Human Nature and the Limits of Science

JOHN DUPRÉ

CLARENDON PRESS • OXFORD
2001

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York
Athens Auckland Bangkok Bogotá Buenos Aires
Cape Town Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul
Karachi Kolkata Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi Paris São Paulo Singapore Taipei Tokyo Toronto Warsaw

with associated companies in Berlin Ibadan

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2001

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

ISBN 0–19–924806–0

1 3 5 7 9 10 8 6 4 2

Typeset by Invisible Ink
Printed in Great Britain
on acid-free paper by
T. J. International Ltd.,
Padstow, Cornwall

7

Freedom of the Will

1. Introduction

So far this book has mainly been about science, or some pretenders to the title of science. It may therefore seem surprising that I choose to conclude the book with a discussion of one of the most ancient and traditional problems of pure philosophy, freedom of the will. I hope the surprise will be short-lived, however. It is the sciences of the human mind, and their aspirations to provide complete explanations for human behaviour, that have naturally focused recent worries about the intelligibility of free action. And, often no doubt at a somewhat inchoate level, it is concerns about the threat to human autonomy that have motivated hostility to some of these scientific, or scientistic, projects. My own view is that the connection between these issues is more indirect. It is the metaphysics that underlies and motivates scientism that also grounds doubts about the possibility of human autonomy. In other words, the sciences that seem superficially inimical to human freedom are based on a metaphysics that really does preclude freedom. My aim in this chapter is to show how disposing of both the bad metaphysics and its scientistic spawn opens the way for a proper account of human autonomy. Such an account finally is one of the vital ingredients we need for a more complete understanding of humans in the societies that are their natural and necessary environments.

The argument of this book so far has been that a proper understanding of a domain as complex and richly connected to diverse factors as that of human behaviour can only be adequately approached from a variety of perspectives. But it is not at all obvious how concerns about human autonomy are defused by the move from a single reductive approach to explanation to a pluralistic approach. In this chapter I try to show how the pluralistic framework I have defended in previous chapters does indeed leave space for real human autonomy. I begin by defending the thesis, in sympathy with most untu-

tored convictions but contrary to the orthodox philosophical view of the subject, that the denial of determinism does remove some of the central difficulties in providing a proper account of human freedom. I shall also briefly explain what I take to be some of the grounds and consequences of the denial of determinism.[1]

2. Free Will and Determinism

It has notoriously been supposed that the doctrine of determinism conflicts with the belief in human freedom. Yet it is not readily apparent how indeterminism, the denial of determinism, makes human freedom any less problematic. It has sometimes been suggested that the arrival of quantum mechanics should immediately have solved the problem of free will and determinism. It was proposed, perhaps more often by scientists than by philosophers, that the brain would need only to be fitted with a device for amplifying indeterministic quantum phenomena for the bogey of determinism to be defeated. Acts of free will could then be those that were initiated by such indeterministic nudges. Recently there has been some inclination to revive such a story as part of the fallout from the trend for chaos theory. Chaotic systems in the brain, being indefinitely sensitive to the precise details of initial conditions, seem to provide fine candidates for the hypothetical amplifiers of quantum events.[2]

But this whole idea is hopeless, and appeals to quantum mechanics merely illustrate the hopelessness.[3] To see this, one needs only

[1] This task is undertaken in greater detail in Dupré, 1993a: chs. 8 and 9.

[2] More recent appeals to exotic physics for the solution of this problem have involved Francis Everitt's many-worlds interpretation of quantum mechanics. On one reading of this, the world is constantly splitting into two, and one cause of some such splittings might be free acts of will. On what appears to be a more sophisticated version of the theory there is only one world, but it consists of a vast superposition of coexisting states. One philosopher who thinks this view helps us to understand the mind is David Chalmers (1996). Although Chalmers does not directly address questions of freedom, he presents the individual mind as simply one among all the possible paths through the superpositions of minds that each mind is at any moment part of (or so I understand him; see especially his p. 353). This view, I suppose, explains the illusion of freedom as made possible by a certain kind of randomness. My own view is that this is one of the several points in Chalmers's book where he has failed to draw the conclusion of an argument *modus tollens*.

[3] No doubt the belief in the existence of indeterministic events at the root of free action was often also connected with the inchoate hope that these might be sufficiently loose and microscopic that even an immaterial soul might have a chance of subtly influ-

to consider that the interest in establishing free will is not the conviction that humans are random action generators, but a concern that human autonomy is inconsistent with the possibility of fully explaining human actions in terms that have no apparent connection with the wishes and beliefs of the human agent. Standard compatibilist claims that human autonomy and mechanistic causal explanation are not mutually exclusive may or may not be defensible. But the attempt to reconcile human autonomy with the complete randomness of human actions is surely a dead end.[4] At first sight it appears that, despite the initial worries about determinism, indeterminism makes the conception of freedom of the will even less supportable.

A great deal of recent discussion of these issues has concerned the question whether, in saying an action was free, we imply that the agent could have done otherwise. It is generally assumed that if determinism is true, then the agent could not have done otherwise and therefore that, if the implication just mentioned holds, we cannot be free. The problem that will be clear from the preceding discussion is that even if determinism is false, it is far from obvious what could make it true that the agent could have done otherwise in a way that does anything to illuminate any doctrine of free action. In particular, we had better not interpret it as meaning just that the action was produced only with probability, not certainty (this may be true; it just doesn't add anything relevant to freedom). At any rate, I am not much concerned here with the question of whether the agent could have acted otherwise—my inclination, in fact, is to think this is largely irrelevant to questions of freedom. My concern is rather with autonomy, in the sense that the agent can be seen as truly the originator of an action. This idea can be found most clearly in the concept of agent causation, deriving from Roderick Chisholm (1964). This is the concept of an agent as an initial source, rather than merely an intermediate link in the chain, of causality—or as he says, a 'prime mover'. For Chisholm it is an idea directly motivated by the thought that a responsible agent must have the power either to act or to refrain from acting. Chisholm also seems to think that this

encing them. Though this makes the idea less absurd from the point of view of understanding human autonomy, it introduces new absurdities that I cannot attempt to address here.

[4] This point was clearly stated by C. D. Broad (1952).

is a kind of causality unique to agents. I differ from Chisholm, first, in doubting the importance of alternative possibilities, and second, in that I take agent causation to be much more similar to causation in the non-human realms than he allows.

My aim in this chapter, then, will be to show that the solution to the problem of the freedom of the will does lie, despite familiar objections, with the truth of indeterminism. But the point of this demonstration will not be to defend the idea of alternate possibilities (though I am, in the end, sympathetic to this), but to provide a proper account of autonomy in something like the sense of agent causation. The first step in this account is to distinguish two very different grades of indeterminism. The indeterminism entailed by the common understanding of quantum mechanics, while it denies that the causal upshot of a situation is a determinate function of any fact about that situation, still insists that there is a complete causal truth about every situation. It is just that this truth is in the form, not of a unique outcome, but of a range of outcomes with specific probabilities attached to their occurrence. Thus situations are still conceived as evolving according to laws, just laws of a somewhat different kind. I shall refer to both traditional determinism, and this brand of moderate indeterminism, as versions of the thesis of *causal completeness.* Even if determinism is false, causal completeness requires that there be some quantitatively precise law governing the development of every situation. If we maintain the doctrine of causal completeness, then the only retreat from physical determination of our actions is in the direction of more or less unreliability, hardly a desirable philosophical goal. However, the indeterminism that I wish to advocate is something quite different, the denial of causal completeness.[5] I shall maintain that few, if any, situations have a complete causal truth to be told about them. Causal regularity is a much rarer feature of the world than is generally supposed. And the real solution to the problem of freedom of the will, I shall argue, is to recognize that humans, far from being putative exceptions to an otherwise seamless web of causal connection, are in fact dense concentrations of causal power in a world where this is in short supply.

