry of the cause and effect of human actions are not up to the task of building an account of causation as an objective phenomenon. Cf. pages 21–5. Finally, Lewis' (1986) analysis of causation in terms of counterfactuals makes it necessary to *stipulate* the existence of an asymmetry between cause and effect, because counterfactuals which are made true by nomological dependencies (or laws of coexistence) are symmetrical, just as are the laws making them true. Lewis tries to overcome this problem by denying that there are so-called backtracking counterfactuals, in which the consequence bears on an event happening before the event mentioned in the antecedent. See note 45 below.
- 35 Cf. Horwich (1987). Physical space is not strictly isotropic. The space of general relativity is neither isotropic nor homogeneous. However, if one looks at local interactions, one can consider space as locally isomorphic to a classical Newtonian space: in portions of space containing no singularities, it is approximately true that space has three orthogonal dimensions none of which is objectively privileged with respect to the others. In this sense, far from singularities, space is locally isotropic.
- 36 See also Dieks (1986, p. 95) and Horwich (1987).

- 37 The term “lineal” denotes the property of the network to possess a unique direction. By contrast, the events situated on a single line of the network are in a *linear* order: it is possible to characterize each of them univocally by attributing it a real number. To attribute time in general, and therefore to the whole network, a linear order, a concept of simultaneity is also required. Cf. Reichenbach (1956, p. 39, note 3).
- 38 Such a system need not *always* be isolated and can therefore be integrated into the causal net.
- 39 There are also known to be intrinsically asymmetric microphysical processes, in particular the disintegration of K-mesons or “kaons”. Such processes of decomposition are said to “violate” the symmetry with respect to temporal inversion (“T”). The reason for thinking that a given process is not reversible, or symmetrical with respect to time, is rather indirect: One begins by observing that the process violates symmetry with respect to the operations of charge conjugation (“C”) and of spatial inversion (or parity, “P”). On the basis of the so-called CPT-theorem, according to which all interactions are symmetrical with respect to the joint operation of conjugation of charge (“C”), spatial inversion (“P”) and temporal inversion (“T”), one concludes that the processes of decomposition of K-mesons also violate the symmetry with respect to time. Cf. Lee *et al.* (1957), Sachs (1987), Dowe (1992b, p. 189).
- 40 Cf. Horwich (1987, p. 138).
- 41 The idea that asymmetry is not contained in the concept of causation is also implicit in Grünbaum’s (1973, p. 190) thesis that there is no contradiction in the concept of a world in which causal relations are symmetric.
- 42 This is however what Dieks (1986), Ehring (1986) and Dowe (1992a, 1992b) conclude, because they take it to be unacceptable to suppose that the direction of a causal relation may be determined indirectly, by *other* causal relations containing intrinsically asymmetrical processes.
- 43 There is an important literature on the question whether there are such processes. Cf. Earman (1976), Grünbaum and Janis (1977), Horwich (1987, Chapters 6 and 7), Dowe (1996). Without going into this debate, here we cannot do more than establish, as we have done above, that the account of causation we have developed is compatible with the possibility of their existence, and show how one can make sense of the notion of a “causal process evolving backwards in time” in this framework.
- 44 In the possible worlds conceived by Tooley, all bases of reduction of causal asymmetry yield this result. They do not, e.g., exhibit any privileged direction of the “propagation” of order (Popper 1956–7) nor any “open forks” with an intrinsic direction (Reichenbach 1956; Salmon 1978).
- 45 Lewis (1979, p. 49ff.) reaches the same conclusion concerning the temporal symmetry of a “simple world”. He explains why we evaluate differently counterfactual conditionals bearing on the dependence of the future on the present and counterfactuals bearing on the dependence of the past on the future (so-called “backtracking” counterfactuals), with the help of the temporal asymmetry of what he calls “tiny miracles”. A tiny counterfactual deviance from actuality has many effects: in this sense, the future overdetermines the past. But a single tiny event suffices for these multiple effects: in this sense, the past does not overdetermine the future. It is because the future overdetermines the past that a great miracle is needed for “changing” the past, i.e. for compensating a change that has happened in the past. It is because the past determines the future without overdetermining it that only a tiny miracle is needed to “change the future”. For a criticism of Lewis’ argument for the general falsity of all “backtracking counterfactuals”, cf. Kistler (2002a).

- 46 As we have already noted, this direction can also be determined by other intrinsically asymmetric processes like the decomposition of K-mesons, in case they exist in the world under consideration.
- 47 Beauchamp and Rosenberg (1981, p. 210); Ehring (1986, p. 255); Dieks (1986, pp. 106–107); Dowe (1992a, p. 203; 1992b, pp. 181–182); Sosa and Tooley (1993, p. 4). Cf. also Papineau (1985, p. 273).
- 48 Cf. Fair (1979, pp. 242–243).
- 49 Price (1992) and Menzies and Price (1993) reach the same conclusion. They take causation to be a secondary quality.
- 50 On this point, see Chapter 5, pages 139–44. We shall come back to the example of the ice cube in Chapter 7, pages 196–8.
- 51 However, he does not take this argument to be conclusive: before he presents it, Hume notes “that experience in most instances seems to contradict this opinion” (Hume 1955, p. 76), i.e. that there are cases of simultaneous causation. And after having presented the argument we are about to analyze he adds: “If this argument appear satisfactory, ’tis well. If not, I beg the reader to allow me the same liberty, which I have us’d in the preceding case, of supposing it such. For he shall find, that the affair is of no great importance” (Hume 1955, p. 76).
- 52 Russell uses a version of this argument in the context of his demonstration that a dilemma results from conceiving the cause to have finite duration: if the cause contains a change, one must conclude that only its last temporal part really plays the role of the cause, and in fact only an infinitesimally short part. But if – and this is where he takes up Hume’s argument we are concerned with – “the cause is purely static, involving no change within itself, then, in the first place, no such cause is to be found in nature, and in the second place, it seems strange – too strange to be accepted, in spite of bare logical possibility – that the cause, after existing placidly for some time, should suddenly explode into the effect, when it might just as well have done so at any earlier time, or have gone on unchanged without producing its effect” (Russell 1912, p. 196).
- 53 Later we shall contest that the latter two relations are causal and that the wagons’ beginning to move is strictly simultaneous with the locomotive’s.
- 54 If gravitation is usually presented as a non-local interaction, this is merely due to the simplicity of Newton’s formulation of the law of gravitation. However, other mathematical expressions of the law represent the fact that gravitation is a field that propagates with a finite speed. The two formulations are strictly equivalent as far as the resultant forces are concerned (cf. Einstein and Infeld 1936, pp. 223–224 and pp. 227–228; Feynman 1965, p. 51). Action at a distance is not a consequence of the law of gravitation itself but only an accidental feature of its particular Newtonian formulation. The same is true for the classical representation of electrostatic interaction: in Coulomb’s representation, electric charges act on each other simultaneously at a distance. But this representation is known to be an approximation, which is useful in the context of macroscopic electrically charged objects moving at low speeds. When a precise representation is needed, one uses an expression in which the causal influence of one charge on the other is strictly local and spreads at a finite speed: Coulomb’s potential is then replaced by a Liénard-Wiechert potential. Cf. Jackson (1975, p. 656). Poincaré was the first among post-Newtonian physicists who defended the thesis that gravitation does not act simultaneously at a distance. “The propagation of gravitation is not instantaneous but happens with the speed of light.” (Poincaré 1906, p. 496). Einstein has shown that the possibility that gravitation propagates faster than the speed of light would entail the possibility of backwards causation. It would have the consequence “that we would have to consider a transmission mechanism possible by the use of which the effect would *precede* the cause (which may be accompanied by an act of will). Even though this result does not, in my opinion,

contain any contradiction from a logical point of view, it contradicts so radically the character of the whole of our experience that it suffices to establish the impossibility of the hypothesis  $W > V$ ” (Einstein 1907, pp. 381–382; italics Einstein’s. “ $W$ ” is the hypothetical speed of the propagation of gravitation; “ $V$ ” is the speed of light.) Poincaré takes the impossibility of backwards causation to be, not a contingent fact, but rather a presupposition of our spontaneous conception of causation: “We shall try to arrange that  $t$  always negative; if indeed one conceives that the effect of gravitation takes a certain time to propagate, it would be more difficult to understand how the effect could depend on a position that the attracting body has not yet reached” (Poincaré 1906, p. 540).

- 55 The force due to gravitational attraction is weaker than the force due to electrostatic attraction by an order of magnitude of  $10^{20}$ .
- 56 Later (pp. 50–2), we shall return to the conception of concrete objects as sequences of events, which are their “temporal parts”.
- 57 We shall come back to the ontological status of the terms of the causal relation in Chapters 5 and 7, where we shall analyse the relation between the concept of causation between events developed in the present chapter and the concept of causal responsibility between facts. Only after having analysed the relations of logical entailment between statements reporting causation among events and statements reporting relations of causal responsibility among facts will it be possible to find out which of the two notions is the most fundamental one. We shall reach the conclusion that causal responsibility between facts is ontologically less fundamental even if it plays a primordial role for causal *explanations*.
- 58 Quine (1985).
- 59 The statement “ $A$  is cause of  $A$ ” may be given yet another sense in which it can be true: the variable  $A$  can be understood to range over the domain of objects rather than events or facts. In this case it is possible to understand “ $A$  is cause of  $A$ ” as meaning that  $A$  at a given instant  $t$  is causally related to itself at a later instant. The statement is then taken to express the causal theory of the identity of objects through time. We shall return to this point later.
- 60 Cf. Sperber and Wilson (1986).
- 61 An analogous analysis applies to the case mentioned earlier, of the change of temperature of a piece of iron. Its temperature is nomically linked to the radiation it is emitting; therefore raising its temperature can make that radiation visible, in other words make the iron glow. Nevertheless, the relation between the iron’s being hot and its glowing is non-causal nomic dependence. What is causal and what makes that relation itself appear as causal, is the process of transference of heat to the piece of iron, which causes the change of both temperature and the quality of the radiation emitted, which thereby becomes visible.
- 62 An object is said to be concrete if it fills a determinate portion of space. No property other than those possessed by the object itself can be instantiated at that position without modifying the object itself.
- 63 Russell does not use the term “genidentity”. Later (Russell 1948, p. 333ff.), he introduces the concept of a “causal line” that he uses to explain the persistence of objects through time. However, Russell’s conception is very different from the one developed here in that it is epistemic rather than ontological. “A ‘causal line’, as I wish to define the term, is a temporal series of events so related that, given some of them, something can be inferred about the others whatever may be happening elsewhere” (Russell 1948, p. 477). Other passages in Russell’s work (cf. Russell 1948, p. 476ff., 500, 510) make it clear that causal lines are determined in terms of our capacity to recognize the same object at different times, rather than in terms of its *being* the same object independently of any observers.
- 64 Here is how Reichenbach characterizes the notion: The “*physical identity* of a thing, also called *genidentity*, must be distinguished from logical identity. An

- event is logically identical with itself; but when we say that different events are states of the same thing, we employ a relation of genidentity holding between these events. A physical thing is thus a series of events; any two events belonging to this series are called genidentical. The relation of genidentity is therefore a two-place propositional function which is symmetrical, transitive, and reflexive.” (Reichenbach 1956, p. 38). Grünbaum speaks of “material objects each of which possesses genidentity (i.e. the kind of sameness that arises from the persistence of an object for a period of time) and whose behavior therefore provides us with genidentical causal chains” (Grünbaum 1973, p. 189).
- 65 Cf. Carnap (1928, p. 170; trans. p. 199).
- 66 In his own analysis of the concept of genidentity, Carnap does not consider this possibility. However, Grünbaum (1973) and Armstrong (1980) defend a “causal theory of identity through time”, which is a relational theory of the diachronic identity of objects, against a rival theory according to which the identity of objects through time is primitive and not based on any relation between its temporal parts. “We should continue to favor a relational account of identity through time, appealing in particular to the relation of causality between the successive phases” (Armstrong 1980, p. 77). Russell (1914, p. 108ff.; 1948, p. 333f., 476f.) has drawn attention to the fact that the concept of substance is not appropriate within the framework of a theory that conceives diachronic persistence as a causal process.
- 67 Aronson (1971a) expresses the same idea with the help of his distinction between natural and non-natural changes: just as inertial movement, the persistence of an object through time belongs to the class of “changes that take place in individuals that can be completely accounted for without having to appeal to the behaviour of other individuals” (Aronson 1971a, p. 421). Aronson calls such changes “‘natural changes’” (Aronson 1971a, p. 421). For him, only ‘unnatural’ changes, in other words changes due to interactions, are causal.
- 68 Cf. Quine (1973, p. 6); Ehring (1986, p. 256); Dieks (1986, p. 87ff.); Dowe (1992a, p. 202f. and 214; 1995b, p. 369).
- 69 Cf. also Reichenbach (1928, 1956). We come back to this distinction in the next paragraph.
- 70 They are conserved only within an certain interval of precision whose size is determined according to Heisenberg’s relation of uncertainty. Cf. Chapter 4. The expression “uncertainty” that is traditionally employed in this context may be misleading: according to quantum mechanics, this “uncertainty” is not a limitation of the precision of the *knowledge* we may obtain of the value of the energy *E* of a system, but the *objective* limit to which the value of *E* is determined. Cf. Bohm (1962).
- 71 This presupposes that all interactions are local. Cf. pages 27–9.
- 72 Cf. Quine (1973, p. 6); Salmon (1984, p. 146); Ehring (1986, p. 256); Dowe (1992a, p. 202f. and 214).
- 73 Cf. Reichenbach (1928, p. 307–308; trans. p. 135–138); Reichenbach (1956, sec. 23), Salmon (1978, p. 690; 1984, p. 148).
- 74 Russell (1948, p. 333f. and 476f.) introduces the concept of “causal line”, which is similar to that of causal process. See Russell’s definition of “causal line” quoted above, note 63. We have already noted (cf. note 63) that the concept introduced by Russell differs from that of a causal process in a fundamental respect. The concept of causal line is epistemic rather than ontological. It is defined in terms of the *knowledge* that we have of the events under consideration. Cf. also Russell (1948, pp. 500 and 510). However, the concept of causal process is not relative to knowledge. If one tried to take Russell’s concept in an ontological sense this would have the undesirable consequence that certain pseudo-processes would count as causal. Cf. Salmon (1984, p. 144). In particular, there are pseudo-processes

exhibiting nomic regularity (cf. pages 60–2). If the relevant law is known, it is in their case possible to infer the description of certain events in the pseudo-process from the description of others. Nevertheless those events are by hypothesis not causally related.

