carried out by nonphilosophers on causation and related matters. For example, the "definitions" of "causal effect" that one finds in writers like Pearl (2000a) and Holland (1986) are not just attempts to describe commonly accepted usage—either of scientists or of ordinary people—and should not be evaluated just on the basis of whether they provide such a description. Again, they have, to be sure, an important continuity with ordinary usage and with scientific practice—otherwise, they could hardly claim to be characterizations of "causal effect" in any sense—but they are also intended as clarifications or regimentations of that usage and practice, introduced with certain purposes in mind (e.g., statistical applications). In fact, I believe that a similar point holds for much of the philosophical literature on cause and explanation: this is also not "conceptual analysis" in the purely descriptive sense described above. We lack an accurate, common, accepted vocabulary for describing the activity carried on in this literature, but it is legitimate and important nonetheless.[2]

1.3 The Manipulability Conception of Causal Explanation

I turn now to some brief remarks that are intended to illustrate and motivate the manipulability conception; details come in subsequent chapters. I emphasize that my aim at this point is simply to sketch the general picture I advocate in a very rough (some may think reckless) way; qualifications and refinements are added later.

The manipulability conception plays an important role in the way that scientists themselves think about causal explanation but has received rather less attention from philosophers. The basic idea is nicely illustrated by a contrast drawn between descriptive and explanatory science in a paper by Robert Weinberg (1985) on recent developments in molecular biology. Weinberg tells us that "biology has traditionally been a descriptive science," but that because of recent advances, particularly in instrumentation and experimental technique, it is now appropriate to think of molecular biology as providing "explanations" and identifying "causal mechanisms." What does this contrast between description and explanation consist in? Weinberg explicitly links the ability of molecular biology to provide explanations with the fact that it provides information of a sort that can be used for purposes of manipulation and control. New experimental and instrumental techniques have played such a decisive role in the development of molecular biology into an explanatory science precisely because such techniques make it possible to intervene in and manipulate biological systems and to observe the results in ways that were not previously possible. Molecular biologists correctly think that "the invisible submicroscopic agents they study can explain, at one essential level, the complexity of life" because by manipulating those agents it is now "possible to change critical elements of the biological blue print at will" (p. 48).

This passage suggests the underlying idea of my account of causal explanation: we are in a position to explain when we have information that is

relevant to manipulating, controlling, or changing nature, in an "in principle" sense of manipulation characterized in chapter 3. We have at least the beginnings of an explanation when we have identified factors or conditions such that manipulations or changes in those factors or conditions will produce changes in the outcome being explained. Descriptive knowledge, by contrast, is knowledge that, although it may provide a basis for prediction, classification, or more or less unified representation or systemization, does not provide information potentially relevant to manipulation. It is in this that the fundamental contrast between causal explanation and description consists. On this way of looking at matters, our interest in causal relationships and explanation initially grows out of a highly practical interest human beings have in manipulation and control; it is then extended to contexts in which manipulation is no longer a practical possibility. This interest is importantly different from a number of the other interests philosophers have associated with explanation, for example, from our interest in prediction or even in nomically grounded prediction, or from our interest in constructing theories that unify, systematize, and organize in various ways, or that trace spatiotemporally continuous processes. As we shall see, one can have information that is relevant to prediction (including prediction based on generalizations that many philosophers are prepared to regard as laws), or information about spatiotemporally continuous processes, or information that allows for the sort of unification and systemization that many philosophers have thought relevant to explanation, and yet lack the kind of information that is relevant to manipulation on which my account focuses. When this is the case, my view is that one doesn't have a (causal) explanation. Conversely, one can have information that is relevant to manipulation and hence to explanation, even though one lacks the other features described above. What one needs for manipulation is information about *invariant* relationships, and one can identify invariant relationships even in cases in which one doesn't know laws, cannot trace spatiotemporally continuous processes, or unify and systematize.

I said above that explanatory information is information that is potentially relevant to manipulation and control. It is uncontroversial, however, that causal relationships exist and that explanation is possible in circumstances in which actual manipulation is impossible, whether for practical or other sorts of reasons. For example, we construct causal explanations of past events and of large-scale cosmological events, and in neither case is manipulation of these phenomena possible. The notion of information that is relevant to manipulation thus needs to be understood modally or counterfactually: the information that is relevant to causally explaining an outcome involves the identification of factors and relationships such that *if* (perhaps contrary to fact) manipulation of these factors *were* possible, this would be a way of manipulating or altering the phenomenon in question. For example, it is currently believed that the explanation (1.2.1) of the mass extinctions at the end of the Cretaceous period has to do with the impact of a large asteroid and the killing effects of the dust it created. Clearly, we cannot now do anything to affect whether this impact occurred, and quite possibly humans could have