The solution to the problem of human autonomy that I propose, then, is a complete reversal of traditional non-compatibilist

[5] Causal completeness remains the orthodox assumption in the philosophy of science. It has been criticized in most detail by Cartwright (1999).

approaches. Such solutions have assumed that the non-human world consists of a network of causal connections, the links in which instantiate lawlike, exceptionless generalizations, but have tried to show that humans, somehow, lie outside or partially outside this web.[6] By contrast, I am suggesting that causal order is everywhere partial and incomplete. There is no such causal web. But humans, by virtue of their enormously complex but highly ordered internal structure, provide oases of order and predictability; they are potent sources of causality. Thus the significance of recognizing indeterminism is not at all to show that human actions are unreliable or random. It is rather to show that the causal structure that impinges on human beings, whether externally from macroscopic causal interaction, or internally from constitutive microstructural processes, is not such as to threaten the natural intuition that humans are, sometimes, causally efficacious in the world around them.

This picture immediately accords with some obvious empirical facts: among the most apparently orderly features of the external world, such as straight roads and vertically stable edifices, not to mention complex machines, are products of human action; and among the most predictable entities in the world, as Hume, to a rather different purpose, argued, are people. Plans can be coordinated among many people, and complex human institutions can function, because human behaviour is to a substantial degree reliable; as Hume remarked, purses of gold don't remain undisturbed for long at Charing Cross. All of this is quite unproblematic if we see humans as sources of causal order rather than either as exceptions to a universal external order or as insignificant components of some all-encompassing cosmic order. Thus a radical rejection of the traditional mechanistic assumption of causal completeness does indeed do something to defuse the traditional problem of free will.

I shall expand on these claims in the course of this chapter. Prior to that, however, my main task will be to render its presuppositions plausible. The orthodox, though certainly not the universal, contemporary view of free will is compatibilist: it holds that everything we have any right to want from freedom of the will can be had in a deterministic world. In the next part of the chapter I shall argue, on the other hand, that determinism, specifically microphysical

[6] A classic statement of such a position is that of William James (1884/1956).

determinism, or, in fact, merely microphysical causal completeness, really is a problem for an adequate account of human autonomy. But, as I argue in the succeeding section, we have, fortunately, no reason to believe in determinism—or even causal completeness, whether microphysical or of any other kind. The chapter will conclude with some further discussion of how I conceive the rejection of causal completeness to provide a way out of the traditional problem of free will.

3. Microphysical Determinism and the Causal Inefficacy of Everything Else

Suppose that there is some set of microscopic entities undecomposable into any smaller constituents, and of which all larger entities are composed. Assume that all putative entities that might appear not to be composed of anything (numbers, abstract objects, universals, etc.) are either wholly dependent for their existence and behaviour on objects made of these microscopic entities, or nonexistent. Though these suppositions are hardly uncontroversial, I believe that they would be widely accepted among the many philosophers who think of themselves as physicalists. Now suppose that we also have a fully deterministic account of the behaviour of these microscopic entities. Although heroic attempts have sometimes been made to deny it, it seems to follow inevitably from this set of assumptions that the behaviour of everything is fully determined by the laws at the microlevel. This seems to follow immediately from the assumption that objects at higher levels are composed entirely and exhaustively of the microscopic objects. For, given the assumption of determinism, it is true of every individual microscopic object that its behaviour is fully determined by the laws governing microscopic objects. And surely if the behaviour of every constituent of a thing is determined, so is the behaviour of that thing.

This point can be made more graphic by thinking of a constituent of a human being, say an electron in my finger. I might be inclined to explain the movement of that electron by saying, for example, that I was waving to a friend, and my waved hand just brought the electron along with it. But clearly this explanation is going to have to be consistent, at the very least, with the explanation in terms of the microphysical laws acting on the electron. If we now consider the

same condition applying to all the various electrons and suchlike in my arm, it would appear that only cosmic coincidence or some kind of dependence of the higher level on the processes at the lower level could ensure this overall compatibility. The bold conclude at this point that either the higher-level phenomena are reducible, in the sense of derivable, from the lower-level phenomena, or that the former cannot really exist at all (eliminativism). The more cautious fall back on claims of supervenience. They claim, that is, that even if the connections between the lower and higher levels are far too complex for us to discern any systematic relationships, the latter do, nonetheless, depend entirely, or *supervene*, on the former. As far as I can tell, however, this is merely reductionism with a modest reticence about the capacity of humans to carry it out. At any rate, none of these positions allows any genuine autonomy to the higher structural level.[7]

One might wonder whether, even given that the laws at the microlevel fully determine the physical trajectory of my arm, they might nevertheless fail to determine that that movement constitutes waving to a friend. So the microphysical can explain why my arm waved about, but fails to say what I was doing. But if this is right, it can only mean that we need to push the microphysical explanation further back. Facts such as that my friend was leaving and that I thought it polite to wave to him and wanted to be polite, etc., etc., cannot, on the picture under consideration, deflect the electrons from their predetermined paths. If such facts, the facts that make it true that what I was doing was an act of waving to a friend, have a role to play in the explanation, it can only be through their correspondence with, perhaps even identity to, the underlying microscopic facts. So perhaps to capture what we might otherwise have supposed to be the correct, mentalistic, explanation we need to bring in more microphysical facts than we might if we were only looking for microphysical antecedent conditions that cause my arm to move. But nothing has suggested, or could suggest, that the resources required could lie outside the realm of microphysics.

[7] This is at the basis of Kim's well-known arguments against non-reductive physicalism (Kim, 1993, especially essays 14 and 17). Kim shows that such a position requires 'downward causation', the causal influence of macroscopic on microscopic entities. I accept the argument but, as will be clear below and as I have explained elsewhere (Dupré, 1993a), I see no problem with downward causation.

It is important to stress the concept of causal completeness rather than merely determinism here, since nothing is significantly altered in the preceding argument by moving from a deterministic to an indeterministic but complete system of laws at the microlevel. Given my intention to drink from the glass of water in front of me, the probability that the electron referred to two paragraphs ago will move in a certain direction is very high. Again there must be some parallel explanation at the microlevel that also attributes a similar high probability to such a move. And again, when we aggregate all the particles that compose my arm, some explanation is required of the apparently extraordinary coincidence between the phenomena at the two levels.

The compatibilist will not, perhaps, be particularly disturbed by any of this. My action, according to the compatibilist, is caused by my beliefs and desires or whatever internal states, and these internal states are also physical states of my brain. She will invite us to recall that autonomy *requires* that our states of mind should be causally efficacious, and to agree that to be so they must be part of the causal nexus of the physical. I have just two comments to make to such a compatibilist. First, if the arguments that have driven so many physicalists away from robust reductionism towards doctrines of supervenience, anomaly,[8] or even eliminativism are correct, the supposedly causally efficacious mental states appear to be in a precarious state. If eliminativism were true, they would be in the worse than precarious state of nonexistence. But even on other such putatively non-reductive physicalist doctrines, there will be no determinable principles on the basis of which a physical causal process will give rise to the causal processes at the mental level, and the co-occurrence of processes at these levels will be something of a mystery.