- 75 The fact that the propagation of the phase of a wave obeys to a law allows us to mention it in a scientific explanation. The non-causal character of the propagation of the phase is compatible with its integration in the scientific theory of wave phenomena. Our account can be seen as generalizing Sober's (1985) thesis that *certain types* of pseudo-processes are scientifically respectable, i.e. those composed of events which are indirectly related by an underlying causal process. Sober's thesis allows him to account for the intuition that the parent's phenotype may explain the offspring's phenotype without being its cause: the parent's genotype is the common cause of the parent's and the offspring's phenotypes. However, Sober's suggestion that "perhaps we should conclude that pseudo-processes are explanatory only if the events they link are also connected by causal processes" (Sober 1985, p. 308) wrongly excludes the explanation of the position of a wave phase in terms of an earlier position and the phase speed. Pseudo-processes governed by laws can provide non-causal explanations even if no causal process underlies them.
- 76 On this condition, cf. Salmon (1984, p. 141).
- 77 We have already accounted for the origin of the difference between these two cones, in the sense that one contains the future and one the past, from the point of view of event *O*.
- 78 Our very general conception of events allows us to conceive of such a pseudo-process. According to most rival accounts of events, its constituents are not events at all because "nothing happens" in them, so that one cannot make reference to them with predicates expressing change.
- 79 Kitcher (1985) and Dowe (1992b, pp. 200–201) have noted the circularity of the definitions of "mark" and of "causal interaction". Salmon (1994) has acknowledged this circularity and therefore abandoned his earlier account.
- 80 Dowe repeatedly says that "the precise characterization of 'object' is unimportant" (1992c, p. 126) or that "what counts as an object is unimportant; any old gerrymandered thing qualifies" (Dowe 1995b, p. 371). These statements are contradicted by his own efforts (Dowe 1995a, pp. 326–331) to exclude temporally arbitrary assemblages from the class of genuine objects. As we have seen, these efforts are not successful. Apart from the suggestion we have considered above that an object must be "wholly present at a time", Dowe quotes with approval Quine's (1965, pp. 229–231) suggestion that a physical object is "an intrinsically determinate portion of the spacetime continuum" (Dowe 1992c, p. 126). But *this* is certainly too weak to exclude either the light spot or the series of illuminated portions of the wall from the class of objects.
- 81 Curiously, Dowe says of the light spot that the reason for which it does not count as a causal process is "that no energy is [ . . . ] carried off by the spot" and that "it cannot transmit a conserved quantity" (Dowe 1992c, p. 127). This seems indeed the correct analysis; but only transference theory is in position to provide it: as long as Dowe denies that individual amounts of energy can be transferred, he cannot characterize the light spot as something that *cannot transmit* energy. Carrying is after all a mode of transmitting.
- 82 These notions have been introduced earlier, on page 60.
- 83 If this argument is correct, it shows that one cannot, as many authors hold following Popper (1934), Carnap (1966) and Hempel (1966), *reduce* causation to the instantiation of a law. We shall see later that laws of nature play an important role in the analysis of causation, for they are an essential constituent of the concept of *causal responsibility*. We shall say that the fact that the cause *c* has



property  $F$  is causally responsible for the fact that the effect  $e$  has property  $G$  if and only if, first,  $c$  and  $e$  are related as cause and effect in terms of transference, and second, there is a law linking  $F$  and  $G$  that is instantiated in the situation. Cf. Chapter 5.

- 84 The conservation law itself is neither causal nor non-causal. Cf. Chapter 6.
- 85 Concrete particulars must be distinguished from abstract particulars, an ontological category to which belong, e.g. particular quantities of energy.
- 86 In Chapters 5 and 7 we shall analyse in more detail the commonsense notion of causation by a way of a semantic analysis of causal statements. It will turn out to have two aspects, only one of which corresponds to the notion of causation between events in terms of transference developed in this chapter.
- 87 The idea that forces are causes is of course older. It is in this sense that Kant interprets Newton's theory and in particular the law that  $F = ma$ . Largeault (1985, p. 112) names several other authors who have explicitly defended the idea that forces are causes. We cannot here enter into a historical analysis allowing us to follow the development of this idea as a function of the notion of force employed by each author. Recently, Bigelow *et al.* (1988) and Bigelow and Pargetter (1990) have proposed a theory of causation in which forces play a major role. Their analysis of the notion of force leads them to the conclusion that "forces are species of the causal relation itself and as such have a different ontological status from the sorts of entities normally considered to be related as causes to effects." (Bigelow *et al.* 1988, p. 614). Zangari (1992) objects to this account that some interactions with observable effects do not bring any forces into play. This is the case with the Aharonov-Bohm effect. Cf. Aharonov and Bohm (1959), Peshkin and Tonomura (1989), Liu (1994).
- 88 The transmission of information cannot ground causation, for the transmission of information requires only reliable statistical dependence between two series of events. Lawful dependencies existing among different effects of a common cause, which could not be linked as cause and effect, create such a dependency, as well as purely accidental correlations. The following situation provides an example of the first type: consider the events constituted by the reception of a radio programme at two points situated in opposite directions from the sender. Given that the signal spreads with the speed of the light, the distance between these events is "spatial" in the sense of relativity theory: it is therefore impossible that one be the cause of the other. Cf. our criterion (N). Nevertheless, information flows from one receiver to the other. The existence of such situations justifies the exclusion of information flow from the processes underlying causation. However, transference theory also indicates the relevant property which information lacks and which it would need to possess to be able to ground causation. Information is not transmitted in the physical sense required for the existence of causal relations because it does not obey to a conservation law.
- 89 Castaneda (1984) proposes the hypothetical "causity". The hypothesis of the existence of a unique property that might be identified with "causity" cannot be tested independently of the phenomenon it is supposed to explain, i.e. causality; therefore it seems to be ad hoc.

## 2 Laws of nature and universal generalizations

- 1 This is widely acknowledged. In Hempel's words, "lawlike sentences cannot be characterized in terms of their form alone" (Hempel 1965, p. 339).

- 2 For this form to be appropriate properties *A* and *B* must be possessed by the *same* object. The analysis can be generalized to cases where a law relates properties possessed by *different* objects.
- 3 This will be justified by our analysis of laws, which accounts for the existence of exceptions. However, from a linguistic point of view, expressing a law in its *generic* (U2) rather than explicitly universal form provides a way to account for the existence of exceptions to the generalization, by virtue of the *vaguer* content of the generic form. Cf. Sorensen (1990). This makes it possible to semantically evaluate generic statements in a way that is sensitive to *context*. Cf. Carroll (1991, p. 259f.). However, I think it is impossible to correctly analyse laws with exceptions by trying, as do Adams (1988) and Carroll (1991), to give a logically precise form to vague or “inexact” generalizations. As Sorensen (1990) notes, construing laws of nature as “vague” generalizations, expressed by generic statements such as “rocks sink in water”, is compatible with the existence of exceptions such as pumice that float on water. However, “the vagueness construal [ . . . ] obscures [ . . . ] [the] logical properties” (Sorensen 1990, p. 617) of statements of law. In the absence of an analysis of the origin of exceptions, Carroll (1991) ends up analysing “*As are Bs*” as meaning “among *As* the frequency of *Bs* is high” (cf. Carroll 1991, p. 257). This is certainly inadequate, if only because it blocks the possibility that a given law has *many* exceptions. Cf. Chapter 6.
- 4 From now on we will often leave the qualification “concrete” implicit: when we speak of objects it will mean *concrete* objects, i.e. particular entities possessing many properties at a time. The concept of object taken in this sense must be distinguished from the concept of property instance or “abstract particular” (Campbell 1990), i.e. the concept of the instance of *one* property. Cf. Chapter 6.
- 5 Cf. Goodman (1955, p. 17ff.); Braithwaite (1959, p. 293ff.); Nagel (1961, p. 58ff.); Hempel (1965, p. 264ff.); Carnap (1966, p. 12ff.); Lewis (1986, pp. 121–126).
- 6 Armstrong (1983, p. 17f.) borrows this example from Reichenbach.
- 7 According to Braithwaite (1959), whether someone considers a given generalization as nomological or as accidental “will then have to depend upon knowledge or belief in the general proposition rather than in anything intrinsic to the general proposition itself” (Braithwaite 1959, p. 301). More precisely, Braithwaite holds that someone considers that a given hypothesis expresses a law of nature if and only if the person considers that this hypothesis can be “deduced from a higher-level hypothesis in a true and established scientific deductive system” (Braithwaite 1959, p. 297). A general hypothesis is considered to be accidental if “there is no established scientific system in which the generalization appears as a consequence” (Braithwaite 1959, p. 304). This makes interpreting a generalization as a law of nature “epistemologically relative” (Braithwaite 1959, p. 307), i.e. relative to the conviction that a given scientific system is, at a given moment, true and well-established.
- 8 This position will be criticized in Chapter 6.
- 9 Cf. Ducasse (1924), Kneale (1949; 1950; 1961), Chisholm (1955), von Wright (1947; 1971).
- 10 Cf. Dretske (1977b), Armstrong (1983), Tooley (1987).
- 11 This argument is due to Dretske (1977b). Popper (1949) expresses its central idea that a non-modal or extensional language is incapable of accounting for the specificity of laws of nature, insofar as laws are not equivalent to universal statements expressing a more or less accidental fact in spite of their universal form. Cf. Popper (1949, p. 63). However, Popper’s account contains several ambiguities. He seems to take the difference to concern the *form* of the statements rather than the relations they express. However, exactly the opposite is the case, for the whole problem takes its origin in the fact that both laws and accidental universal generalizations can be expressed by statements of form (U).

There is a second ambiguity to which we shall come back later: Popper postulates a difference between universality and necessity, in other words between statements which are accidentally though universally true and statements that have modal force. However, later in the same paper, he expresses the idea that modal expressions are *reducible* to extensional expressions: according to Popper, to say that something is possible means that it is real in some spatio-temporal region. But this entails an analogous reduction of necessity: that is necessary which is real everywhere and at all times.

- 12 This is true if one considers only matter but not antimatter, in other words if one does not take positrons – the electrons' antiparticles – into consideration.
- 13 I am here using Quine's terminology: Carnap calls *extensional*, contexts in which existential generalization is valid and in which the substitution of coreferential expressions is possible *salva veritate*, and *intensional* contexts in which these two types of inference are not valid. Quine calls the former contexts *transparent* and the latter *opaque*.
- 14 The concept of a referentially transparent context was first introduced by Frege who attributes the idea to Leibniz (Frege 1892b, p. 49/50; transl. 64). The passage to which Frege refers contains a statement of the "Leibniz law" according to which two expressions designate the same object if one can be substituted for the other in all contexts (more precisely: in all transparent contexts), without thereby changing the truth value of the statement. Leibniz says: "Terms are the *same* [eadem] if one can be substituted in place of the other without destroying truth, as triangle and trilateral, quadrangle and quadrilateral." (Leibniz 1969, p. 241). (*Eadem sunt quorum unum in alteris locum substitui posset, salva veritate, ut Triangulum et Trilaterum, Quadrangulum et Quadrilaterum.*)" (Leibniz 1890, p. 219). It is true that in this definition of identity, Leibniz seems to confuse the term referring to an object with the object itself. Cf. Quine (1960, § 24, note 5). In a transparent context, one can substitute one term for another sharing its reference. It is however not the *terms* that are identical, as Leibniz says, but their reference. Frege (1884, § 65; p. 76f.) says he agrees with Leibniz's explanation of identity, without mentioning this problem. Cf. also Frege (1983, p. 131).
- 15 Cf. Dretske (1977b, p. 250).
- 16 Dretske (1997b, p. 250) uses the word "transparent" to characterize universal material implications, but only by virtue of the fact that their truth value is not affected by the substitution of a predicate by another that expresses a coextensional concept. He does not mention the fact that statements of form (U) are *not* transparent in a strict sense, for extensional generalization is in their case invalid. From the fact that all *As* are *Bs* it does not follow that there are any *As*. See Moravcsik (1991) for the suggestion that the contrary intuition has a pragmatic source.
- 17 Concepts of numbers are of this sort: "If I say 'Venus has 0 moons', there simply does not exist any moon or agglomeration of moons for anything to be asserted of; but what happens is that a property is assigned to the *concept* 'moon of Venus', namely that of including nothing under it." (Frege 1884, p. 59; italics Frege's; transl. p. 59e). Another example of a second-order concept is the concept of unicity corresponding to the property of having exactly one object in its extension. Cf. Frege (1884, p. 65).
- 18 We have already noted that Dretske uses the term "transparent" in the sense we have chosen to give it in the framework of the present chapter: although extensional generalization over statements of form (U) is not valid, we call them nevertheless "transparent" by virtue of the fact that the substitution of predicates expressing coextensional concepts preserves their truth value.
- 19 From a realist point of view it is equivalent to speak, with Dretske, in terms of "properties" or, with Frege, in terms of "concepts".

- 20 Frege introduces this distinction in the context of his analysis of the semantics of indirect discourse. Subordinate phrases put in *indirect* discourse must be evaluated according to the mode appropriate for opaque contexts, whereas phrases put in *direct* discourse must be evaluated according to the mode appropriate for transparent contexts.
- 21 Several contemporary realist philosophers have developed versions of such an ontology, in particular Armstrong (1978; 1983) and Tooley (1987).
- 22 The analysis of the concept of property developed in Chapter 6 tackles questions closely related to the status of laws. If it turns out to be conceivable that properties exist independently of the linguistic predicates denoting them, this adds plausibility to the idea that there may be relations, and in particular nomic relations, between these properties, independent of their expression in statements.
- 23 This position is also defended by Armstrong (1983, p. 88ff.; 1993, p. 422). Armstrong (1992; 1993) also considers, without adopting it, the doctrine according to which universals are to be understood not as singular entities but as *types* of property instances.
- 24 This is why our account is not threatened, at least at first sight, by van Fraassen's (1989, chap. 5) criticism of realist theories of laws. Cf. Chapter 4, pages 127–31.
- 25 It would also be valid for “steel laws”, a term which Armstrong uses to speak of laws that are *necessarily* of iron, in other words, that are necessarily without exceptions. Armstrong himself doubts whether such laws actually exist.
- 26 Fuhrmann (1991) has also suggested conceiving of laws as of universal relations between particular properties, which he calls “tropes”. (We shall return to this term shortly.) However, he develops this idea in a way that makes it useless as a solution for the problem of exceptions. Indeed, Fuhrmann conceives of the nomic relation between tropes according to the model of mereological inclusion. This leads him to the result that if a law links tropes of type *F* to tropes of type *G*, it is logically necessary that every *F* be accompanied by a *G*. Fuhrmann's conception does not allow him to account for the phenomena of exceptions any more than Dretske's (1977b). However, his ontological choice – that we share – to locate the nomic constraint at the level of particular properties, in other words neither at the level of concrete objects nor at the level of universals, allows him to escape from a logical problem that plagues theories in terms of universals: the so-called “inference problem”. Cf. Chapter 4, pages 127–31.
- 27 Coffa (1974) makes a similar suggestion in the context of the justification of the objective validity of statistical inferences.
- 28 In Armstrong's (1978; 1983; 1992) and Tooley's (1987) accounts, properties are conceived of as “universals”, which are entirely present in all their instances. Tooley's conception of universals is Platonist, in the sense of taking universals to be able to exist independently of their exemplification by concrete objects. This entails a disagreement between Tooley and Armstrong about the existence of properties without any instances. Armstrong, whose position is often called “Aristotelian”, denies the existence of universals that are not instantiated at all. Against realism of universals, Lewis argues for a nominalist position according to which properties are just extensions of predicates. An extension, being a set, is not present in each of its members. There is a third way to justify the idea that there is a real difference, and not only a pragmatic difference, between the predicates “being blue” and “being grue”. A somehow neglected ontological category, which is mentioned in Aristotle's *Categories*, are property instances or “individual accidents”, such as *the green of this coat*, said of a particular coat. The realist can say that the coat contains a property instance of green but none of grue. During the modern epoch, property instances have been conceived of in many ways and given many names. Mulligan, Simons and Smith (1984, p. 293, note 13) give the following list of words that have been used to speak of property

instances: “particularised quantities” (J. Cook Wilson), “tropes” (D.C. Williams), “individualised forms” (P.T. Geach), “concrete properties” (G. Küng), “quality-instances” (D.C. Long), “cases or aspects” (N. Wolterstorff), “instances” (R. Grossman), “accidents” (A. Kenny). For their own account, they choose the Husserlian term “moment”. Campbell (1990) who has recently defended the idea that properties fundamentally exist as particular instances rather than as universals, calls them “abstract particulars”. We take concrete individual objects and events to be bundles of property instances. We shall often omit the qualifier “natural” when speaking of properties, and use the term “property” to designate those simple qualities which entertain nomic relation with each other. These properties are the referents of the terms that appear in true law statements. Any concrete object or event possesses a large number of such properties.