done nothing to alter the impact even if they had existed with current levels of technology at the time of impact. My suggestion is that if this explanation is correct, it nonetheless will be true that if it had been possible to alter or prevent the impact, this would have altered the character of or prevented the extinction. Put differently, my idea is that one ought to be able to associate with any successful explanation a hypothetical or counterfactual experiment that shows us that and how manipulation of the factors mentioned in the explanation (the *explanans*, as philosophers call it) would be a way of manipulating or altering the phenomenon explained (the *explanandum*). Put in still another way, an explanation ought to be such that it can be used to answer what I call a *what-if-things-had-been-different question*: the explanation must enable us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways. We can also think of this as information about a pattern of counterfactual dependence between explanans and explanandum, provided the counterfactuals in question are understood appropriately. As we shall see, even when actual manipulation is impossible, it is heuristically useful to think of causal and explanatory claims in this way: it both clarifies their content and enables us to understand why they have many of their distinctive features.

On this view, our interest in causal explanation represents a sort of generalization or extension of our interest in manipulation and control from cases in which manipulation is possible to cases in which it is not, but in which we nonetheless retain a concern with what would or might happen to the outcome being explained if various possible changes were to occur in the factors cited in the explanans. If we had been unable to manipulate nature—if we had been, in Michael Dummett's (1964) example, intelligent trees capable only of passive observation—then it is a reasonable conjecture that we would never have developed the notions of causation and explanation and the practices associated with them that we presently possess. Once developed, these notions and practices were then extended to contexts in which actual manipulation was infeasible or impossible. This extension was very natural and perhaps inevitable because, as we shall see in chapter 3, it is built into the notion of a relationship that is usable for purposes of manipulation and control that whether such a relationship holds does not depend on whether the manipulation in question can be actually carried out.

Although it will not be news to historians that the aim of manipulating or controlling nature has played a central role in the development of modern science, this aim has received relatively little attention from philosophers. Most philosophers have distinguished sharply between pure science and applied science or technology, and have regarded explanation as a characteristic aim of pure science and manipulation and control as aims of applied science. To the extent that philosophers have concerned themselves with applied science, they often have seen it as primarily focused on prediction and have failed to appreciate how different prediction is from control. To readers in the grip of this conventional picture of science, my association of our interest

in explanation with our practical interest in control over nature will seem misguided and counterintuitive.

However, a variety of more recent developments in history and philosophy of science and in science studies challenges this sharp distinction between pure science and its application. Part of my intention in writing this book is to contribute to a conception of the role of causal explanation in science that fits with these new developments. I include among these developments recent work in the history of science that emphasizes how concerns with technological application have heavily influenced the content of the more theoretical parts of science (e.g., Smith and Wise 1989; Barkan 1999) and recent work by sociologists, philosophers, and historians on experimentation, which has emphasized in various ways how important our ability to intervene and manipulate nature is in the development of a scientific understanding of nature. Broadly similar ideas can be found in some of the recent philosophical literature on explanation, for example. in Paul Humphreys's (1989) work, with its criticisms of "passive empiricism." On the conception of science that I favor, two aims that are often regarded as quite separate—the "pure science" aim of representing nature in a way that is truthful and accurate and the "applied science" aim of representing nature in a way that permits manipulation and control—are deeply intertwined.

My association of our interest in explanation with our interest in manipulation and control will also seem less surprising when one reflects that a very central part of the commonsense notion of cause is precisely that causes are potential handles or devices for bringing about effects. We find this idea in the manipulability theories of causation defended by writers like Collingwood (1940), Gasking (1955), and von Wright (1971). As we shall see in chapter 2, it is also widely endorsed by social scientists and statisticians, who have shown that this idea can play an important heuristic role in both elucidating the meaning of causal claims and clarifying how statistical evidence can be used to test them. Unfortunately, however, standard philosophical statements of the manipulability theory lead to accounts of causation that are unacceptably anthropocentric and subjectivist. I show in chapter 2 how a manipulability account of causation/explanation can be developed in a way that satisfies reasonable expectations about the objectivity of causal relationships.

1.4 Causal Explanation, Invariance, and Intervention Illustrated

My discussion so far has been rather abstract. It will be useful to have a concrete example in front of us to illustrate some of the ideas to which I have been referring. Consider (1.4.1) a block sliding down an inclined plane with acceleration a (fig. 1.4.1). What accounts for or explains the motion of the block? The standard textbook analysis proceeds as follows. The block is subject to three forces: a gravitational force due to the weight of the block; a normal force N, which is perpendicular to the plane; and a force due to