But the second and more important point is that even if it is clear how a physical process of the kind occurring must at the same time give rise to mental processes or events of the appropriate kind, it is impossible to escape the charge that these processes are redundant. All the physical movements of the agent would have happened even if the mental occurrences had not, if, that is to say, there had been no principles or laws requiring mental processes or events to come along for the ride with the physical ones. The mental is at best

[8] I refer, of course, to the anomalous monism of Davidson (1970).

epiphenomenal,[9] which is to say lacking in any autonomous causal power. In summary, then, causal completeness at the microlevel must entail reductionism, at the very least in the sense of the supervenience of everything else on the microphysical. And even supervenience, I claim, is sufficient to deny any real causal autonomy to higher structural levels.

The alternative picture I would like to advocate denies causal completeness at any level. Objects at many, probably all, levels of the structural hierarchy have causal powers. One of the reasons why these causal powers are never displayed in universal laws (deterministic or probabilistic) is that objects at other levels often interfere with the characteristic exercise of these powers.[10] I take it that the example of the electron in my hand is best seen as such a case, a case, that is, of interference with the microphysical level by the macrophysical.[11] If that is right, then the behaviour of microlevel objects is very frequently consequential on processes at higher structural levels. As a simple example in the opposite direction, a person's plans can be seriously impeded by a dose of radiation. Just as I advocate ontological pluralism at particular structural levels against the essentialism that tries to insist on a uniquely privileged position for one set of kinds, I want to insist that the same ontological tolerance should be accorded between structural levels. As objects are united into integrated wholes they acquire new causal properties (perhaps that is exactly what it is for a whole to be—more or less—integrated). I see no reason why these higher-level wholes should not have causal properties just as real as those of the lower-level wholes

[9] The recognition of this consequence of the completeness of physics has led to some extraordinary discussion of consciousness, the thing that seems most clearly to be shown thereby to be epiphenomenal (see Chalmers, 1996).

[10] As of course may objects at the same level. This kind of objection to universal regularities has been emphasized by Cartwright (1983), who has developed the idea in terms of the inevitability of open-ended *ceteris paribus* conditions.

[11] It will be objected—and may already have been objected—that the electron will be pushed by the microscopic object or objects immediately behind it and will push those in front of it, and thus all the particles are moving in response to microlevel forces. I do not mean to deny this: certainly it would be absurd to suppose that my intention independently acted on each particle in my arm. The real issue is whether all these arm-particles are moving as part of a much wider set of microphysical events (photons bouncing of the glass, hitting my retina, stimulating my brain, etc.) of which my intention to drink the water is ultimately a mere epiphenomenon, or whether, rather, the fundamental explanation for all those particles pushing one another in a certain direction is that I am thirsty and see a glass of water I plan to drink. Evidently I prefer the latter view.

out of which they are constructed. The reasons that many philosophers have seen for denying such higher-level reality are all grounded in the conceptual nexus that links determinism (or at least causal completeness) and reductionism. To the former, which is both still widely believed and traditionally linked to concerns about human freedom, I shall now turn.[12]

4. Causal Incompleteness

The thesis I shall now defend is that there is no plausible ground for the belief in causal completeness. I shall address most of the argument to the doctrine of determinism, but I intend that everything I say will apply equally to indeterministic versions of causal completeness unless I explicitly differentiate the two cases. The basic strategy of my argument will be as follows. Presumably determinism is a very strong metaphysical assumption. To claim that everything that happened had to happen, given the totality of prior conditions, is to impose an enormously strong—indeed the strongest possible—restriction on the possible evolution of the universe. And even the claim that the state of the universe at any time fully determines a set of objective probabilities for its subsequent state is a strong assumption. My point is then that such strong assumptions require persuasive reasons if they are to have any plausibility. I do not take seriously the idea that determinism might be established by means of a transcendental argument of some kind, simply because, as I shall explain below, an indeterministic, causally incomplete world seems to me entirely possible. Thus the first step in the argument is to insist that the onus of proof belongs with the determinist. Determinism has been so closely linked with the philosophy underlying the rise of modern science that it has come to seem obvious, something to deny which is to call in question the whole scientific project. But in my view determinism is a philosophical free rider on the scientific world view, something for which the latter provides no warrant, and something that is, despite the success of the scientific world view, inherently implausible. Assuming that what is most fundamental to the scientific world view is empiricism—very broadly, answerability of

[12] The various theses summarized in this paragraph are defended in detail in Dupré (1993a).

our beliefs to what we actually experience in the world—my question will be whether there is any basis in our empirical interaction with the world for supposing that it is causally complete. My answer will be in the negative.

There are two main kinds of experience that might be held to legitimate a belief in determinism. These are, first, our familiarity with scientific laws and, second, our everyday causal experience. An important special case of the latter is our experience of highly organized systems, especially machines and organisms. I shall deal with these topics in turn, but reserving the special cases of machines and organisms to a separate section. First, then, do the results of scientific investigation lend support to the idea that the world is deterministic, or at any rate causally complete? Here I must first dispose of a troublesome red herring. It is often claimed that science must assume determinism as a methodological imperative. The idea is that it would be sheer defeatism, when confronted with a phenomenon anomalous in the light of current belief, to assume that this was simply a phenomenon outside the causal nexus. We naturally and correctly attempt to broaden our understanding of the range of phenomena in question so as to remove the appearance of anomaly. But that, it is claimed, is to assume that the anomalous phenomenon is in fact part of a uniform and complete causal nexus. Thus it might be suggested (and this is the non sequitur) that science must assume determinism; and then, perhaps, that the successes of science provide evidence that the presuppositions of science, in particular determinism, must be true. But of course to say that science aims to explain phenomena does not entail that all phenomena can be fully explained. And to say that science has had some explanatory successes hardly implies that everything that happens can be fully explained as part of an underlying universal regularity. We can be optimistic about life without inferring that this must be the best of all possible worlds.

So do the actual results of scientific research provide more direct evidence for determinism? The most compelling such results, for the reasons spelled out in the preceding section, would be those that provided evidence for causal completeness at the microlevel. But clearly there is no such evidence. Although certain very specialized phenomena in extremely carefully controlled conditions do exhibit some impressive regularities, this is the entire extent of such evidence. (As should become apparent later on, the fact that these

regularities are produced in extremely elaborate *machines*—machines painstakingly designed for the very purpose of producing these regularities—is of great significance.) Evidence for causal completeness would require that increasingly complex systems of physical particles could be shown to be amenable to causal explanation in terms of the laws said to govern individual particles, evidence, that is to say, for general reductionism. I cannot here go into the general difficulties that confront the project of reductionism. But I do not need to do so. No one has claimed to be able to explain the behaviour even of very small collections of particles in terms of the behaviour of individual particles; the reduction even of relatively simple parts of chemistry to physics is now looked on with considerable scepticism (Scerri, 1991; 1994); and even physics itself is acknowledged to consist of laws the relations between which are obscure, though at least the unification of physics is still looked upon by some physicists as an attainable goal. At any rate, the view that every physical particle has its behaviour fully determined by microphysical laws must derive any plausibility it has from some source other than the development of microphysics.[13]

It appears then that microphysical determinism must be motivated, somewhat paradoxically in view of the connections between determinism and reductionism, by experience at the macroscopic level. But before turning to our everyday experience of causal regularities we might consider the possibility that microphysical determinism could be motivated by our knowledge of macrophysical laws. The obvious candidates, since they remain the most widely admired paradigm of scientific knowledge, would be the laws of Newtonian mechanics. But here we encounter exactly the same difficulty that we saw at the microlevel. Whereas scientists have been able to subsume very simple systems such as the solar system under impressively reliable regularities, the ability to apply Newtonian laws to more complex systems has proved severely limited. The notorious failure to solve the three-body problem, let alone *n*-body problem,