- 29 For a critical analysis of Armstrong’s account of laws in the light of his “combinatorial” theory of possibility, cf. Kistler (2004).  
 30 Cf. Dretske (1977b, p. 263); Armstrong (1983, p. 39f.).

### 3 Applicability conditions and the concept of “strict law”

- 1 Cf. Scriven (1961), van Fraassen (1980; 1989), Cartwright (1983), Giere (1988). For example, Scriven says that “physical laws [i.e., in our terminology, nomological statements] are not even probably true – they are already known *not* to be true” (Scriven 1961, p. 92; italics Scriven’s).
- 2 Cf. Ramsey (1929), Wittgenstein (1964), Schlick (1931), Lange (1993). Schlick says that “natural laws, strictly analysed, are not propositions that are true or false, but are, rather, ‘directions’ for the construction of such propositions. [ . . . ] Natural laws are not [ . . . ] ‘general implications’, because they cannot be verified for *all* cases, but are rules, instructions, aiding the investigator to find his way about in reality, to discover true propositions, to expect certain events” (Schlick 1931, p. 285f.). Wittgenstein holds a similar position: “An hypothesis is a law for forming propositions. You could also say: An hypothesis is a law for forming expectations” (Wittgenstein 1964, § 228, p. 285; transl. p. 285). Von Wright (1984, p. 142ff.) tackles the problem of exceptional situations. According to him, “one can say that the standard formulation of the law *conceals* the open character of the conditions of its application, whereas the *reformulation* of the law as a universal implication just *cuts out* these circumstances by giving to the law a sharp logical structure which it does not ‘in itself’ possess” (von Wright 1984, p. 143; italics von Wright’s). This analysis leads von Wright to adopt an instrumentalist position. “One cannot disentangle the existence of laws from considerations of an epistemic nature. To say that there is a law such that . . . is *like* saying that we have a ‘device’ or ‘method’ or ‘formula’ for *predicting* that *p* at *t* under some circumstances and for *explaining* why *p* at *t* once this is an established fact” (von Wright 1984, p. 144, italics von Wright’s). However, as we shall see, the mere observation that the applicability of laws of nature is subject to conditions that are not explicitly expressed in the law statement is no good reason to adopt instrumentalism. Indeed, our own account is intended to be a form of realism compatible with this fact.
- 3 Cf. Popper (1934, chap. III, section 12); Hempel and Oppenheim (1948). Popper says: “To give a *causal explanation* of an event means to deduce a statement which describes it, using as premises of the deduction one or more *universal laws*, together with certain singular statements, the *initial conditions*” (Popper 1934, p. 31; transl. p. 59).
- 4 Cf. Canfield and Lehrer (1961); Stegmüller (1966); Hempel (1988).

- 5 Cf. Davidson (1970); Fodor (1974; 1975); Carroll (1991).
- 6 In this paper, Davidson introduces his famous “anomalous monism” as an account that allows him to explain the relation between body and mind. According to this doctrine, mental events can take part in causal relations, but only by virtue of the fact that such events, being particulars, can also be described in physical terms.
- 7 The third thesis says that there are uncontroversial cases of causal relations between mental events and physical events, in perception and action.
- 8 Davidson has later withdrawn this thesis. In (Davidson 1991) he defends the idea that causation is incompatible with strict laws. This makes his position close to that of Russell’s (1912) and Carnap’s (1928).
- 9 Davidson’s statement that every event taking part in a causal relation falls under strict *and deterministic* laws is certainly false by virtue of the statement of determinism. The radioactive decomposition of an atomic nucleus is a physical event that has effects in an ordinary sense (the omission of the particles that result from the decomposition), but which does not fall under a deterministic law – at least this is so according to presently accepted physical theories about radioactive decomposition. In what follows we shall leave to one side the distinction between deterministic and probabilistic laws, and concentrate instead on the distinction between strict and non-strict laws.
- 10 Cf. Fodor (1974; 1975; 1987; 1989; 1991).
- 11 In particular, I agree with Schiffer in rejecting the idea that a “*ceteris paribus*” clause is part of the content of some laws. As several authors have noted (cf. Giere 1988, Lange 1993), this would make the law trivial. The opposite hypothesis, defended by Armstrong (1983), will be examined later in this chapter (pages 117–20).
- 12 The modern sense of the word “enthymeme”, introduced by Boethius in the sixth century, must be distinguished from an older sense according to which it designates a formally valid deduction the truth of whose premises is not established or judged to be merely probable.
- 13 This is the example used by Canfield and Lehrer (1961).
- 14 It is important that the number of such premises is indefinite. This entails in particular that they cannot be explicitly stated. Cf. pages 117–18 in this chapter. As long as one is not conscious of this fact one may get “trapped in the idea that the ‘*ceteris paribus* clause’ is a premiss which is joined *conjunctively* with the obvious premisses.” (Lakatos 1978, p. 98; emphasis Lakatos). Lakatos detects this “confusion” in Canfield and Lehrer (1961) and Stegmüller (1966), but admits that his own way of speaking can sometimes be misleading in this respect.
- 15 The superposition of the forces that the magnet and the gravitational field exert on the iron weight results in a net force of zero, because these forces are of equal size but oriented in opposite directions.
- 16 The inference is in fact less direct. An intermediate step is required, at which one infers that the total force acting on *c* is the vector sum of the gravitational and magnetic forces.
- 17 The first group is equivalent to the conditions Hempel (1988) calls conditions of “theoretical ascent”, the second is equivalent to what he calls “provisos”. Our subclassification within the second group diverges however from Hempel’s.
- 18 It is necessary to make it clear that these are conditions of the applicability of laws as they are explicitly expressed in linguistic statements. Within a realist conception of laws, the conditions of the applicability of the type analysed in this section appear as trivially satisfied insofar as the laws themselves are concerned. The realist position is grounded on the distinction between the law and the statement of the law; this distinction allows the realist to reason directly about

- laws, abstraction made of their explicit formulation. On this point, cf. Achinstein (1971, p. 2) and Schiffer (1991).
- 19 A number  $a$  is said to be an accumulation point of a set  $E$  if every open interval containing  $a$  contains at least one number of  $E$  other than  $a$ .
  - 20 Cf. Ashcroft and Mermin (1976, p. 428).
  - 21 This is a causal process. The law  $L(H)$ , applied to a particular situation, can justify statements of the type “the fact  $Fa$  is causally responsible for the fact  $Gb$ ”. Cf. Chapter 5.
  - 22 We use the term “observational predicate” only for ease of exposition. To be precise we should systematically say “predicate belonging to a theory understood independently of the theory to which belongs the law used in the explanation under consideration”.
  - 23 Scriven expresses a thesis similar to Hempel’s (1988) when he says that inferences between propositions expressed in theoretical language ( $L_T$ ) and propositions expressed in observational language ( $L_O$ ) “are probability inferences and not deductions. It has commonly been recognized that this is true of  $L_O \rightarrow L_T$  inferences.” (Scriven 1961, p. 95). However, Scriven’s main argument aims at establishing that inferences in the opposite direction, from statements in theoretical language to statements in observational language, are not deductive but probabilistic. “It is also true of  $L_T \rightarrow L_O$  inferences, contrary to a more common assumption in the literature of philosophy of science.” (Scriven 1961, 95; immediately after the preceding quote). The following section of this chapter establishes that Scriven’s distinction is not exhaustive and that  $L_T \rightarrow L_O$  inferences are neither complete deductions nor probabilistic inferences. To say that they fall in the latter category even when no probabilistic laws are at play, is to reason incorrectly as follows: from the observation that laws are applicable only under certain conditions one concludes that their application must follow the logic of a deduction covered by a probabilistic law.
  - 24 In the present case, the systematic character of the behaviour of objects of the same type as  $o$  is not due to the systematic presence of an external source of exceptions. We shall examine the case of such external sources in the following section.
  - 25 Hempel (1988, p. 26f.) notes that the invalidity of the deductive inference from  $(L) T_1 \rightarrow T_2$  to  $(U) O_1 \rightarrow O_2$ , threatens the applicability of the method invented by Frank Ramsey, which allows us to eliminate theoretical terms from a theory. It is impossible to “complete”  $O_1$  by statements  $P$  specifying the applicability conditions of the law, in such a way that the law  $T_1 \rightarrow T_2$  would entail  $((O_1 \wedge P) \rightarrow O_2)$ , for such statements  $P$  cannot be expressed in a purely observational language. “The need for provisos precludes the general avoidability of theoretical expressions by those elimination methods” (Hempel 1988, p. 27).
  - 26 Cf. Hempel (1988, p. 23).
  - 27 Cf. Hempel (1988, p. 28).
  - 28 On this point, cf. Stegmüller (1966), Coffa (1968), Hempel (1988).

## 4 Consequences

- 1 Cf. Goodman (1955, p. 19), von Wright (1971, p. 71). Fodor (1991, p. 32) suggests this idea when he says that it might perhaps turn out to be impossible to consider non-strict laws as confirmed by their instances and grounding the evaluation of counterfactual conditionals. Fodor takes these criteria to be part of the *concept* of (strict) law.
- 2 Cf. Lewis (1973).
- 3 Cf. Davidson (1970, p. 219ff.), Fodor (1989), Schiffer (1991).

- 4 Putnam (1967) was the first to draw attention to the theoretical importance of multiple realizability in the context of mental properties.
- 5 Later, Fodor (1991) has withdrawn his thesis that this nomologically fundamental level, at which all laws are strict and exceptionless, is identical with physics as a whole:

The legitimacy of *ceteris paribus* laws is an issue not just in the philosophy of psychology but in the philosophy of the special sciences at large. And maybe not just in the special sciences. [...] Considerations recently raised by C.G. Hempel [Hempel 1988] make it seem plausible that there are *no* strict laws of nature according to this understanding.

(Fodor 1991, p. 21; emphasis Fodor's)

- 6 I do not pretend that van Fraassen's argument has no relevance for the version of the realist theory of laws defended here, but only that it does not apply to it in the particular form the argument takes in van Fraassen (1989).
- 7 Before van Fraassen, Lewis (1983, p. 366) and Carroll (1987, p. 265f.) have produced versions of this argument against theories that reduce laws to relations between universals. Van Fraassen (1993) has taken up this problem again on the occasion of retorting to Armstrong's (1993) argument that his earlier formulation of the relation between laws and their instantiations (Armstrong 1983) suffices to solve the problem of inference.
- 8 In the theory of universals (cf. Armstrong 1978; Tooley 1977; 1987), particulars are considered to be of order zero, properties of particulars and relations between particulars to be of order one. In general, a universal (property or relation) is of order  $k+1$  if  $k$  is the highest order of the entities possessing the property or standing in the relation. Cf. van Fraassen (1989, p. 100).
- 9 Criteria of the identity of natural properties will be considered in Chapter 6.
- 10 We do not need to endorse Reichenbach's verificationism according to which a statement is not objectively true or false if it is impossible to obtain conclusive knowledge of it. For Reichenbach such statements can only be more or less probable according to the degree to which they are justified by the evidence available to a subject.
- 11 "Eine Theorie heisst *falsifiziert*, wenn sie nicht nur mit einem wirklichen einzelnen Ereignis, sondern mit einem wiederholbaren Effekt in Widerspruch steht; formal gesprochen: wenn sie unverträglich ist mit einer empirischen Hypothese, die sich bewährt" (Carnap 1935, p. 290; emphasis Carnap's).
- 12 Cf. Popper (1972, p. 197ff.).
- 13 Cf. Popper (1934, 14–17, transl. 40–42).
- 14 Cf. Popper (1934, 14ff., transl. p. 40ff. and Chapter 4).
- 15 Hempel (1988, p. 25f.) and Lange (1993, p. 238f.) insist on the difference between Duhem's and Quine's arguments and the thesis developed here, according to which it is invalid to deduce a statement about the possession of a given property by a particular object from laws involving that property, because the conditions of applicability of the relevant laws are not always satisfied.
- 16 This justifies Zahar's (1982, p. 26) reply to Lakatos: He reminds him that both Duhem and Popper knew how to overcome this problem.
- 17 Later in the same text, Lakatos (1978, p. 98; note 3) regrets that he may have given the impression that *ceteris paribus* clauses have the form of premises that can be added to the other premises in the nomological deduction of an observable phenomenon.
- 18 Russell's (1912) and Carnap's (1928) theses about the inappropriateness of the concept of causation within a scientific worldview may seem to give rise to a similar scepticism. Perturbation theory shows us how to avoid it: it shows how



exact quantitative laws can be applied to situations which do not satisfy their conditions of applicability in any *absolute* sense.

- 19 Scriven (1961) says that laws which do not strictly speaking apply to any situation can nevertheless be good approximations to the phenomena in their domain. About approximations, see Coffa (1967). Pietroski and Rey's (1995) account can be understood as a philosophical interpretation of the procedure by which theoretical models are adapted to concrete situations, following the prescriptions of perturbation theory. For a critical evaluation of Pietroski and Rey's account, cf. Kistler (2003).