[13] It is of course true that microphysical laws *purport* to apply to indefinitely complex systems, in the sense that they determine how the formalism should, in principle, be applied to such systems. But in practice they certainly cannot be so applied. And one need hardly be a radical sceptic about induction to resist extrapolation from a very narrow and limited set of data to every phenomenon whatever that could in principle be subsumed under the purported regularity. For a detailed and subtle account of how results really are generated in physics experiments, see Galison (1987).

marks this failure. Thus we have no empirical evidence for the general truth of Newtonian mechanics as applied to complex systems of bodies unless we are prepared to countenance inductions grounded on one kind of case (very simple systems) to all cases, most of which are very different from those empirically studied. Moreover, to reiterate a point emphasized by Cartwright (1983), we know that laws such as those of Newtonian mechanics are true only under a very stringent *ceteris paribus* condition, a condition we know to be generally false. Thus, far from knowing that these laws are universally true, we know that they are generally false. The assumption that the laws of Newtonian mechanics are, in some sense, carrying on regardless under the overlay of increasingly many interfering and counteracting forces is not merely sheer speculation, but actually of dubious intelligibility. What are these laws supposed to be doing, given that the objects, subject to such diverse other influences, are not behaving in any sense in accord with them? Certainly this can hardly be a good *empirical* ground for the alleged universality of microphysical laws.[14]

The other common idea, mentioned above, is that determinism is evident from our everyday experience of causality. This assumption can be seen in classical regularity theories of causality from David Hume to J. S. Mill and J. L. Mackie.[15] Hume appeared to take determinism outside the human sphere to be so obvious as not to need much discussion. He was more concerned to show, with well-known examples such as the sure and swift appropriation of a purse of gold abandoned at Charing Cross, that humans were subject to regularities just as immutable as those governing the natural world. Mill was a good deal more sensitive to the complexities of regularities of the latter kind, realizing that the regularities of common experience could easily enough be defeated by either the absence of necessary background or auxiliary conditions, or by the presence of interfering conditions. Thus a lighted match thrown onto a pile of dry straw will always start a fire—unless, that is, there is no oxygen, or a fire extinguisher is simultaneously directed at the straw, etc. While thus acknowledging the complexity of everyday causal regularities, Mill appears to have thought that with sufficient care to include all the relevant auxiliary conditions and exclude all possible

[14] See Suppes (1993) for a more detailed argument complementary to the present one.
[15] See principally Hume (1748), Mill (1875), and Mackie (1974).

blocking conditions, a truly universal regularity could be discovered. This idea reached its most sophisticated expression with Mackie's analysis of an everyday cause as an insufficient but non-redundant part of an unnecessary but sufficient condition, or an 'inus' condition. The sufficient condition in this analysis is the cause with all the auxiliary conditions and the negation of possible interfering conditions. The non-necessity of such conditions points to Mackie's additional recognition that there might be many such complex sufficient conditions of which none, therefore, would be necessary (a bolt of lightning might equally well have ignited the pile of straw).

Many objections can be raised against this picture, at least if it is assumed that it intends one to take seriously the universality of the implied laws rather than merely to illuminate the relations between miscellaneous items of causal lore. One may well doubt, to begin with, whether there is any definite limit beyond human imagination to the number of conditions that we might need to add to produce a fully universal generalization. More seriously, the more conditions are added, the further these putative regularities recede from any possibility of empirical support or refutation. Indeed the reason we are forced to move from simple regularities (e.g. lighted matches cause fires in flammable materials) to increasingly complex and qualified regularities is simply because we recognize the general falsity of the simpler ones. But as we move to such ever more complex regularities, first, the amount of evidence even bearing on the truth of the regularity will rapidly decline; and second, in keeping with the process that brought us the complex regularity in the first place, were we to find an exception to the complex regularity we would presumably respond by looking for a further interfering condition rather than by rejection of the entire regularity. This suggests that the Mill/Mackie programme might better be seen as embodying a methodological rather than a metaphysical conception of determinism.

A second kind of objection casts doubt on the empirical basis of everyday causal determinism from a rather different perspective. Many everyday phenomena give no superficial appearance of being deterministic or even nearly deterministic. Consider, for example, a tossed coin. Now it is often asserted that a coin spinning through the air is a fundamentally deterministic phenomenon, and the only reason we are unable to predict the outcome is that we have an insufficiently precise knowledge of the initial conditions. It is much less clear why this is asserted. Presumably it must be because the

kinds of laws involved in such a process (mainly Newtonian) are assumed to be deterministic. But I have already considered the weakness of that line of thought. The present case, since it is one in which we cannot in fact make any such predictions, provides further support for the argument against basing determinism on macroscopic scientific laws. At any rate, the thesis that everyday causal experience, suitably refined in the style of Mill and Mackie, provides grounds for the belief in determinism, simply ignores the fact that a great deal of our experience, whether of gambling devices such as tossed coins and roulette wheels, or just of seemingly quite erratic natural phenomena such as falling leaves or swirling smoke, provides no such grounds.

The final argument I shall mention is perhaps the most telling. It is that if there is causal indeterminism anywhere, it will surely be (almost) everywhere. Suppose, as is sometimes rather bizarrely suggested, that the only locus of indeterminism is in quantum mechanics. But surely—and here phenomena such as hypothetical quantum amplifiers in the brain have genuine significance—it must be impossible to insulate the indeterminacy of quantum events so fully from consequences at the macroscopic level. Consider again, for instance, the tossed coin, and suppose that its trajectory deterministically produces—*ceteris paribus*—its final outcome. Suppose the coin is at a point at which it is about to land heads. And suppose finally that a collision with a fast-moving air molecule is sufficient to reverse this outcome and produce a toss of tails. If the situation is sufficiently delicately balanced this must surely be possible. Assuming that the molecular trajectory is a sufficiently microscopic event to be subject to some degree of quantum indeterminacy, then we can easily see that the claim to determinacy of the coin-tossing event cannot be sustained. We cannot treat this as merely another interfering factor, because whether or not it has any effect on the final outcome cannot be determined by any amount of knowledge of the initial conditions.

It is a further advantage of this example that a coin toss is the kind of event that might imaginably have massively ramifying consequences. Perhaps the last degenerate scion of some aristocratic line is wagering his fortune on this coin toss. The outcome will dramatically affect the lives of his dependants, servants, creditors, etc. and their fortunes will have an increasing cascade of consequences. The general point that this example is intended to illustrate is that in-

determinism anywhere, by virtue of the variety of causal chains that might be initiated by an indeterministic event, is liable to infect putatively deterministic phenomena anywhere. It is significant that this applies equally within and across levels of structural complexity.

One final point will conclude this part of the discussion. The last argument presented is an argument against determinism, but not necessarily against causal completeness. In the case of the coin toss, provided only there is no correlation between interfering molecular events and outcomes, we should expect that these would be equally likely to change heads to tails and vice versa. So even if these interfering events occurred in accordance with no law even of a statistical nature, they might not render incomplete the supposed law that coins of a certain kind come up heads 50 per cent of the time. On the other hand the preceding arguments, based ultimately on the lack of empirical support for determinism, seem if anything even more pressing against an indeterministic version of causal completeness. For any investigation of a range of phenomena will provide statistical facts. That, for some x, x per cent of events of type A are followed by an event of type B, is a matter of logic. But for this very reason, even if we have excellent grounds for believing that As really do have a tendency to produce Bs, it is difficult to see why we should be led to believe that there is any x such that it is a *law* that x per cent of As produce (or are followed by) Bs. The most plausible basis for such a belief, I suppose, would be microphysical reductionism, a topic about which I shall say no more here. We might better ask, 'What would it *mean* for there to be a law of this kind, as opposed to there merely being a tendency of As to produce Bs, and a statistical correlation of a certain strength between As and subsequent Bs?' Ignoring for the present purposes a range of widely explored subtleties concerning spurious and genuine correlations, joint effects of a common cause, and so on, which would be required for a detailed answer to this question, the simple answer which is sufficient for my present purposes is just that a precise causal law should license us to expect that the proportions measured in a suitably large number of trials should be (approximately) repeated in the future. It seems to me, on the contrary, that in practice such an expectation would often be foolhardy.

In real life, the level of confidence with which we treat statistical experience as a guide to future expectations will vary from almost total to almost none. No doubt many explanations could be given of

the reasonableness of such a perspective, some consistent with causal completeness. The explanation that seems to me most consistent with both investigative practice and the experience of causal regularity, however, has nothing to do with laws or statistical uniformities at all. Correlations reveal, I believe (subject to well-known qualifications), the causal powers of certain objects or events to produce particular effects. Whether we expect the production of such effects to occur with a fairly constant frequency depends whether we think that the frequency of other relevant causal factors is likely to remain reasonably stable. But without some apparently quite arbitrary way of privileging a particular constellation of background conditions, there is no such thing as the quantitatively precise, constant, and timeless tendency of As to produce Bs *ceteris paribus*. Other things can be a particular way, and they can be more or less reliably that way. But except in the very simplest cases, as in Newtonian mechanics where we imagine that there are only two bodies in the universe, and everything else is supposed not equal but absent, I do not know what 'everything else being equal' even means. Thus once we have fully appreciated the complexity of the causal nexus, the thesis of indeterministic causal completeness is seen not only to be devoid of empirical support, but even to be, once again, of dubious intelligibility.

5. Machines and Organisms

As I have tried to show in the preceding section, I do not think that direct reflection on our (extremely limited) knowledge of universal regularities lends much support to the idea of a universe with a complete causal structure. However, it may well be that deterministic intuitions derive more from reflection on complex and highly organized structures, especially machines and biological organisms. Since the overall metaphysical vision out of which the whole problem of free will arose is aptly referred to as mechanism, it is certainly appropriate to consider the artefacts that have somehow come to provide a model for the universe; the consideration of organisms, notoriously liable to be treated as a kind of naturally occurring machine, will bring us back to a topic fundamental to the central theme of this book, the causal status of humans.

It is easy enough to see why machines should have some tendency

to inspire deterministic intuitions. Machines, good ones anyhow, are extremely predictable. I am confident that the text I type into my computer is exactly what will eventually come out of my printer when I connect them up in the right way. (Though not so confident that I do not occasionally make a hard copy; and some people, I am told, even make back-ups of their computer files on disks.) But a little further reflection makes it very puzzling that something like this, rightly admired as one of the great triumphs of modern technology, should be taken as a model for the universe in general. If the sort of regularity that is characteristic of a good computer or car were typical of the universe it would, one might imagine, be fairly easy to make, or perhaps even just find, such things. But it is not at all easy, which is why such technological achievements are admired. If the universe is a machine, it is far from obviously so.

Perhaps a more sympathetic interpretation of the tendency for machines to inspire determinism is the idea that only if determinism were true would it be possible to make reliable machines. And since we can make reliable machines, determinism is proven to be true. Underlying what seems to me a great exaggeration in the first premise of this argument there is, nevertheless, a very interesting question: what degree of order must exist in the world for the kinds of reliable machines we possess to be possible? The beginning of a more temperate answer to this question than the immediate appeal to determinism is the observation that no machines are completely reliable, and some are very unreliable. The point of this observation is not to insist—though strictly speaking it is no doubt true—that there is some possibility, however remote, that when I type the word 'type' on my computer a four-letter obscenity will instead appear on the screen; or that when the spark ignites in the combustion chamber of my car the gasoline inside it will spontaneously liquefy. Rather I want to focus on the question, what it is that makes machines more or less reliable. And of course the answer is not, at any rate, that reliable machines have access to more universal laws.

Consider, then, what is by modern standards a fairly simple machine, an internal combustion engine. If we ask how such a machine operates we may be content with a very simple story: a mixture of air and petrol is exploded in a cylinder, pushing a piston down the cylinder; the cylinder is connected to a shaft which is rotated by the moving piston. A number of similar cylinders are connected to this shaft, and a sequence of explosions keeps the shaft rotating

continuously. It seems to me that this is, roughly speaking, a correct answer to the question how an internal combustion engine works. But if, on the basis of this explanation, someone lined up some coffee cans partially filled with petrol on the kitchen floor, stuck toilet plungers in the cans, tied the ends of the plungers to a broomstick, and then posted lighted matches through little holes in the sides of the coffee cans, they would certainly not have built an internal combustion engine (though I suppose the broomstick might jump about a bit).

I suggest that it is useful to think of how a machine works in two stages. First there is the question what makes it even possible for the machine to do what it is supposed to do. A slightly more elaborate version of the answer sketched in the previous paragraph might be an answer to this question for an internal combustion engine. Having got that far, however, most of the details of the internal combustion engine concern the more or less ingenious auxiliary devices that make sure it really does do what it is intended to do rather than one of the many other things it has an initial capacity to do. So, for instance, the cylinder must be strong enough to avoid simply disintegrating when the petrol explodes; the crankshaft must be extremely strong and rigid if it is to reliably convert the linear momentum of the pistons to rotational motion; piston-rings prevent the energy of the explosion from being dissipated between the piston and the cylinder; oil must be provided to prevent the cylinders getting so hot as to seize in the cylinder, or for that matter melt; some way must be found to dissipate excess heat from the running engine; and so on. Even a Trabant has the capacity to run and sometimes does so. The difference between this and a well-designed car is that the behaviour of the parts of the latter is so tightly constrained that it can do nothing but what it is designed to do—though eventually, of course, even the best-designed machine will break free of its constraints. My point so far is just that this kind of constraint is not something characteristic of nature generally, but something that engineers devote enormous efforts to attempting, never with total success, to achieve.

Of course, this account of the reliability of machines does assume the reliability of various causal relations: gasoline and air mixtures almost invariably explode when sparked; heat will flow from a hot engine to cooling water circulating over it; and many others. It is interesting that many such regularities can be seen as reflecting the

overall upshot of very large numbers of similar though indeterministic processes at the microlevel, which suggests the hypothesis that it is just those macrolevel processes that can be roughly reduced in this way that reveal this near determinism. But I do not want to insist on this here. While machines could presumably not work without exploiting extremely reliable regularities such as those just mentioned, the regularities that characterize the machines themselves, as with many other macroscopic causal regularities are only more or less reliable. And in keeping with a general philosophical theme I endorse, it is best to think of these regularities as involving the reliable exercise of the capacities of things when properly triggered and unimpeded. The capacity of suitable mixtures of petrol and air to explode when ignited is an extremely reliable one and very difficult to impede. Such reliably exercised capacities are no doubt a precondition of the possibility of building reliable machines. But the existence of such capacities provides no basis at all for the conclusion that everything that happens is the exercise of a similarly reliable capacity. Indeed a great deal of experience—experience of the generally more or less unreliable and unpredictable natures of things—speaks against it.

Reflection on how good machines are engineered, far from making us think of mechanism as generally characteristic of the world, should make us realize how difficult it is to turn even little bits of the world into bits of mechanism. Though I have admitted that machines could not be made to work if there were not things in the world with capacities that, under certain circumstances, are exercised with (almost) complete regularity, it is important to note that these are quite different from the much more complex and much more tenuous capacities of machines. A fortiori, the things in the world are not limited to simple, reliable capacities of the first sort; and the things that happen in the world are not always, or even generally, the simple exercise of such reliable capacities.