## 5 The nomological theory of causation and causal responsibility

- 1 As far as our *practical* needs of action planning are concerned, it may be sufficient to know nomic relations and unnecessary to know causal relations. This gives those versions of the nomological theory a certain *prima facie* plausibility, which analyse the concept of causation in terms of the concept of "efficacious strategies". Cf. Gasking (1955), Cartwright (1979). We have already critically examined these accounts above, in Chapter 1. However, the fact that the difference between a causal relation and the instantiation of a law can be *without practical relevance* does not justify considering these two relations as *objectively* identical.
- 2 Hanson's position is similar to Scriven's. He argues that causal relations exist only by virtue of our theories: "Causes certainly are connected with effects; but this is because our theories connect them, not because the world is held together by cosmic glue" (Hanson 1958, p. 64).
- 3 Some of these reasons have been exposed by Kim (1981) in an article criticizing the position defended by Scriven (1975).
- 4 Cf. Chapter 1, pages 48–9.
- 5 Cf. van Fraassen (1980, p. 132ff.); Salmon (1984, p. 95).
- 6 Humphreys (1989, p. 30f.) provides a list of types of explanations in which the explanans does not designate the cause and the explanandum not the effect. His list contains item (4) in our earlier list of counterexamples and item (2) in the following list. Humphreys does not distinguish between explanations dealing with causal relations in which the explanans does not designate the cause and the explanandum does not designate the effect (our first list) and explanations dealing with nomic but non-causal correlations (our second list).
- 7 Cf. Chapter 1, pages 48–50 and Chapter 6, pages 177–80.
- 8 Cf. Chapter 1, pages 61–2.
- 9 Cf. Chapter 1, pages 60–1. This provides a simple counterexample to a sophisticated version of the nomological theory: Clendinnen suggests reducing the concept of causal relation to the concept of "most direct nomic chain" (Clendinnen 1992, p. 352). His central idea is that it suffices to identify the lawful link between two events at the level of the most fundamental laws to be able to produce a criterion for establishing when one causal chain is more direct than another. This analysis seems to be able to predict that nomic relations between the effects of a common cause are not always related as cause and effect. However, there seem to be law-governed pseudo-processes at a fundamental nomological level which refute Clendinnen's proposal: the law of the propagation of the phase of a wave may be considered as fundamental, and not derived from more fundamental laws. The positions of the summit of a wave at two moments are lawfully related to each other by a "most direct nomic chain"; nevertheless, they are causally isolated from each other.
- 10 This remark is of course not intended to resolve a famous debate in the philosophy of action, about whether reasons can nevertheless play a causal role in

bringing about actions. Cf. Davidson (1963). Whether this is so or not is independent of our thesis that one can explain an action by making reference to a future event that is its goal, where that future event plays no causal role in bringing about the action.

- 11 Cf. Railton (1978; 1981; 1989); Salmon (1990).
- 12 On this point, cf. Pereboom and Kornblith (1991).
- 13 The fact that one of these eventive causal predicates is symmetrical (“... is causally related to ...”) whereas the other reflects the asymmetry between cause and effect (“... is the cause of ...”), should be considered in the light of the analysis offered in the first chapter of the origin of causation’s asymmetry. We have suggested that this asymmetry is a contingent property of the causal relations in this world, which is not part of the concept of causation between events (or “eventive” causation). However, the concept of causal responsibility is fundamentally asymmetrical.
- 14 Davidson says of the predicate “cause” in general, suggesting that he speaks of ordinary language, that it is “an ordinary two-place predicate in an ordinary, extensional first order language” (Davidson 1967, p. 161). Strictly speaking, what can be called “extensional” or “intensional” are not the causal relations themselves, but the statements making reference to them. Nevertheless, we shall occasionally speak of the causal relation itself as being extensional, meaning thereby that the *concept* expressed by the causal *predicate* at issue is extensional.
- 15 “Direct reference” picks out a particular object as such, whereas “reference by description” refers to an object insofar as it satisfies a description, in other words possesses the property mentioned in the description.
- 16 The old friend is supposed to be the president; and the two statements are supposed to be uttered on the same occasion. I borrow this example from Dretske (1977a).
- 17 The pair of examples (3)/(4) has two particularities irrelevant in the present context that do not stand in the way of analysing it in the same framework of event causation as the pair of statements (1)/(2): First, in (3)/(4), the cause does not bring about any change but guarantees on the contrary the *absence* of any change. Without the cause, the effect, i.e. the relative stability of the electrons with respect to the nucleus, would disappear. As we have seen in the first chapter, the absence of change can be an effect just as well as a change. Second, the events playing the role of cause and effect are situated on a temporal continuum, in other words they are temporal parts of enduring states. The details of the causal processes underlying the causal relation expressed by statements (3) and (4), i.e. the exchange of photons that realizes the electromagnetic attraction, justifies slicing these continuing states into temporarily limited events, which are causes and effects. Cf. Chapter 1, pages 43–4.
- 18 Cf. Quine (1960, §§ 31–2). Quine analyses only predicates creating opaque contexts. Our reason for introducing the concept of semi-opacity stems from the particular semantic properties of causal predicates. Contrary to the semi-opaque contexts studied here, opaque contexts do not in general allow existential quantification *salva veritate*. The predicate “desire” is a case in point: the truth of a statement in which the predicate “desire” creates an opaque context does not entail that the objects of the desire exist. From

John desires that his gain in this week’s lotto exceed one million pounds,

One cannot infer that there exists such a gain. We shall henceforth neglect the difference between opaque and semi-opaque contexts and use these expressions indifferently, given that we shall concentrate on the causal context which is strictly speaking semi-opaque.

- 19 Cf. pages 145–7.  
 20 “FACT” and “EV” represent particular expressions, whereas “*e*”, “*f*” and “*P*” represent variables bearing respectively over events, facts and predicates.  
 21 Cf. pages 149–50.  
 22 Parsons (1990) attributes to Reichenbach the idea, which he endorses, that facts can be analysed as containing information either about objects or about events, and that these analyses are equivalent. I shall criticize the thesis that they are equivalent. In particular I shall deny the equivalence of:

(i) “A flight by Amundsen over the north pole occurred” (Parsons 1990, p. 136)

and

(ii) “Amundsen flew over the north pole” (Parsons 1990, p. 136)

Indeed, one of the goals of the present chapter is to clarify the difference between the event designated by

(iii) “A flight by Amundsen over the north pole”

- and the fact expressed by (ii). The event designated by (iii) is a particular entity possessing properties not explicitly named in (iii). (i) says of a particular entity, an event, that it exists, whereas (ii) says of an object involved in this event that it possesses a certain property. On this point, cf. Davidson (1966, p. 115ff.).
- 23 Pariente (1982, p. 46ff.) explains that this “chronological function” is played by the copula. In the statement “*a* is *P*” that, if it is true, expresses the fact that *a* is *P*, the copula “is” plays both the predicative function of attributing the property *P* to *a*, and the chronological function of attributing *P* to *a* at a certain moment. If the copula is in the present tense, it is the moment of utterance. If the fact that *a* is *P* is the cause of the fact that *b* is *R*, the first and the third occurrences of “is” determine that, the instant at which *Pa* and *Rb* are facts, is the instant of utterance. The causal relation expressed by the predicate “is the cause of” holds because a causal process runs from *a* to *b* at that time.
- 24 Geach (1969, p. 321) has famously called a particular type of external relations “Cambridge changes”. The truthful attribution of such a change to an object *a* does not require any real change *within the object a*, but only in other objects standing in an external relation to *a*, for example in a relation of spatio-temporal distance or in a juridical relation. Statement (7) is a case of the former type. On the distinction between internal and external relations, cf. Mulligan and Smith (1986); Mulligan (1991).
- 25 We presuppose that a state is a continuous series of events of the same type. Cf. Chapter 1.
- 26 (8) expresses a generalization about a series of eventive causal relations. This statement sums up the net effect of a whole series of causal relations bringing about at each moment the events that together constitute the climate of Rome.
- 27 A property of an object is “intrinsic” if it is entirely determined within the space occupied by the object, and a property of an event is intrinsic if it is entirely determined within the spatio-temporal zone occupied by the event. Cf. Chapter 6.
- 28 Bennett takes this condition to define what he calls “V-facts”.
- 29 Cf. also Bennett (1988, p. 37ff.).
- 30 Cf. Horgan (1978), Bennett (1988).
- 31 On this point, cf. Mellor (1987, pp. 202, 207).

## 6 Efficacious properties and the instantiation of laws

1 Cf. Chapter 5.

2 The terms “intrinsic” and “extrinsic” are used in this sense in the debate about mental causation, more precisely about the causal role of the contents of mental states. Cf., e.g., McGinn (1982). According to an alternative terminology, used by Wilson (1995), the criterion of the determination of the property within the spatio-temporal zone occupied by the bearer of the property, yields a distinction between “internal” and “external” properties: mental properties are called “internal” or “external” according to whether they depend on something else than their bearer. For Wilson, “‘intrinsic’ and ‘extrinsic’ are labels for different *descriptions* of mental states instantiated in individuals” (Wilson 1995, p. 155; my emphasis). We adopt the former terminology because it helps us express the possibility that science makes reference to causally efficacious properties by using geometrical expressions, e.g. when the position of a particle in a field is used to designate the cause of its acceleration. Wilson’s terminology would have it that in this case, one uses an extrinsic description for an internal property. By avoiding this way of speaking, I would like to stress the fact that relational properties are as objective as properties that are entirely determined in the zone occupied by the event possessing them. In particular, a property can be relational relative to a narrow event  $e_1$  but intrinsic (“internal” in Wilson’s terminology) relative to a larger event  $e_2$  including  $e_1$ . It is important for our concerns to distinguish between two types of objective properties, not between two types of *descriptions* of properties.

Our thesis that the properties of one type are causally inert entails that explanations making reference to them are either non-causal (such as explanations making use of the concept of a planet) or indirect: for pragmatic reasons of explanatory economy, it is often useful not to mention causally efficacious processes, as is the case when the issue is the determination of the acceleration of a charged particle in an electric field.

An independent motivation for the use of the term “intrinsic” rather than “internal”, and “extrinsic” rather than “external”, is that the latter couple of terms is strongly associated with the spatial aspect of the distinction, whereas the conception of the spatio-temporal zone allows us to treat (spatially) relational properties in exactly the same way as temporally relational and historical properties. The former couple of terms is preferable because it seems to be more neutral in this respect.

3 The word “shares” suggests a realist conception of properties as universals. In an ontology in which the fundamental entities are property instances, it would be more correct to say “all samples of copper stand to each other in an internal relation of similarity, with respect to their natural properties”.

4 Cf., e.g., Tooley (1987, p. 174).

5 When the analysis of causation is not itself at issue and when the meaning of the word “cause” is taken for granted, one can use the expression “causal property” to mean “intrinsic property”. When Stalnaker, e.g., speaks of “dispositional states understood as real causal properties” (Stalnaker 1984, p. 16), he is criticizing a theory that takes those dispositional states to be extrinsic properties of their bearers. According to this theory, a person  $P$  has a mental state  $M$  only by virtue of the fact that *someone else* attributes it to her. This would make of  $P$  a purely relational property. However, according to our analysis, a property  $P$  can be causally efficacious only if it is intrinsic.

6 Blackburn considers the possibility that “science finds only dispositional properties, all the way down” (Blackburn 1990, p. 63). The contemporary discussion about the concept of disposition and its logical link to the concept of natural

- kind and counterfactual conditions begins with Carnap (1936). Cf. Mumford (1998) and Kistler and Gnassounou (forthcoming).
- 7 Cf. Posch (1981).
  - 8 According to Goodman's famous definition, "the predicate 'grue' [ . . . ] applies to all things examined before *t* just in case they are green but to other things just in case they are blue" (Goodman 1955, p. 74), where *t* is a particular moment of time.
  - 9 Cf. Lewis (1973, p. 73; 1983, p. 367; 1986, p. 124; 1994, p. 478ff.). John Earman who has also defended it, has called this account of laws of nature, by tribute to its main advocates, the "Mill-Ramsey-Lewis [ . . . ] account of laws" (Earman 1984, p. 197; 1993, p. 416).
  - 10 The qualification "contingent" is intended to exclude from laws of nature generalizations that are true for logical or other purely formal reasons.
  - 11 Cf. van Fraassen (1989, p. 53).
  - 12 Van Fraassen makes a similar objection against Lewis's account of laws: "Truth and simplicity just do not add up to necessity, as far as my intuitive reactions are concerned" (van Fraassen 1989, p. 47).
  - 13 Cf. van Fraassen (1989, p. 57/8).
  - 14 Cf. Quine (1953; 1960, p. 200ff.) and Sober (1981).
  - 15 Achinstein notes that "if principles of coextensiveness and necessity were sufficient for property identity then we should have to accept as identical all properties related by laws asserting or implying that something has the one iff [i.e. if and only if] it has the other" (Achinstein 1974, p. 266).
  - 16 This is of course not the only possibility. According to one conception of events, developed by Goldman (1973) and Kim (1973), an event is characterized by a constitutive property that may be considered as essential to it. Lewis (1986) and Yablo (1992a; 1992b) also take certain properties of events to be essential in an absolute sense, distinguishing in particular between essentially "mental" and essentially "cerebral" events.
  - 17 Davidson seems to take the concept of causal law to be equivalent to that of strict law. Cf. Davidson (1967; 1969, p. 165). In Davidson's nominalist framework, a "law" is what we call a "law statement". However, the existence of strict laws (and thus, for Davidson, causal laws) can be justified within nominalism even less than it can in a realist conception: to be strict such laws would have to contain an explicit mention of all their conditions of applicability. Cf. Chapter 3.
  - 18 These notions are often used, though rarely explicated, in philosophy of mind (cf., e.g., LePore and Loewer 1987, p. 630) and philosophy of science (Cf., e.g., Tooley 1987, pp. 175, 246–247, 274).
  - 19 Many authors, such as Kim (1974), Salmon (1978, p. 687) and Cummins (1983), use the distinction between causal and non-causal laws without justifying it. Others, such as Armstrong (1993, p. 422), maintain, in the spirit of the nomological theory of causation, that *all* laws are causal. For Tooley (1987, p. 274), all these distinctions are derived from a fundamental distinction between two kinds of nomic relation between universals, in other words a fundamental distinction at the level of laws. Cf. below, pages 180–5.
  - 20 These concepts have been introduced in the first chapter, pages 57–67.
  - 21 This has been suggested by Mackie (1974, p. 210ff.). Mackie calls laws of evolution that ground causal relations, "laws of working" and distinguishes them from "functional laws" expressing "concomitant variation". Mackie does not simply attribute a causal character to the first class of laws and deny it to the second class. Rather, he takes there to be a kind of "functional causation" (Mackie 1974, p. 147) which depends on the instantiation of a law of association. However, it is for him an indirect type of causation that is immediately dependent on a law of