Turning now to organisms, it is a familiar idea, especially following Descartes, that organisms just are machines. Natural theology until the late nineteenth century considered organisms quite explicitly as the products of a divine mechanic.[16] Mechanistic modes of investigation have had extraordinary successes in uncovering how metabolism, reproduction, and other basic biological processes

[16] The *locus classicus* is Paley (1802).

work. And even in the domain of behaviour, the complex but highly stereotyped performances of many insects in, for example, constructing and provisioning burrows for egg-laying have many of the characteristics of a well-designed machine. To the extent that the analogy is appropriate, the same remarks that I made about machines will apply to the relevance of organisms to the prevalence of causal regularity. Looking, however, at the other end of the organic scale, and most especially at humans, the parallel with machines has serious limitations.[17]

The fact that when, for example, I intend to walk down the garden path, my legs move in just the right way to maintain my balance and propel me forward is, I suppose, something that could be explained in a manner strongly analogous to the performance of a machine, though perhaps one more complex than any machine we have yet managed to construct. I suppose that the physiology and cell chemistry of muscle tissue explains how the physical movements are obtained, and a variety of sensory and neural mechanisms bring it about that the motion is steady and in the right direction, and that a vertical posture is maintained. Although this seems significantly analogous to the account I offered of the working of the internal combustion engine, we should now note that an internal combustion engine is in reality not a machine but a part of a machine. If we think now not just of an engine but of an entire car, an important class of features has yet to be mentioned. I am thinking of such things as the ignition key, the steering wheel, and the brake pedal, those devices by which the machine is made to act in a way conducive to the ends of its human operator. A reliable car, as opposed to a reliable engine, the latter of course being a necessary but insufficient component of the former, is one in which there is a reliable correlation between inputs to these controls and the behaviour of the whole machine. Thus machines are not sources of causal autonomy; they are, at most, instruments for furthering the causal autonomy of their users. The superficial, and I think also deep, disanalogy between humans and machines is that humans have no controls.

[17] I focus here only on what I take to be the extremes of the animal scale. I assume that higher mammals, birds, and perhaps higher molluscs, are more like humans than they are like the most machine-like of insects. But I shall make no attempt here to draw any more specific lines between different kinds of organisms, though this may be an important task for those concerned with our ethical responsibilities towards animals.

It may rightly be objected at this point that insects with simple stereotyped behaviours have no controls either, yet I have claimed that they are closely analogous to machines. There are two possible responses. First, a stereotypical performance might simply be produced in response to nothing at all. More typically and interestingly, a kind of behaviour might be triggered by some sensory input, the sense organs thus serving as devices for producing behaviour appropriate to the external circumstances. This is primarily what I have in mind in talking of the stereotypic and machine-like behaviour of certain insects: a certain stimulus triggers a sequence of behaviour. One might reasonably suggest that the sense organs in such an organism serve the functions of controls. There is, of course, a tradition of psychological investigation of humans that applies just this model to humans. Though in its crudest behaviourist versions it has been almost wholly rejected, the idea that sensory inputs, mediated by 'information-processing mechanisms', somehow elicit the appropriate 'emission' of behaviour is still widely, perhaps generally, assumed. This is a mechanistic model, though one in which the complexity of the machine is such that we as yet have no idea what it is designed to do in the innumerable situations it encounters.[18] Against this model, and as I have argued in more detail in earlier chapters, I propose that we should recognize that we were not designed at all, and consequently there is nothing we were designed to do in any situation.

Between two views that I have rejected, that we are random action generators and that we are machines, can be found the view that makes sense of human autonomy. Many parts of humans have just the characteristics of machines that I have emphasized in the preceding discussion, namely complex constraints that ensure the predictable exercise of some capacity of an organ or physiological system. But humans are fundamentally different from machines in that they have no controls. Self-control, in the sense of the absence of external controls, is of course nothing but the autonomy, or free will, that it was the goal of this chapter to illuminate. I have not attempted to refute the idea that sense organs might sometimes function as

[18] As should be clear from earlier chapters, evolutionary psychologists and their fellow-travellers do think we are designed, and do think they know exactly what we are designed to do, to survive and reproduce. I have said enough about what I see as the inadequacies of this unwitting attempt to appropriate the creationist tradition.

controls, in the sense that the input to sense organs might determine, via a complex intermediate causal chain, the behaviour of the whole organism. This is presumably roughly true of simple organisms. But it does not appear to be true of ourselves, except perhaps in purely reflexive actions, such as ducking to avoid a flying object. The reason we are so liable to think of ourselves in this machine-like way is because we are tempted by determinism. If the world is deterministic then my behaviour is causally necessary given the stimuli that impinge on me; and presumably the most important stimuli are sensory ones. The point of all the complex machine-like parts of me would then have to be just to make sure that the causally elicited behaviour was appropriate to the circumstances disclosed by my sense organs. And this is the concordance that, according to evolutionary psychology, millions of years of evolution have succeeded in bringing about. But the rejection of causal completeness allows a more natural view of things. My complexity of structure gives me a vast array of causal powers, a range of powers that would be inconceivable without that intricate machine-like internal structure. But the exercise of those powers, though obviously influenced by the circumstances I perceive myself to be in, ultimately depends on an autonomous decision-making process.[19] Once we see causal order as something special rather than something universal, there is no obstacle to seeing the human will as an autonomous source of such order.

One very important point should be added, which should ultimately profoundly modify the preceding point. For much of human behaviour, context is far more than a trigger that prompts the emission of behaviour. Human behaviour, recalling a philosophical truism, consists not merely of movements, but of actions, and social context plays a central part in determining what action is constituted by a particular movement. The action of signing a cheque could be no more than a meaningless wiggling of my hand without an extremely elaborate social context. Most important of all, what would otherwise be merely noise becomes articulate language only

[19] Philip Kitcher, commenting on this passage, raised the pertinent question whether the 'I' introduced in this discussion, if not fully comprehensible in terms of neurons, receptors, and so on, did not recapitulate old-fashioned dualism. For now I want only to insist that the 'I' refers to the whole organism, not just some neurologically salient bits of it. I shall try to make clearer below how I understand the 'autonomous decision-making process' that it undergoes.

because there is a society in which noises have meanings.[20] I have said earlier that the usefulness of reductive analysis is that whereas knowledge of the inner workings of things is typically quite insufficient to tell us what they will do, it is generally the way of explaining why they have the capacities to do the things they do. In the case of humans this is only true if we limit consideration to capacities described in the broadest way. The structure of our brains is no doubt such as to enable us to learn a language. But what we can say or do with that capacity is entirely determined by the social context that we have the good or bad fortune to find ourselves in. Although it has been the main purpose of this chapter to defend the idea that individual humans are potential sources of causal efficacy, it is with regard to capacities the possibility of which derives from the relation of an individual to a social context that this potential causal efficacy takes the most significant forms.[21] And as I argued in Chapter 2, distinctively human capacities are almost all dependent on a social context, or at least a relation between an individual and a social context. In the next section I shall try to make clearer how such capacities, the distinctively human capacities that are most thoroughly concealed by the scientistic and doggedly individualistic perspectives that have been criticized in this book, can help to make sense of human freedom.

6. Moral Autonomy

Pleasant though it would be to do so, I do not expect to resolve all the problems that have troubled philosophers over the ages concerning the human will. What I have, more modestly, been trying to show in this chapter is that, contrary to a notorious tradition of philosophical controversy, a reasonable metaphysics of causality presents no special difficulties for the idea of human autonomy, and requires neither ghostly nor random nudges of the physical causal order. In order to give more positive philosophical substance to the view, I shall begin this discussion by attempting to locate my views more perspicuously

[20] As I have mentioned earlier, the elaboration of this point is one of the most famous contributions of Wittgenstein's masterpiece, *Philosophical Investigations* (1953).

[21] Searle (1995, esp. ch. 6) provides an excellent though rather different account of the way social context shapes human capacities, though I doubt whether he would agree with the way I develop this thought.

within some aspects of the traditional debate, and in particular relate what I have said to some very famous views on the subject, those of Hume and of Kant.

Although it may well remain the dominant view of the subject, Hume's attempt to reconcile human autonomy with a classically deterministic structure of causal relations seems to me unconvincing.[22] On the other hand, Hume was surely right that exercises of human freedom were much better understood as instances of causality than of its complete absence. The intuition developed by Hume and subsequent compatibilists that a connection, probably causal, between the psychological states of the agent and the action is, far from a bar to autonomy, a necessary condition of autonomy, seems basically correct. But given the assumption that the agent is located in an otherwise seamless causal nexus, this insight closes off the only, acausal, escape from this nexus, and inevitably leaves the psychological causes of action at best supervening ineffectually above this self-sufficient nexus. His problem, therefore, was his commitment to a universalistic regularity theory of causality. Given such an account of causality, any departure from determinism is a failure of causality itself.[23] But the solution I have advocated was certainly not an option for Hume, since even more renowned than his compatibilist account of free will are his arguments against an ontology of causal powers; and it is just such an ontology that I present as one of the main resources for escaping the causal inefficacy of the human will. Although many philosophers remain convinced by Hume's arguments on this topic, an increasing number do not. This, however, is not a matter that can appropriately be addressed here, and I note only that I and others have attempted to rebut Hume's attack on causal powers at some length elsewhere.[24] Those who

[22] I cannot begin to discuss the enormous literature on this question. Strawson's classic paper (1974) perhaps brings out as clearly as possible some of the most troubling consequences of taking physical determinism fully seriously, though Strawson himself does not appear to think we should be led by such reflections to doubt the truth of physical determinism.

[23] Recently there has been a prominent movement to provide regularity theories of indeterministic causality (e.g. Eells, 1991). I believe there are deep internal problems with such a position (see Dupré and Cartwright, 1988; Dupré, 1993a: ch. 9). But at any rate, for reasons already explained, I do not think the move to such a theory will make any significant difference to the problems currently under consideration.

[24] Some recent advocates of causal powers include Harré and Madden (1975), Cartwright (1989), and Dupré (1993a: ch. 9). Anscombe's (1971) brilliant attack on Hume's account of causation is also highly relevant.

remain convinced that the idea of a causal power or capacity has been shown by Hume to be incoherent will not be convinced by this part of my argument.

Kant also appears to have thought that a deterministic causal structure was compatible with human autonomy, though the metaphysical excesses to which he was led in effecting this reconciliation have convinced almost no one. However, his conception of human autonomy offers some promise of providing a vital and final piece in the picture that I wish to present.[25] My point at the end of the last section was that human decision could be a real source of causal order in the world. However, this claim is liable to seem shallow without some further account of the origins of this order. In particular if one traces human decisions ultimately to contingent human desires, desires which presumably can themselves be traced either to our biological heritage or our upbringing, human autonomy seems at best a focus rather than a source of order. And this is just the conception of human decision-making discussed in the last chapter that has been cultivated for well over a century by economists and more recently by exponents of so-called rational choice theories in a variety of disciplines. It does not matter much at this point whether one says, with the economists, 'tastes are exogenous' (and, moreover, *de gustibus non disputandum*), or with the evolutionary psychologists, that they well up atavistically from our evolutionary past. Either way human autonomy turns out to amount to little more than the more or less effective attempts of a want-satisfying machine to satisfy its wants. Some account of the ultimate springs of behaviour that does more to explain the sense in which goals belong authentically to the agent who pursues them seems needed to give real interest to the account of human autonomy suggested in this chapter.

It seems to me that Kant's account of human action points in the right direction in which to look for this final ingredient of an account of human autonomy. Kant, as is well known, distinguished sharply between action motivated by desire and action motivated by principle. And for Kant only action motivated by duty, by the commitment to conform one's actions to the moral law, counted as truly free. Though this account has struck many as intolerably austere, and others as positively paradoxical in seeing free action as

[25] The ideas I borrow from Kant here are most accessibly presented in *The Groundwork of the Metaphysics of Morals* (1948).

law-governed,[26] it is easy to see how it addresses the concern of the last paragraph, that action merely directed at whatever a person happens to want seems to lack the authentic connection with the person acting that an account of free action would ideally incorporate. It does seem to me, however, that once the restrictions of the deterministic framework have been dispensed with, it is possible to gain this benefit of a very loosely Kantian account without either the extreme moralistic distinction in favour of action motivated by duty, or the metaphysically murky appeals to the noumenal world, which have combined to cast deep suspicion on Kant's conception. My suggestion is that it is quite generally the possibility of acting from a principle broader than the maximization of immediate satisfaction that grounds human autonomy. It is this, after all, that enables human action to produce order in the world, and it is this capacity above all that we aim to instil in our children through education.

In a world where order is a local and incomplete phenomenon, the importance of principle as a source of human action is easily stated: it explains how ideas, the creative acts of the human mind, can change the world. But unlike Kant, I do not make a fundamental distinction here between moral and more mundane principles, though perhaps it will be appropriate to make one of degree. I conceive of the principle 'Follow the architect's blueprints in determining where to build the wall' as as genuine a source of autonomous action as 'Do whatever is necessary to end hunger'. One does not have to be as severe in one's moral demands as Kant to see that doing whatever one feels like at the moment, if it is an intelligible human life at all, is not one that realizes the important human capacity for freedom. And despite this moderation of the Kantian position, the enormous importance of moral principles in this context should not be downplayed. The most fundamental reason why we should care about human autonomy is that it holds out the hope that human

[26] This again raises one of the main strands of debate that I have largely avoided during this foray into the free will problem, the question whether the agent could, when she acted, have done otherwise (see e.g. Frankfurt, 1969). For Kant, it is clear at least that if the agent had acted otherwise, she would not have acted freely. Being, up to a point, sympathetic with what I take to be the basic compatibilist insight, I am not convinced that this should be a fundamental issue. If it is, or should be, it is at any rate not one I have anything to say about here.

One theorist of free will whose views are in many respects highly compatible with my own is Eleonore Stump. Stump also argues against the view that a libertarian must insist on the possibility of the agent having acted differently (Stump, 1996).

action might produce a better world. And what that requires, I suppose, is action grounded in moral principles.[27] This is something I believe we are free to choose; and making this choice, I claim, can make a difference.

It is the curse of this topic that any suggestion of a basis on which autonomy might be grounded will inevitably provoke the question, 'What is the origin of that basis?' And hence what I have just said will surely invite the question, 'Where do principles themselves come from?' And the questioner is likely to perceive the standard dilemma: either principles are indeterministically embraced, reintroducing all the problems with naive indeterminism, or they are caused by the circumstances of the agent's upbringing and so on, thus again reducing the agent to just a complex part of the causal nexus. A better solution, I believe, lies with the point at which I ended the previous section of this chapter. Principles, I take it, are essentially linguistic phenomena. The ability to adopt a principle, and to make it part of one's nature that one aims to act in accordance with that principle, is, I take it, a wholly language-dependent possibility. And language is essentially social. Thus the condition for genuinely free individual action is the embedding of the individual in society. Thus, finally, the causal capacities most characteristically and uniquely human are capacities that derive not solely from the internal structure of humans, or human brains, but that depend essentially on the relationship between an individual and society.