- evolution or “neolithic law”. “A functional law will entail the holding of neolithic laws in appropriate fields” (Mackie 1974, p. 147).
- 22 Cf. Tooley (1987, p. 272). It is not easy to interpret the dependency of relations  $C$  on relations  $\rightarrow$  in Tooley’s theory. Crucial for our argument is that Tooley gives ontological priority to relation  $\rightarrow$ , insofar as he takes the other option to entail a singularist theory of causation that he rejects. (Tooley changes his mind on this point in 1990. See below.) “Any relation,  $C$ , which is the special, first-order relation that obtains between any two causally related states of affairs, has that status *not by virtue of its intrinsic nature*, but by virtue of the fact that it is the relation which enters into all causal laws in a certain way – and where causal laws are laws which obtain by virtue of a second-order relation between universals which determines the direction of the logical transmission of probabilities. The primary locus of causal fact [ . . . ] is the second-order relation which enters into causal laws” (Tooley 1987, p. 274; emphasis Tooley’s). Cf. Van Fraassen (1989, chap. 5) for criticism of Tooley’s account of the relation between  $\rightarrow$  and  $C$ . Later, Tooley has changed his mind and set out to defend the opposite thesis according to which “it is possible for one event to cause another, without the causal relation being an instance of any causal law” (Tooley 1990b, p. 272) and according to which “events can be causally related even in worlds where there are no causal laws” (Tooley 1990b, p. 322). However, Tooley keeps thinking that it is “very unlikely that a reductionist analysis can be given for causation if causal relations do not presuppose causal laws” (Tooley 1990b, p. 321). Our own reduction of causation is not explicitly incompatible with this thesis but it only entails the dependency of causal relations on conservation laws and there is no reason to consider conservation laws as “causal laws”.
- 23 Most “probabilistic” theories of causation (Cartwright 1979, Skyrms, 1980; Eells and Sober 1983; Cartwright 1989; Eells 1991) start from the idea that causation is a relation between two *types* of events or between two properties  $F$  and  $G$ , where  $F$  causes  $G$  if and only if (i)  $F$  does not lower the probability of  $G$  in any homogeneous situation allowing to test whether  $F$  causes  $G$  and (ii)  $F$  enhances the probability of  $G$  in *at least one* causally homogeneous situation allowing to test whether  $F$  causes  $G$ . Skyrms (1980, p. 108) calls this the condition of “pareto-dominance”. A large number of counterexamples to such theories have been found. Cf. Otte (1981; 1985), Sober (1982), Dupré (1984), Carroll (1991). They seem to show that the increase in probability is an indicator of causation which is often but not always reliable. The only version of that theory I shall consider a little more closely is Tooley’s; it goes without saying that I do not pretend to do justice to the others.
- 24 Salmon (1984) notes clearly that statistical relations are only indicators for the existence of causal relations but can provide no basis for their ontological reduction.
- 25 “The probability of the coin’s coming up heads gets ‘transmitted’, so to speak, to the die’s coming up six” (Tooley 1987, p. 252).
- 26 Fales (1990, p. 609) makes a similar objection. He asks what excludes the possibility that the particular place imagined by Tooley be the *common cause* that brings about both the deviant behaviour of the die and of the coin. The problem is that the way Tooley sets up his thought experiment is not detailed enough to decide between these possibilities. Fales notes also that Tooley’s account is not compatible with the existence of causal relations in a situation where no deviation of frequencies is observed: in one such situation, two dice fall on “six” exactly once in six trials, but do so exactly in the same trials. Fales’ objections confirm our criticism that Tooley mistakes the deviations of frequency that causal relations normally bring about, for these causal relations themselves.
- 27 We agree with Papineau in this respect. He argues against the attempt to *define* causation in terms of relations of increase in probability, pointing out that

“probabilistic relationships [ . . . ] rather [ . . . ] provide indirect evidence for the existence of ‘underlying’ causal connections” (Papineau 1985, p. 276). Coffa expresses an analogous idea concerning the relation between judgments of the existence of nomic relations and observations of probabilities: “When nomicity and statistical relevance enter into conflict, it seems clear that nomicity always wins the day. They seldom enter into conflict, for statistical relevance is the evidence that we may have for the presence of nomicity” (Coffa 1974, p. 162).

- 28 In two other clauses Tooley stipulates that the link between  $P$  and  $Q$  must not be logically necessary, and must not be logically equivalent to facts concerning particulars.
- 29 Tooley really develops this argument with the help of a set of postulates ( $R_1$ ) – ( $R_6$ ). These new postulates differ from ( $M_1$ ) – ( $M_6$ ) only concerning the relation of causal necessitation, designated by “ $\rightarrow_i$ ” in postulates ( $R_1$ ) – ( $R_6$ ). This difference has no incidence on the present discussion.
- 30 Carroll’s (1991) criticism of accounts of causal laws working with the concept of “probabilification” leads him to a similar conclusion. According to Carroll, “property-level causation should not be a central topic in the philosophy of science” (Carroll 1991, p. 245), for the “interest of property-level causation [ . . . ] is [ . . . ] exhausted by the interest of singular causation” and “by the interest of lawhood” (Carroll 1991, p. 267). His examples and his analysis make it clear that his objections against “property-level causation” are in fact objections against the concept of causal laws. However, Carroll’s own proposal is not satisfying. First, he says nothing about singular causation; second, his analysis reduces the relation of nomic dependence to a frequency. For example, Carroll analyses the causal generalization “dogs bark” by “the frequency of barkers in dogs is high”, in symbols, “ $Fr(Gx/Dx)$  is high” (Carroll 1991, p. 257). He introduces the notion of “high frequency” in order to take exceptions into account: there are dogs that do not bark. However, it is inappropriate to analyse laws with exceptions in terms of events correlated with a high frequency. Indeed, the *number* of exceptions for a given law is irrelevant for their status as exceptions. To take an example used by Carroll (1991, p. 258–259) himself: if by virtue of a law of nature, sunspots cause electrical perturbations on Earth, but if four times out of five, another law prevents the effect from occurring, the frequency of the effect – the electrical perturbation – given the cause – a sunspot – is low. But this does not prevent the sunspots from being, in the remaining cases, the causes of the electrical perturbations. However, Carroll’s analysis, reducing nomic dependency to high frequency, is bound to predict the opposite.

## 7 Causal responsibility and its applications

- 1 After a brief analysis of this phenomenon in terms of facts, Bennett concludes by saying that “highlighting can henceforth be neglected, as having little philosophical significance” (Bennett 1988, p. 35).
- 2 The intuition that a change in emphasis between (1) and (2) brings with it a change in truth value is weaker than the analogous intuition with respect to the substitution of a predicate or a descriptive expression, as in examples (1)/(2) and (3)/(4) of Chapter 5. One of the aims of the present chapter is to show that the situation is nevertheless fundamentally the same. Taking into account the complete sense of the causal concept expressed by the predicate, one takes into account not only the cause event but also a specific causally efficacious property. In (1) and (2), the emphasis contains the information about which property of the cause has been efficacious in bringing about the effect. At the same time, according to our account, the concept of causal responsibility ( $C_F$ ) contains an

extensional element ( $C_E$ ), whose existence explains the intuition that there is “something true” in a statement such as (2).

- 3 We shall use the expressions “aspect”, “facet” and “property” interchangeably with respect to events; this is justified insofar as those terms seem equally interchangeable when applied to objects. When we intend to draw explicit attention to the difference between objects and events, we shall use the term “property” for objects and the terms “aspect” and “facet” for events.
- 4 (2) contains an ambiguity that is inessential for our present purposes. Instead of structure (2’), one can interpret (2) as having the logical structure (2’’):

(2’’)  $C_F(\text{drinks hemlock at dawn (Socrates at } t), \text{ dies (Socrates at } t + \Delta t))$ .

The ambiguity concerns the exact interpretation of the nomic link. (2’’) entails that there is a law linking the ingestion of hemlock at dawn to the death of the drinker, whereas (2’) entails the existence of a nomic link between the fact of drinking *anything* at dawn and the death of the drinker.

- 5 Cf. Chapter 5, page 152, Thesis I.
- 6 Cf. our Thesis III introduced below.
- 7 This distinction is due to Dretske (1972) although he uses a different terminology. Achinstein (1979) uses it when he justifies his rejection of Thesis II (cf. Chapter 5, page 152), according to which the causal context is transparent.
- 8 The non-semantic use of “contrastive statements”, as Dretske (1972) calls them, or of “emphasis” as Achinstein (1975; 1979) and Kim (1977) call it, play a pragmatic role for the communication of information. Emphasis can be used for denial, for highlighting new information, or for drawing attention to aspects of particular interest, importance, significance, surprise etc., for the expected auditor of the statement. Cf. Dretske (1972) and Achinstein (1975).
- 9 This is Achinstein’s (1975) terminology.
- 10 It is necessary to imagine a context of utterance as unusual as that, in order to avoid that the context plays the role of explicit emphasis: the role of disambiguating the statement with respect to the property of the cause that is said to be efficacious. In most situations, general knowledge shared by all participants in the communication will make it clear that hemlock but not dawn is capable of playing the role of the property causally efficacious for bringing about the death of Socrates.
- 11 Without emphasis, such as in (1a), it would not be possible to determine the first factive argument of the causal verb  $C_F(Fc, Ge)$ . Therefore, statement (1a) does not express a complete proposition and cannot be semantically evaluated, at least if we suppose that the context of utterance does not provide this information either. However, what we have called the extensional structural element of the content of (1a), which corresponds to  $C_E(c, e)$ , is perfectly determinate, insofar as (1a) specifies the eventive arguments  $c$  and  $e$ . (1a) can therefore be partially evaluated by taking into account only the extensional part of its content  $C_E(c, e)$ .
- 12 (2-O) follows interpretation (2’) of (2). In the analysis corresponding to interpretation (2’’) (cf. above, note 4) the last clause would read: “... the causal process by which  $c$  brings about  $e$  depends on  $c$ ’s property of being a drinking of hemlock by Socrates occurring at dawn and affects  $e$ ’s property of being Socrates’ death.”
- 13 Cf. Føllesdal (1965, p. 271).
- 14 Levin (1976) shares our judgment that there is a sense of (7), or a legitimate interpretation of it, in which it is true in Føllesdal’s situation. However, his analysis of the logical structure of causal statements, and in particular of Føllesdal’s example, is quite different from ours. According to Levin, all argument places in a causal statements are transparent (or “extensional”, as Levin says). He for-



- malizes and develops Davidson's (1967) account of causal statements (cf. Chapter 5, pages 145–7 above), with the result that the logical structure of causal statements in the proper sense is  $C(c,e)$ , where  $c$  and  $e$  represent events. In his analysis, statements making reference to the efficacious property of the cause are *causal explanations*. Their logical form is  $C^*(c,e,D)$ , where  $D$  represents a particular *description* of the cause. This allows Levin to say that the argument place  $D$  is transparent. He combines this analysis with the nomological theory of causal explanations, according to which  $D$  is the antecedent predicate of a law statement linking it to a description of the effect.
- 15 The analysis of the problem cases discussed in this section is further developed in Kistler (1999).
  - 16 In our terminology, what Aronson calls a “result” is a factive effect, what he calls an “occasion” is a factive cause. However, one can only conclude that a result is not an effect and that an occasion is not a cause, if one supposes that only events can be causes. This is precisely what I deny.
  - 17 Cf. Chapter 5, page 154.
  - 18 The indirect way consists in studying *other* causal processes, to which the process at issue is related within a causal network. Cf. Chapter 1.
  - 19 Cf. Chapter 6, pages 166–76.
  - 20 In a similar way, Horgan (1978) contests the ontological relevance of the analysis of the structure of ordinary language: “The surface grammar of ordinary discourse does not in itself commit us to the existence of entities such as avalanches and catastrophies” (Horgan 1978, p. 39). On some occasions, it is indeed justified to reject what ordinary language statements suggest about the nature of the entities they designate. It turns out, e.g., that the verb “cause” does not always express the concept of causation reducible to physical transmission or the concept of causal responsibility. In some circumstances, this word can be used to refer to an abstract relation of moral responsibility, between an agent and a consequence of his action (or omission). However, one needs a *justification* for rejecting the ontological implications of true statements. Mellor and Horgan seem to refuse in general the capacity of the analysis of linguistic structures to provide information about the entities they refer to. They owe us an alternative account of how to do ontology.
  - 21 Negative facts pose a problem for modern realist analyses of causation; but they did not pose any problem for the Aristotelian conception of causation in which our distinction between facts and events would make no sense. As the example in the quote from Aristotle's *Physics* shows, negative causes raise no conceptual difficulty for his account. To understand why, it may suffice to remember that Aristotle's concept of causation is not yet distinct from the concept of explanation. Indeed, Aristotle's famous four types of causes are often interpreted as four types of explanation. Cf. Wieland (1970, p. 262).
  - 22 Cf. Chapter 1, pages 26–9.
  - 23 This point has famously been made by Davidson (1967).
  - 24 A similar point has been made by Quine: “One's interest in contributory causes is conspicuously independent of the proportion of energy contributed. [ . . . ] In practice, indeed, the words ‘cause’ and ‘because’ often connote no contribution of energy at all. They are stretched to apply to logical premises, to purposes, to dispositions.” (Quine 1973, p. 7).
  - 25 Cf. pages 197 on Aronson's distinction between “effect” and “result”, which allows him to say that the cooling of the water is the result, though not the effect, of throwing an ice cube in the water.
  - 26 Both Ehring (1986, p. 252) and Krajewski (1997, pp. 195–196) argue that accounts reducing causation to transference are incapable of handling cases of triggering or release adequately.

- 27 Krajewski (1997, p. 195) describes a case of an electrically triggered explosion. This may be a case in which the effect of the triggering is part of the cause of the explosion in the sense of the transference theory, although, as Krajewski makes clear, a part whose quantitative contribution to the cause of the explosion is negligibly small.
- 28 This point is further illustrated in the next section.

# References

- Achinstein P. (1971), *Law and Explanation*, Oxford: Clarendon Press.
- (1974), The Identity of Properties, *American Phil. Quarterly* 11(4), pp. 257–275.
- (1975), Causation, Transparency, and Emphasis, *Canadian J. of Phil.* 1, pp.1–23.
- (1979), The Causal Relation, *Midwest Studies in Philosophy* IV, pp. 369–386.
- Adams E.W. (1988), Confirming Inexact Generalizations, *PSA (Philosophy of Science Association) 1988* vol. 1, pp. 10–16.
- Aharonov Y. and Bohm D. (1959), Significance of Electromagnetic Potentials in Quantum Theory, *Physical Review* 115, p. 485.
- Albert D.Z. (1994), Bohm's Alternative to Quantum Mechanics, *Scientific American* 5: pp. 32–39.
- Anderson J.L. (1967), *Principles of Relativity Physics*, New York: Academic Press.
- Anscombe, G.E.M. (1971), Causality and Determination, in: Sosa and Tooley (1993), pp. 88–104.
- Aristotle (1984), Posterior Analytics, translated by J. Barnes, in: *The Complete Works of Aristotle, The Revised Oxford Translation*, J. Barnes ed., Vol. I, Bollingen Series 71.2, Princeton University Press, Princeton, NJ, pp. 114–166.
- (1970), *Physics, Books I and II*, transl. by W. Charlton, Oxford: Clarendon Press.
- Armstrong D.M. (1978), *Universals and Scientific Realism*, Cambridge: Cambridge University Press.
- (1980), Identity through Time, in: P. van Inwagen (ed.), *Time and Cause*, Dordrecht: Reidel, pp.67–78.
- (1983), *What is a Law of Nature?*, Cambridge: Cambridge University Press.
- (1992), Properties, in: Kevin Mulligan (ed.), *Language, Truth, and Ontology*, Dordrecht: Kluwer.
- (1993), The Identification Problem and the Inference Problem, *Phil. and Phen. Research* 53, pp. 421–422.
- Aronson J.J. (1971a), On the Grammar of 'Cause', *Synthese* 22, pp. 414–430.
- (1971b), The Legacy of Hume's Analysis of Causation, *Stud. Hist. Phil. Sci.* 2, pp. 135–165.
- (1982), Unentangling Ontology from Epistemology in Causation, *Erkenntnis* 18, pp. 293–305.
- Ashcroft N.W. and Mermin N.D. (1976), *Solid State Physics*, Fort Worth: Saunders.
- Beauchamp T. and Rosenberg A. (1981), *Hume and the Problem of Causation*, New York: Oxford University Press.
- Bell J.S. (1964), On the Einstein–Podolsky–Rosen paradox (1964), in: Bell 1987.

- (1976), Einstein–Podolsky–Rosen experiments (1976), in: Bell 1987.
- (1987), *Speakable and Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press.
- Bennett J. (1988), *Events and Their Names*, Indianapolis/Cambridge: Hackett.
- Bigelow J., Ellis B. and Pargetter R. (1988), Forces, *Phil. of Science* 55, pp. 614–630.
- Bigelow J. and Pargetter R. (1990), Metaphysics of Causation, *Erkenntnis* 33, pp. 89–119.
- Blackburn, S. (1990a), Hume and Thick Connexions, *Philosophy and Phenomenological Research*, 50, Supplement, pp. 237–250; repr. in: Blackburn (1993).
- (1990b), Filling in Space, *Analysis* 50, pp. 62–65; repr. in: Blackburn (1993).
- (1993), *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Bohm D. (1951), *Quantum Theory*, Englewood Cliffs, NJ: Prentice Hall.
- (1962), Classical and Non-classical Concepts in the Quantum Theory, *British Journal for the Philosophy of Science*, 12 (48), pp. 265–280.
- Braithwaite R.B. (1959), *Scientific Explanation*, Cambridge: Cambridge University Press.
- Bunge M. (1959), *Causality and Modern Science*, 3rd edn. Dover: New York, 1979.
- Campbell K. (1990), *Abstract Particulars*, Oxford: Blackwell.
- Canfield J. and Lehrer K. (1961), A Note on Prediction and Deduction, *Phil. of Science* 28, pp. 204–208.
- Carnap R. (1928), *Der logische Aufbau der Welt*, Meiner, Hamburg 1998; Engl. translation by Rolf A. George (1967) *The Logical Structure of the World; Pseudo problems in Philosophy*, London: Routledge & Kegan Paul.
- (1931), Die physikalische Sprache als Universalsprache der Wissenschaft, *Erkenntnis* 2: pp. 432–465.
- (1934), *Logische Syntax der Sprache*, Wien, Julius Springer, 1934, 2nd edn, 1968; Engl. translation by Amethe Smeaton (Countess von Zeppelin): *The Logical Syntax of Language* (1937), London: Routledge & Kegan Paul, 6th impression 1964.
- (1935), Review of: Karl Popper, Logik der Forschung, *Erkenntnis* 5, pp. 290–294.
- (1936), Testability and Meaning, *Phil. of Science* 3, pp. 420–471.
- (1966), *Philosophical Foundations of Physics. An Introduction to the Philosophy of Science*, Martin Gardner (ed.), New York, Basic Books.
- Carroll J.W. (1987), Ontology and the Laws of Nature, *Australasian Journal of Philosophy* 65(3), pp. 261–276.
- (1991), Property-Level Causation?, *Philosophical Studies* 63, pp. 245–270.
- Cartwright N. (1979), Causal Laws and Effective Strategies, in: Cartwright 1983.
- (1983), *How the Laws of Physics Lie*, Oxford: Clarendon Press.
- (1989), *Nature's Capacities and their Measurement*, Oxford: Clarendon Press.
- Castaneda H.-N. (1984), Causes, Causity, and Energy, *Midwest Studies in Philosophy* IX, pp.17–28.
- Chang H. and Cartwright N. (1993), Causality and Realism in the EPR Experiment, *Erkenntnis* 38, pp. 169–190.
- Chisholm R. (1955), Law Statements and Counterfactual Inference, *Analysis* 15, pp. 97–105.
- Clendinnen F.J. (1992), Nomic Dependence and Causation, *Phil. of Science* 59, pp. 341–360.
- Coffa, J. A. (1967), Feyerabend on Explanation and Reduction, *J. of Phil.* 16, pp. 500–508.

- (1968), Discussion: Deductive Predictions, *Phil. of Science* 35, pp. 279–283.
- (1974), Hempel's Ambiguity, *Synthese* 28, pp. 141–163.
- Couty R. and Ezra J. (1965), *Analyse*, Paris: A. Colin.
- Cummins R. (1983), *The Nature of Psychological Explanation*, Cambridge, MA: MIT Press.
- Davidson D. (1963), Actions, Reasons, and Causes, in: Davidson 1980.
- (1966), The Logical Form of Action Sentences, in: Davidson 1980.
- (1967), Causal Relations, in: Davidson 1980.
- (1969), The Individuation of Events, in: Davidson 1980.
- (1970), Mental Events, in: Davidson 1980.
- (1980), *Essays on Actions and Events*, Oxford, Clarendon Press 1980.
- (1991), Three Varieties of Knowledge, in: Griffiths (ed.), *A.J. Ayer: Memorial Essays*, pp. 153–166.
- Dieks D. (1981), A Note on Causation and the Flow of Energy, *Erkenntnis* 16, pp. 103–108.
- (1986), Physics and the Direction of Causation, *Erkenntnis* 25, pp. 85–110.
- Dowe P. (1992a), Process Causality and Asymmetry, *Erkenntnis* 37, pp. 179–196.
- (1992b), Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory, *Phil. of Science* 59, pp. 195–216.
- (1992c): An empiricist defence of the causal account of explanation, *International Studies in the Philosophy of Science* 6, pp. 123–128.
- (1995a): Causality and Conserved Quantities: A Reply to Salmon, *Phil. of Science* 62, pp. 321–333.
- (1995b): What's Right and What's Wrong With Transference Theories, *Erkenntnis* 42, pp. 363–374.
- (1996): Backwards Causation and the Direction of Causal Processes. *Mind* 105, pp. 1–22.
- Dretske F. (1972), Contrastive Statements, *Phil. Rev.* 81: 411–437.
- (1977a): Referring to Events, *Midwest Studies in Philosophy* II, pp. 90–99.
- (1977b), Laws of Nature, *Phil. of Science* 44: 248–268.
- Ducasse C.J. (1924), Causation and Types of Necessity, *University of Washington Publications in the Social Sciences*, 1, pp. 69–200.
- (1926), On the Nature and Observability of the Causal Relation, in: Sosa and Tooley (1993), pp. 125–136.
- Duhem P. (1906), *La théorie physique: son objet et sa structure*, Paris: Vrin, 1981.
- Dupré J. (1984), Probabilistic Causality Emancipated, *Midwest Studies in Philosophy*, Vol. IX, pp. 169–176.
- Earman J. (1976), Causation: A Matter of Life and Death, *J. of Phil.* 73, pp. 5–25.
- (1984), Laws of Nature: The Empiricist Challenge, in R. J. Bogdan (ed.), *D.M. Armstrong*, Dordrecht: Reidel, pp. 191–223.
- (1993), In Defence of Laws: Reflections on Bas van Fraassen's *Laws and Symmetries*, *Phil. and Phen. Research* 53, pp. 413–419.
- Eells E. (1991), *Probabilistic Causality*, Cambridge: Cambridge University Press.
- Eells E. and Sober E. (1983), Probabilistic Causality and the Question of Transitivity, *Phil. of Science* 50, pp. 35–57.
- Ehring D. (1986), The Transference Theory of Causation, *Erkenntnis* 14, pp. 249–258.
- (1997), *Causation and Persistence: A Theory of Causation*, New York: Oxford University Press.

- Einstein A. (1907), Über die vom Relativitätsprinzip geforderte Trägheit der Energie, *Annalen der Physik* 23(7), pp. 371–384.
- Einstein A., Podolsky B. and Rosen N. (1935), Can the Quantum-Mechanical Description of Reality be Considered Complete? *Physical Review* 47(10), pp. 777–780.
- Einstein A. and Infeld L. (1936), *L'évolution des idées en physique*, transl. M. Solovine, Paris: Flammarion 1983.
- Fair D. (1979), Causation and the Flow of Energy, *Erkenntnis* 14, pp. 219–250.
- Fales E. (1990), Review of: Michael Tooley, Causation. *Phil. and Phenomenol. Res.* 50, Suppl.
- Feynman R. (1965), *The Character of Physical Law*, Cambridge, MA: MIT Press.
- Fodor J.A. (1974), Special Sciences, repr. as Chapter 5 in: *Representations*, pp. 127–145, Cambridge, MA: MIT Press, 1981.
- (1975), Introduction, in: *The Language of Thought*, Cambridge, MA: Harvard University Press 1979.
- (1987), *Psychosemantics*, Cambridge, MA: MIT Press.
- (1989), Making Mind Matter More, *Philosophical Topics* 17(1), pp. 59–79.
- (1991), You Can Fool Some of The People All of The Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations, *Mind* 100, pp. 19–34.
- Føllesdal, Dagfinn (1965), Quantification into Causal Contexts, in: Robert S. Cohen and Marx W. Wartofsky (eds), *Boston Studies in the Philosophy of Science, vol. 2*, New York: Humanities Press, 263–274.
- Frege G. (1884), *Die Grundlagen der Arithmetik*, repr. and trad. angl. par J.L. Austin, *The Foundations of Arithmetic*, Oxford: Basil Blackwell, 1959.
- (1891), Funktion und Begriff; repr. in Frege (1962); transl. by Peter Geach, Function and Concept, in Frege (1980), pp. 21–41.
- (1892a), Über Begriff und Gegenstand, repr. in: Frege (1962); transl. by Peter Geach, On Concept and Object, in Frege (1980), pp. 42–55.
- (1892b), Über Sinn und Bedeutung, *Ztschr.f.Philos.u.philos. Kritik* NF 100, pp. 25–50, repr. in: Frege (1962); transl. by Max Black, On Sense and Meaning (in the 1st edn, Basil Blackwell, Oxford, 1952, the title is “On Sense and Reference”), in: Frege (1980), pp. 56–78.
- (1904), Was ist eine Funktion?, repr. in: Frege (1962), *Funktion, Begriff, Bedeutung*, ed. Günther Patzig, Göttingen: Vandenhoeck & Ruprecht.
- (1980), *Translations from the Philosophical Writings of Gottlob Frege*, ed. by Peter Geach and Max Black, 3rd edn (1st edn 1952), Oxford, Basil Blackwell, 1980, repr. 1988.
- (1983), Ausführungen über Sinn und Bedeutung, in: *Nachgelassene Schriften, vol. 1*, 128–136; eds. Hans Hermes, Friedrich Kambartel and Friedrich Kaulbach, Hamburg: Felix Meiner; translation: Comments on Sense and Meaning, in: Gottlob Frege, *Posthumous Writings*, ed. by Hans Hermes, Friedrich Kambartel and Friedrich Kaulbach, transl. by Peter Long and Roger White, Basil Blackwell, Oxford, 1979.
- Fuhrmann, A. (1991), Tropes and Laws, *Phil. Stud.* 63, pp. 57–82.
- Gasking D. (1955), Causation and Recipes, *Mind* 64, pp. 479–487.
- Geach P. (1969), God's Relation to the World, repr. in: *Logic Matters*, Berkeley, CA: University of California Press, 1972, pp. 318–327.
- Giere R.N. (1988), Laws, Theories and Generalizations, in: Grünbaum and Salmon 1988.

- Goldman A. (1970), *A Theory of Human Action*, Englewood Cliffs, NJ: Prentice-Hall.
- Goodman N. (1955), *Fact, Fiction, and Forecast*, Cambridge, MA.: Harvard University Press.
- Grünbaum A. (1973), *Philosophical Problems of Space and Time*, Dordrecht: Reidel.
- Grünbaum A. and Janis A.I. (1977), Is there Backward Causation in Classical Electrodynamics? *Journal of Phil.* 74, pp. 475–482.
- Grünbaum A. and Salmon W. (eds) (1988), *The Limitations of Deductivism*, Los Angeles, CA: University of California Press.
- Hacking I. (1983), *Representing and Intervening*, Cambridge University Press.
- Hanson N.R. (1958), *Patterns of Discovery*, Cambridge: Cambridge University Press.
- Hausman D. (1986), Causation and Experimentation, *American Phil. Quarterly* 23(2), pp. 143–154
- Heathcote A. and Armstrong D. (1991), Causes and Laws, *Noûs* 25, pp. 63–73.
- Hempel C.G. (1965), *Aspects of Scientific Explanation*, New York: Free Press.
- (1966), *Philosophy of Natural Science*, Englewood Cliffs, NJ: Prentice Hall.
- (1988), Provisos: A Problem Concerning the Inferential Function of Scientific Theories, in: Grünbaum and Salmon 1988.
- Hempel C.G. and Oppenheim P. (1948), Studies in the Logic of Explanation, in: Hempel 1965.
- Holt P.J. (1976), Causality and Our Conception of Matter, *Analysis* 37, pp. 20–29.
- Horgan, T. (1978), The Case Against Events, *Phil. Rev.* 87(1), pp. 28–47.
- Horwich P. (1987), *Asymmetries in Time*, Cambridge, MA: MIT Press.
- Hume D. (1955), *A Treatise of Human Nature* (1739), L.A. Selby-Bigge and P.H. Nidditch (eds), Oxford, Clarendon Press.
- Humphreys P.W. (1989), Scientific Explanation: The Causes, Some of the Causes, and Nothing But the Causes, in: Kitcher and Salmon 1989, pp. 283–306.
- Jackson J.D. (1975), *Classical Electrodynamics*, New York: Wiley.
- Kerner E.H. (ed.) (1972), *The Theory of Action-at-a-Distance in Relativistic Particle Dynamics*, New York: Gordon and Breach.
- Kim J. (1969), Events and Their Descriptions: Some Considerations, in: Rescher *et al.*
- (1973), Causation, Nomic Subsumption and the Concept of Event, *J. of Phil.* 70(8), pp. 217–236
- (1974), Noncausal Connections, *Noûs* 8, pp. 41–52.
- (1976), Events as Property Exemplifications, in: M.Brand and D.Walton (eds), *Action Theory*, Dordrecht: Reidel, pp. 159–177.
- (1977), Causation, Emphasis, And Events, *Midwest Studies in Philosophy II*, pp. 100–103.
- (1981), Causes as Explanations: A Critique, *Theory and Decision* 13, pp. 293–309.
- Kistler, M. (1999), Causes as Events and Facts, *Dialectica* 53, pp. 25–46.
- (2002), The Causal Criterion of Reality and the Necessity of Laws of Nature, *Metaphysica* 3(1), pp. 57–86.
- (2002a), Causation in contemporary analytical philosophy, *Quaestio-Annuario di storia della metafisica*, vol. 2, ed. C. Esposito and P. Porro, Turnhout (Belgium), Brepols, pp. 635–668.
- (2003), “Laws of Nature, Exceptions and Tropes”, *Philosophia scientiae* 7(2), pp. 189–219.
- (2004), “Le combinatorialisme et le réalisme nomologique sont-ils compatibles ?”, in: J.-M. Monnoyer (ed.), *La structure du monde: objets, propriétés, états de choses*, Paris, Vrin, pp. 199–221.

- Kistler, M. and Gnessounou, B. (forthcoming), *Dispositions and Causal Powers*, Aldershot: Ashgate.
- Kitcher P. (1989), Explanatory Unification and the Causal Structure of the World, in: Kitcher and Salmon 1989, pp. 410–505.
- Kitcher P. and Salmon W.C. (eds) (1989), *Minnesota Studies in the Philosophy of Science, Vol. XIII: Scientific Explanation*, Minneapolis, MN: University of Minnesota Press.
- Kneale W. (1949), *Probability and Induction*, Oxford: Oxford University Press.
- (1950), Natural Laws and Contrary-to-Fact Conditionals, *Analysis* 10, pp. 121–125.
- (1961), Universality and Necessity, *Brit. J. Phil. Science* 12, pp. 89–102.
- Krajewski W. (1982), Four Conceptions of Causation, in: W. Krajewski ed., *Polish Essays in the Philosophy of the Natural Sciences*, Reidel: Dordrecht, 1982.
- (1997): Energetic, informational, and triggering causes. *Erkenntnis* 46, pp. 193–202.
- Lakatos, I. (1978), Falsification and the Methodology of Scientific Research Programmes, in: *The Methodology of Scientific Research Programmes*, third edn, Cambridge: Cambridge University Press, 1983.
- Lange M. (1993), Natural Laws and the Problem of Provisos, *Erkenntnis* 38, pp. 233–248.
- Largeault J. (1985), *Systèmes de la nature*, Paris: Vrin.
- Lee T.D., Oehme R., and Yang C.N. (1957), Remarks on Possible Noninvariance under Time Reversal and Charge Conjugation, *Phys. Rev.* 106, pp. 340–345.
- Leibniz G.W. (1890), II. Specimen of universal calculus, in *Die philosophischen Schriften*, ed. C.J. Gerhardt, vol. 7, Weidmannsche Buchhandlung, Berlin, p. 218–227; transl. by Leroy E. Loemker, in: Gottfried Wilhelm Leibniz, *Philosophical Papers and Letters*. A Selection Translated and Edited, with an Introduction by Leroy E. Loemker, 2nd edn, Reidel, Dordrecht, 1969, pp. 240–247 (1st edn, 1956, Chicago, IL: Chicago University Press).
- Lemmon E.J. (1967), Comments on D. Davidson's "The Logical Form of Action Sentences", in: *The Logic of Decision and Action*, Nicholas Rescher (ed.), Pittsburgh, PA: University of Pittsburgh Press.
- LePore E. and McLaughlin B. (eds.) (1985), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell.
- Le Pore E. and Loewer B. (1987), Mind Matters, *J. of Phil.* 84(11), pp. 630–642.
- Levin M.E. (1976), The Extensionality of Causation and Causal-Explanatory Contexts, *Phil. of Science* 43, pp. 266–277.
- Lewin, K. (1923), Die zeitliche Geneseordnung. in: C.-F. Graumann (ed.), *Kurt-Lewin-Werkausgabe*, vol. I (A. Métraux ed.), Bern: Hans Huber, and Stuttgart: Klett, 1981, pp. 213–232.
- Lewis D. (1973), *Counterfactuals*, Oxford: Basil Blackwell.
- (1979), Counterfactual Dependence and Time's Arrow, repr. in: Lewis (1986b).
- (1983), New Work for a Theory of Universals, *Australasian J. of Phil.* 61(4), pp. 343–377.
- (1986), Events, in: *Philosophical Papers, vol. 2*, Oxford: Oxford University Press.
- (1994), Humean Supervenience Debugged, *Mind* 103, pp. 473–490.
- Liu C. (1994), The Aharonov–Bohm Effect and the Reality of Wave Packets, *Brit. J. Phil. Sci.* 45, pp. 977–1000.



- Lloyd G. (1995), Ancient Greek concepts of causation in comparativist perspective, in: Sperber D., Premack A. and Premack J. (eds), *Causal Understandings in Cognition and Culture*, New York: Oxford University Press.
- Mackie J. L. (1965), Causes and Conditions, in: Sosa and Tooley (1993), pp. 33–55.
- (1974), *The Cement of the Universe*, Oxford: Clarendon Press.
- McGinn C. (1982), The Structure of Content, in: A. Woodfield (ed.), *Thought and Object*, Oxford, Clarendon Press.
- Mellor D.H. (1987), The singularly affecting facts of causation, in: *Matters of Metaphysics*, Cambridge University Press 1991.
- Menzies P. and Price H. (1993), Causation as a Secondary Quality, *British J. Phil. of Science* 44, pp. 187–203.
- Mill J.S. (1843), *A System of Logic, Ratiocinative and Inductive*, London, Parker; E. Nagel (ed.), *John Stuart Mill's Philosophy of Scientific Methods*, Macmillan, New York 1950.
- Moravcsik J. (1991), "All A's are B's": Form and Content, *J. of Pragmatics* 16, pp. 427–441.
- Mulligan K., Simons P., and Smith B. (1984), Truth-Makers, *Phil. and Phen. Res.* 44, pp. 287–321.
- Mulligan K. and Smith B. (1986), A Relational Theory of the Act, *Topoi* 5, pp. 115–130.
- Mulligan K. (1991), Colours, Corners and Complexity: Meinong and Wittgenstein on Some Internal Relations, in: Spohn *et al.*, pp. 77–101.
- Mumford S. (1998), *Dispositions*, Oxford, Oxford University Press.
- Nagel E. (1961), *The Structure of Science*, London: Routledge and Kegan Paul.
- Otte R. (1981), A Critique of Suppes' Theory of Probabilistic Causality, *Synthese* 48, pp. 167–189.
- (1985), Probabilistic Causality and Simpson's Paradox, *Phil. of Science* 52, pp. 110–125.
- Papineau D. (1985), Causal Asymmetry, *British Journal for the Philosophy of Science* 36, pp. 273–289.
- Pariante J.-C. (1982), Le nom propre et la prédication dans les langues naturelles, *Langages* 66, pp. 37–65.
- Parsons, T. (1990), *Events in the Semantics of English: A Study in Subatomic Semantics*, Cambridge, MA: MIT Press.
- Pereboom D. and Kornblith H. (1991), The Metaphysics of Irreducibility, *Phil. Studies* 63, pp. 125–145.
- Peshkin M. and Tonomura A. (1989), *The Aharonov-Bohm Effect: Lecture Notes in Physics*, No. 340, Berlin: Springer.
- Pettit P. (1993), A Definition of Physicalism, *Analysis* 53.4, pp. 213–223.
- Pietrosky P. and Rey G. (1995), When Other Things Aren't Equal: Saving *Ceteris Paribus* Laws from Vacuity. *Brit.J.Phil.Sci.* 46, pp. 81–110.
- Plato (1961), *Sophist*, translated by Francis Macdonald Cornford, in: *The Collected Dialogues of Plato, Including the Letters*, Edith Hamilton and Huntington Cairns (eds), Bollingen Series 71, Princeton, NJ: Princeton University Press, pp. 957–1017.
- Poincaré H. (1905), Sur la dynamique de l'électron, in: *Oeuvres* vol. IX, Paris Gauthier-Villars, 1954.
- Popper K.R. (1934), *Logik der Forschung*, 10th edition, Tübingen: J.C.B. Mohr, 1994, transl. by the author, with the assistance of Dr. Julius Freed and Lan Freed: *The Logic of Scientific Discovery* (1959), London and New York: Routledge, 1992.

- (1949), A Note on Natural Laws and So-called “Contrary-to-Fact Conditionals”, *Mind* 58, pp. 62–66.
- (1956–7), *Nature* 177(1956), p. 538; *Nature* 178(1956), p. 382; *Nature* 179 (1957), p. 1297; *Nature* 181(1957), p. 402.
- (1972), *Objective Knowledge: An Evolutionary Approach*, Oxford: Clarendon Press, revised edition, 1986.
- Popper, K. R. (1983), *Realism and the Aim of Science, from the Post-Script to the Logic of Scientific Discovery*, W.W. Bartley, III (ed.), Totowa, NJ: Rowman and Littlefield.
- Posch G. (1981), Zur Problemlage beim Kausalitätsproblem, in: Günter Posch (ed.), *Kausalität. Neue Texte*, Stuttgart: Philipp Reclam.
- Price H. (1991), Agency and Probabilistic Causality, *British J. Phil. of Science* 42, pp. 157–176.
- (1992), Agency and Causal Asymmetry, *Mind* 101, pp. 501–520.
- Putnam H. (1967), The Nature of Mental States, in: Rosenthal D. (ed.), *Materialism and the Mind–Body–Problem*, Englewood Cliffs, NJ: Prentice Hall, 1971.
- Quine W. V. O. (1948), On what there is, in: Quine 1953.
- (1961), Two Dogmas of Empiricism, in W.V.O. Quine, *From a Logical Point of View: 9 Logico-Philosophical Essays*, 2d edition (first ed. 1953), Harvard University Press, Cambridge, MA, repr. 1980, pp. 20–46.
- (1960), *Word and Object*, Cambridge, MA: Technology Press of MIT.
- (1965), *The Ways of Paradox*, New York: Random House.
- (1969), Natural Kinds, in: Rescher *et al.* (1969).
- (1973), *The Roots of Reference*. LaSalle, IL: Open Court.
- (1985), Events and Reification, in: LePore and McLaughlin (1985).
- Railton P. (1978), A Deductive-Nomological Model of Probabilistic Explanation, in: Joseph C. Pitt (ed.), *Theories of Explanation*, New York: Oxford University Press, 1988.
- (1981), Probability, Explanation, and Information, *Synthese* 48, pp. 233–256.
- (1989), Explanation and Metaphysical Controversy, in: Kitcher and Salmon 1989.
- Ramsey F.P. (1927), Facts and Propositions, in: Ramsey 1990.
- (1928), Universals of Law and of Fact, in: Ramsey 1990.
- (1929), General Propositions and Causality, in: Ramsey 1990.
- (1990), *Philosophical Papers*, D.H. Mellor (ed.), Cambridge: Cambridge University Press.
- Reichenbach H. (1928), *Philosophie der Raum-Zeit-Lehre*, Berlin, de Gruyter; ed. anglaise: *The Philosophy of Space and Time*, New York: Dover, 1958.
- (1936), Über Induktion und Wahrscheinlichkeit, *Erkenntnis* 5, pp. 267–284.
- (1951), *The Rise of Scientific Philosophy*, Berkeley, CA: University of California Press.
- (1956), *The Direction of Time*, Berkeley, CA: University of California Press, 1991.
- Rescher N. *et al.* (ed.) (1969), *Essays in Honor of Carl G. Hempel*, Dordrecht: Reidel.
- Rosenberg A. and Martin R.M. (1979): The Extensionality of Causal Contexts, *Midwest Studies in Philosophy* IV, pp. 401–408.
- Ruby J.E. (1986), The Origins of Scientific “Law”, *Journal of the History of Ideas* 47(3), pp. 341–359.
- Russell B. (1912), The Concept of Cause, repr. in: *Mysticism and Logic*, London, Unwin, 1986, pp. 173–199, and in: *The Collected Papers of Bertrand Russell*, vol.

- 6: *Logical and Philosophical Papers 1909–1913*, pp. 193–210; page references are to the latter edition.
- (1914): *Our Knowledge of the External World*. London: Allen and Unwin, 1926.
- (1918), *The Philosophy of Logical Atomism*, LaSalle, IL: Open Court, 3rd edn, 1990.
- (1948), *Human Knowledge: Its Scope and Limits*, London: Allen and Unwin, 5th edn, 1966.
- Sachs R.G. (1987), *The Physics of Time Reversal*, Chicago, IL: University of Chicago Press.
- Salmon W. (1978), Why Ask “Why?”? – An Inquiry Concerning Scientific Explanation, *Proceedings and Addresses of the American Philosophical Association* 51(6), pp. 683–705.
- (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton, NJ: Princeton University Press.
- (1990), *Four Decades of Scientific Explanation*, Minneapolis, MN: University of Minnesota Press.
- (1994), Causality Without Counterfactuals, *Phil. of Science* 61, pp. 297–312.
- Scheibe E. (1991), EPR-Situation and Bell’s Inequality, in: Spohn *et al.* 1991, pp. 115–129.
- Schiffer S. (1991), Ceteris Paribus Laws, *Mind* 100, pp. 1–17.
- Schlick M. (1931), Causality in Contemporary Physics, *Brit. J. Phil. Science* 12 (1961): 177–193, 281–298.
- Scriven M. (1961), The key property of physical laws – inaccuracy, in: Herbert Feigl and Grover Maxwell (eds.), *Current Issues in the Philosophy of Science*, New York: Holt, Rinehart & Winston.
- (1975), Causation as Explanation, *Noûs* 9, pp. 3–16.
- Skyrms B. (1980), *Causal Necessity*, New Haven, CT: Yale University Press.
- (1984), EPR: Lessons for Metaphysics, in: *Midwest Studies in Philosophy* IX, pp. 245–255.
- Sober E. (1981), Evolutionary Theory and the Ontological Status of Properties, *Phil. Studies* 41, pp. 147–176.
- (1982), Frequency-Dependent Causation, *Journal of Philosophy* 79, pp. 247–253.
- (1985), A Plea for Pseudo-Processes, *Pacific Philosophical Quarterly* 66, pp. 303–309.
- Sorensen R. (1990), Process Vagueness, *Linguistics and Philosophy* 13, pp. 589–618.
- Sosa E. and Tooley M. (eds.) (1993), *Causation*, Oxford: Oxford University Press.
- (1993a), Introduction, in: Sosa and Tooley (1993), pp. 1–32.
- Sperber D. and Wilson D. (1986), *Relevance*, Blackwell: Oxford.
- Spohn W., van Fraassen B.C. and Skyrms B. (eds.) (1991), *Existence and Explanation*, Dordrecht: Kluwer.
- Stalnaker R. (1968), A Theory of Conditionals, in: W. L. Harper, R. Stalnaker, and G. Pearce (eds.), *Ifs. Conditionals, Belief, Decision, Chance, and Time*, D. Reidel: Dordrecht, 1981.
- Stalnaker, R. C. (1984), *Inquiry*, Cambridge, MA: MIT Press.
- Stegmüller W. (1966), Explanation, Prediction, Scientific Systematization and Non-explanatory Information, *Ratio* 8, pp. 1–24.

- (1969), *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, vol. 1: *Wissenschaftliche Erklärung und Begründung*, Berlin: Springer.
- Suppes P. (1970), *A Probabilistic Theory of Causality*, Amsterdam: North Holland Publishing Company.
- Tooley M. (1977), The Nature of Laws, *Canadian J. of Phil.* 7(4), pp. 667–698.
- (1987), *Causation: A Realist Approach*, Oxford: Clarendon Press.
- (1990a) Causation: Reductionism Versus Realism, *Phil. and Phen. Res.* 50, Suppl., pp. 215–236.
- (1990b) The Nature of Causation: A Singularist Account, *Canad. J. of Phil. Suppl.* Vol. 16, pp. 271–322.
- van Fraassen B. (1980), *The Scientific Image*, Oxford: Clarendon.
- (1989), *Laws and Symmetry*, Oxford: Clarendon.
- (1993), Armstrong, Cartwright, and Earman on *Laws and Symmetry*, *Phil. and Phen. Research* 53, pp. 431–444.
- von Wright G.H. (1947), *Logical Studies*, London: Routledge.
- (1971), *Explanation and Understanding*, London, Routledge and Kegan Paul.
- (1973), On the Logic and Epistemology of the Causal Relation, in: Sosa and Tooley (1993), p. 105–124.
- (1983), On Causal Knowledge, in: C. Ginet and S. Shoemaker (eds.), *Knowledge and Mind*, New York: Oxford University Press.
- (1984), Logical Modality, in: *Truth, Knowledge and Modality, Philosophical Papers, Vol. 3*, Oxford: Basil Blackwell.
- Wieland W. (1970), *Die aristotelische Physik*, Göttingen: Vandenhoeck & Ruprecht.
- Williams D. (1953), On the Elements of Being, *Rev. of Metaphysics* 7, pp. 3–18 and 171–192.
- Wilson, R. A. (1995), *Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind*, Cambridge: Cambridge University Press.
- Wittgenstein, L. (1964), *Schriften 2: Philosophische Bemerkungen*, Frankfurt a.M., Suhrkamp; transl. *Philosophical Remarks*, ed. from his Posthumous Writings by Rush Rhees and translated into English by Raymond Hargreaves and Roger White, Oxford, Basil Blackwell, 1975.
- Yablo S. (1992a), Mental Causation, *Phil. Rev.* 101, pp. 245–280.
- (1992b), Cause and Essence, *Synthese* 93, pp. 403–449.
- Zahar E.G. (1982), The Popper–Lakatos Controversy, *Fundamenta Scientiae* 3, pp. 21–54.
- Zangari M. (1992), Adding Potential to a Physical Theory of Causation, *PSA 1992*, Vol. I, pp. 261–273.
- Zucchi A. (1993), *The Language of Proposition and Events*, Dordrecht: Kluwer.

# Index

- a posteriori 26, 224  
absence 202, 205, 207  
accidental universal correlation 75, 77–8, 85, 87–8, 93, 96–7, 100, 122, 137, 170, 219–20, 234  
Achinstein P. 245  
action: *see* intervention  
anomalous monism 101, 238  
Anscombe G.E.M. 16–17  
anthropocentric, anthropomorphic 4, 21, 25, 71, 226  
antisymmetry 141, 183, 185–6  
applicability condition 99, 104, 107–9, 113, 115–18, 120–1, 123–6, 134–7, 144, 178, 221, 238, 240–1  
Aristotle 4–5, 90, 202, 225, 249  
Armstrong, D.M. 90–1, 94–5, 117–19, 171–3, 220, 225, 236–7  
Aronson, J.J. 39, 45, 47, 68–70, 197, 205, 250  
asymmetry 20, 30–4, 36, 38, 72, 142, 183, 185–7, 197, 217, 227–8, 242
- basic statement 131  
Bell, J.S. 28  
Bennett J. 209, 244  
Blackburn S. 23, 245  
Bohm, D. 29  
Braithwaite, R.B. 78, 234
- Carnap R. 2, 50, 131, 140–1, 144–6, 165, 224–6  
Carroll J. 247  
causal law 180, 183, 186, 245–6  
causal line 51, 230  
causal network: *see* network  
causal responsibility 7, 8, 10, 140, 142, 146, 148, 150, 156–60, 163–4, 166, 177, 179, 188–9, 192, 195–6, 199–200, 203, 208–11, 214–18, 220–4, 233, 248–9  
Causality, Principle of 1  
causation, backwards 20, 34, 38–9, 72, 226, 228  
causation, mental 223  
causation, simultaneous 20, 27–8, 39–40, 42–3, 48, 50, 70, 72  
causation, static 43  
ceteris paribus 6, 94, 100–2, 134, 221, 238, 240  
change, non-change 20, 47, 201–2, 211, 229, 231  
Clendinnen J.F. 241  
common cause 143, 216, 246  
concept 81–5, 88–9, 235, 242  
conserved quantity 9, 26, 47, 51, 53, 55, 57, 58, 60–1, 64–6, 68–9, 71, 73, 136, 144, 151, 215–16, 222, 232  
context: *see* opaque context, semi-opaque context, transparent context  
contiguity, non-locality 12–16, 19, 26–7, 29, 227, 229  
counterexample 104, 133, 225, 241  
counterfactual 122–3, 137, 172, 209–11, 245  
Cummins R. 143
- Davidson D. 100–2, 123, 145–7, 153, 159–60, 165–6, 238, 242, 245  
description 141–2, 145, 191, 242, 249  
determinable, determinate 174  
deterministic 100–1, 213, 238  
Dieks D. 29, 39, 54–6, 68–9  
disposition 168, 244–5  
Dowe P. 29–30, 56, 58, 65–7, 70, 208, 232  
Dretske F. 80–1, 85–6, 89–93, 95, 97, 100, 128, 220, 235

## 262 *References*

- Ducasse 17, 19–20, 225  
 Duhem P. 133–4, 137, 240
- Ehring D. 29, 47, 204, 206, 208  
 Einstein A. 27–8, 229  
 emphasis, highlighting 188–92, 194, 248  
 enthymeme 103–4  
 epiphenomenon 143–4  
 epistemology, epistemological 22–5, 28,  
 72–3, 140, 144, 173, 180–1, 230–1,  
 234  
 event 7, 10, 15, 26, 45, 48, 59–60, 63–  
 64, 68, 139, 141–2, 144–5, 147–9,  
 154–6, 163–5, 176, 186–9, 191, 196,  
 198–201, 203–4, 208, 210–12, 214–16,  
 218, 222, 224–5, 238, 242–3, 248  
 eventive (statement, expression) 149,  
 151, 153–8, 161–5, 197–8, 200, 204,  
 208, 210, 218, 221, 248  
 exception 76, 92–4, 97–101, 109–10,  
 114–16, 118–20, 123, 125–7, 129,  
 133–5, 137–8, 147, 184, 219, 234,  
 236–7, 239  
 explanation 3–5, 8, 11, 39, 48–9, 52, 95,  
 103, 140–8, 153, 162, 165–6, 173, 176–  
 7, 187, 193, 195–7, 205, 211, 214, 216–  
 17, 225, 230, 232, 237, 241, 249  
 explanation, deductive-nomological 2,  
 100, 107, 217  
 explanation, non-causal 4, 49, 143–4,  
 146  
 explanation, teleological 144  
 extensional 150, 152–3, 159, 165, 190–  
 1, 193–4, 235, 242, 248–9
- fact 7, 10, 48, 140, 142, 146–7, 149,  
 154–6, 159–66, 176, 188, 196–200,  
 203–4, 211, 216, 218, 224–5, 233, 243  
 factive (statement, expression) 149–50,  
 153–9, 161–6, 196–7, 199, 203–4,  
 208, 210–11, 218, 248–9  
 Fair D. 29–30, 38–9, 45, 68–70, 196,  
 198, 207–8  
 falsification 131–3, 137–8, 240  
 flow 56, 60, 72  
 Fodor J. 101–2, 125–7, 137, 239–40  
 Føllesdal D. 194–5, 249  
 formal, material (mode of speech) 13–14  
 Frege G. 81–90, 97, 100, 159, 235–6  
 Fuhrmann A. 236
- Gasking D. 20–2, 41–2, 226  
 genidentity 50, 55, 230, 231
- Goodman N. 78, 122, 168, 170, 245  
 group velocity 60–1, 67  
 grue 172–3, 186, 236, 245  
 Grünbaum A. 231
- Hempel C.G. 2, 110–11, 114, 116, 134,  
 225, 238  
 heteronomic: *see* homonomic  
 highlighting: *see* emphasis  
 Holt P.J. 168–70  
 homonomic, heteronomic 123–5, 137  
 Horgan T. 160, 249  
 Hume D. 2, 9–21, 23, 25–7, 40–2, 44–5,  
 71–3, 224–5, 229  
 Humphreys P. 210, 241
- « identical » particles 56  
 indeterministic 41  
 inductive inference 114, 131–2, 138, 170  
 inference problem 127–8, 130, 137  
 initial conditions 2, 141, 166, 217, 237  
 Instantiation, Principle of 90  
 intervention, action 21–2, 24, 29, 58,  
 64, 71, 144, 187, 241–2  
 intrinsic 7, 168–9, 175, 203, 215, 228,  
 244, 246  
 irreflexive 141  
 irreversible 197
- Kepler J. 132  
 Kneale W. 79  
 Krajewski W. 70, 250
- Lakatos I. 134, 238  
 law (of nature) 2, 11, 75, 78, 80–1, 85,  
 88, 90–1, 93–4, 96, 102–6, 108–10,  
 113–14, 117–18, 120–5, 127, 129–30,  
 134, 136–9, 144–7, 151, 171–4, 176,  
 178, 182, 185, 213–14, 217, 219–22,  
 233–4, 236–7, 240, 245–6  
 law of association 177, 179, 186, 246  
 law, best-system analysis, 79, 171–3,  
 245  
 law, causal: *see* causal law  
 law of conservation 6, 67–9, 74, 124,  
 136, 138–9, 147, 163, 186, 215–16,  
 219, 246  
 law of evolution 177, 179, 186  
 law, strict 6, 76, 91, 98–101, 119, 123,  
 125, 134–6, 221, 238–9  
 law, iron 91, 119, 134  
 law, non-strict 79, 100, 102, 119, 123,  
 239

- law, oaken 118–19  
 Leibniz G. 235  
 Lemmon E.J. 45–6  
 Levin M.E. 249  
 Lewin K. 50  
 Lewis D. 79, 171–3, 227–8, 245
- Mackie J.L. 246  
 mark 58, 64  
 material (mode of speech): *see* formal (mode of speech)  
 Mellor D.H. 199–204, 249  
 Mill J.St. 19–20, 171, 226  
 multirealizability 125–7, 240
- Nagel E. 3, 4  
 natural change 47, 225, 231  
 natural kind 172, 227  
 natural property 167, 170–1, 173–5, 186, 217, 220–2, 244  
 necessity 15–17, 19, 23, 79, 95–6, 128, 180, 183, 194–5, 245  
 negation, negative 156–8, 196, 199–201, 204–5, 207, 211  
 network, causal 30–5, 72, 228  
 Newton I. 132, 135, 229, 233  
 nomic, nomological 2, 5, 7, 74–5, 77, 79–80, 85–6, 88–9, 91–2, 96–7, 99–100, 102–4, 130, 133, 135, 138–9, 162–3, 174–5, 186, 189, 192, 205–7, 214–17, 219, 230, 240–1, 245, 247–8  
 nominalist conception of laws 109–10, 114  
 non-causal (law, explanation) 23, 29, 72, 205–7, 213–14, 217, 226, 230, 232, 241, 245  
 non-locality: *see* contiguity
- observable, theoretical 110–12, 114, 116–17, 133–4, 239  
 opaque context 80, 86–8, 91, 96, 150, 152–3, 189–90, 219–20, 235, 242
- Papineau D. 247  
 Parsons T. 159  
 persistence 48, 51, 54, 56, 73, 147, 170, 231  
 phase velocity 60–2, 67  
 Pietroski P. 241  
 Plato 128–9, 224  
 Poincaré H. 229–30  
 Popper K.R. 131–3, 138, 166, 225, 227, 234–5, 237
- possible world 183, 226, 228  
 prediction 141, 237  
 process 2, 54, 57–9, 63–5, 67, 70, 74, 144, 155, 175, 186, 193, 195, 207, 214–15, 226, 232–3  
 probability, probabilistic, statistical 41, 100, 126, 168, 180–1, 183, 185, 213, 216, 236, 239, 246, 247  
 proviso 102, 116, 123, 134, 144, 238  
 pseudo-process 54, 57–67, 73–4, 139, 144, 176–7, 179, 186, 214, 232
- quantum mechanics 27–9, 136, 227, 231  
 Quine W.V.O. 46–7, 133–4, 137, 152, 173–4, 217, 240, 242, 250
- Railton P. 144  
 Ramsey F.P. 79, 161–2, 171–3  
 realism (laws) 79, 99, 109, 116–18, 127–8, 174–5, 220, 238, 240  
 realism (properties, universals) 90, 109, 129, 236  
 reduction 26, 35–6, 72, 125, 139–40, 143, 146, 165, 208, 217, 222, 225–6, 232, 235, 240, 250  
 regularity 13–14, 17–18, 95, 120, 135, 139, 225  
 Reichenbach H. 30, 131–2, 140, 240, 243  
 relational predicate 199, 203, 211, 215, 218, 223  
 responsibility, moral 202–3, 207, 209, 249  
 Rey G. 241  
 Russell B. 2, 3, 8, 50–2, 160, 187, 224, 229–31
- Salmon W. 54, 58–9, 61, 63–4, 70  
 Schiffer S. 102  
 Scriven M. 141–2, 239, 241  
 semi-opaque context 152, 242  
 Sober E. 143, 232  
 Sosa E. 26, 29, 52  
 spatio-temporal 26, 32–3, 45–6, 48, 53, 59, 62, 73, 143–4, 148, 186, 204, 211, 215, 233, 244  
 special science 101–2, 123, 125, 221, 223, 240  
 Stalnaker R. 244  
 state, static 205, 229  
 statistical: *see* probability  
 Stegmüller W. 140, 165  
 subjective 11–16, 23, 196

## 264 *References*

- substance 51, 53–4, 170  
substitution *salva veritate* 80, 86, 90–1, 152, 159, 212, 220, 235  
succession 12–16, 19, 32, 40, 53  
sufficient reason, principle of 40
- temporal part 51–2, 72–3, 146, 201, 203, 216, 224, 230  
theoretical: *see* observable  
theoretical ascent 111–12, 114, 238  
time, direction of 35–6, 72  
Tooley M. 26, 29, 35–7, 52, 180–6, 220, 225, 236, 246–7  
transference, transmission 9, 25–6, 38–9, 44, 50, 53–4, 57, 60–1, 67–9, 71–4, 142, 144, 146, 151, 179, 182, 197–8, 204–8, 215–18, 221–2, 226–7, 232–3, 246, 250  
transparent context 80, 86, 88, 90, 96, 219, 235
- triggering 204–8  
trope 57, 76, 236–7  
truth-maker 146
- vague 234  
Van Fraassen B. 127–30, 137, 171–2, 236, 240, 245  
velocity: *see* phase velocity, group velocity  
Von Wright G.H. 22–5, 41–2, 237
- Williams D. 94  
Wilson R. 244  
Wittgenstein L. 237  
world line 59, 65, 70
- Zangari M. 43–4  
Zucchi A. 161