None of this should seem surprising to those who take seriously the fundamental biological fact that *Homo sapiens* is a social animal. We should not be surprised that the kind of freedom we possess derives from our being the kind of creatures we are. It may, however, be an unwelcome suggestion for the tradition that connects human freedom with the profoundly individualistic social philosophy and metaphysics dominant in contemporary English-speaking culture. However, I do not mean to imply that autonomy is wholly a social product of which individuals are merely the passive vehicles. Principles or rules do not, as Wittgenstein also famously argued, determine their own application, and rules may be applied with creativity and imagination. And imaginative application or extension

[27] Accounts of free will that attach particular importance to a moral dimension in action have also been developed more recently. See, for instance, Watson, 1975; Taylor, 1976.

of rules may increase the range of possibilities open to the members of a society for whom the rule or principle is available. Thus the social construction of language, meaning, and possible principles of action makes possible, but does not fully determine, human agency.

To employ a concept unfashionable in contemporary anglophone philosophy, the relation between individual and society is a dialectical one.[28] The situation of individuals in society in which principles and goals of action, and systems of belief, are articulated is what endows those individuals with the capacities to embrace principles and pursue goals beyond the momentary satisfaction of desire. But the exercise of these capacities in action, argument, or the pursuit of individual goals, in turn can affect society. And such action can expand (or perhaps sometimes contract) the range of capacities and options available to individuals. The debates in which almost all of us play some small part about what principles should govern our behaviour, how our children should be educated, and so on, and the actions we take in support of our views on such matters, not only involve the exercise of individual freedom, but ultimately contribute to affecting and expanding the possibilities for human action. It is this dialectical interaction between the individual and society that grounds individual autonomy. And finally, the pluralistic metaphysics I have defended here and elsewhere shows that there is no philosophical difficulty in taking seriously such a relationship of mutual determination between entities at different levels of structural complexity. Once we do take such a possibility seriously, it should be no surprise that some of the most interesting and puzzling aspects of reality should depend on such interactions rather than merely be the properties of disconnected individual things.

7. Conclusion

The central positive theme that this book promotes is pluralism, and more specifically pluralism in our approach to our own kind. Though a good deal of lip-service is currently paid to pluralism, the

[28] I use the term 'dialectical' reluctantly, because I do not want to become involved with the great weight of historical baggage the term has collected. However, I know of no other concept that so accurately captures the relationship of mutual interaction and dependence that I wish to convey.

commitment seldom goes very deep. My own project is to insist that pluralism goes all the way down to the basic metaphysical issues of causality and of what kinds of things there are. This metaphysical perspective makes the kind of narrowly focused scientific projects I have been examining look as philosophically misguided as they have proved empirically unrewarding.

The fundamental error with the programmes that I have criticized in this book is the belief, explicit or implicit, that there is some fundamental perspective that will enable us to understand why people do what they do. It hardly needs insisting upon that it is important that humans evolved and have common ancestors with the other creatures we find around us. And nothing could be more important to us than the organization of society and of the labour of individuals in society in such a way as to provide us with a good deal of what we need and want; no doubt economics has something to tell us about such questions. These are important fragments of the picture that we, the uniquely self-reflective animals, have spent the last few millennia trying to put together. But they are fragments, and trying to make one or even a few such fragments stand for the whole presents us with a deformed image of ourselves. One of the most traditional objections to such one-sided, reductive pictures of ourselves is that they leave no room for human autonomy or freedom. In the present chapter I have tried to show that the philosophical context in which I criticize these reductive views does indeed provide an endorsement of this traditional objection.

I maintain, then, that an adequate view of ourselves, were we to acquire one, would include many parts. It would at least include an account of us as biological organisms with immensely complex functioning parts, and an account of how this functioning gave rise to some of the enormously complex capacities we often exhibit. It would require an account of how our societies function, as general as possible but no more general than the empirical facts permit, and an account of how aspects of social organization contribute to the endowment of human individuals with complex capacities that would be in principle beyond the reach of an isolated member of our species. It will include detailed accounts of some of the most important aspects of our social organization, such as economics, and of the history of our societies and our species. And at a philosophical level it will include an account of the nature and limits of our powers to act autonomously to create real change in the world. Or so I suppose.

But it may well be asked whether this open-ended intellectual shopping list really offers any illumination as to what such a portmanteau account of human life would look like. And it is, I suppose, inevitable that I have no very satisfactory answer to this question. To a considerable extent this book represents an exercise in philosophy as Lockean underlabourer, clearing the ground of rubbish in advance of the construction of a sound edifice of knowledge. I do not consider this task of negligible importance: there is much rubbish to be cleared away. Any bookshop will display a fair selection of sometimes best-selling volumes devoted to simplistic accounts of the essence of human nature, and the authors of these works are constantly to be read or heard peddling their wares in the middle-brow print and broadcast media. One might mention, for instance, Steven Pinker's physically if not intellectually weighty, but very widely noticed, *How the Mind Works.*[29] The view presented therein, that the mind is a computer programmed by natural selection in the Stone Age, is as reductive and simplistic an approach to its topic as anyone is likely seriously to propose, and is as lacking in serious insight into the human condition as such an attempt is likely to prove. Yet this is the work of a respected scientist and is treated with considerable public interest. This points to an intellectual pathology well worth critical attention. But it remains easier to say what does not work than what does.

There is, however, a more positive thesis implicit in my negative arguments. I am suggesting that nothing will serve to provide us with insight into human nature of quite the kind we are currently inclined to imagine. We are tempted, to put it crudely, to think that such insight into human nature could be provided by a text grounded in one or a few comprehensive theories (as the example of Pinker's book illustrates). My most radical opposition to this view is to suggest that we might better think of the cultivation of a skill, the ability to understand, or have insight into, human nature and human life, than the writing of a text. Such a change might reflect the move from the very simple phenomena (relatively, of course: I do not suggest that they were so simple as to be easily understood) which science has so successfully explained, to the vastly more complex phenomena of which human nature is no doubt the most complex of all. The search for simplistic theories such as those of contemporary socio-

[29] I discuss this work in more detail in Dupré, 1999a.

biology reflects above all the failure to see this divide. Complexity in this context is not just a matter of very difficult sums that we do not yet know how to solve, but the concurrence of different kinds of factors, each of which may well be complex in this same sense, that we do not know how to fit together. Moreover, there is no reason to suppose it is even possible to fit them together in the systematic, even algorithmic, way that is sometimes assumed. We may be able to construct sets of very simplified problems that we can solve quite effectively, but it is quite erroneous to infer from this that we have discovered a method that will in principle solve any arbitrary problem we might be interested in. This is where I want to discern the limits of science referred to in the title of this book. Without in any way refusing the extraordinary range of knowledge that science has provided for us, there are subject matters that require a more synoptic and integrative vision than the analytic methods of science allow. And here, perhaps, there is the possibility for philosophy to graduate from underlabourer to Queen of the Sciences.

Though I advertise no key to the understanding of human nature, and even doubt whether any such key exists, there are important consequences following from the recognition of the complexity and multi-dimensionality of the problem. We might, for instance, try to promote a more cautious and sceptical attitude towards claims to offer scientific solutions to social and even medical problems. In the first chapter of this book I mentioned pharmacological responses to such psychological problems as Attention Deficit Disorder. I do not suggest (nor consider myself qualified to suggest) that such responses are always or generally inappropriate. I do worry, though, that such solutions will tend to encourage the assumption that such problems are unitary conditions with unitary, generally physiological, causes. This seems to me unlikely. Such a perspective may even encourage the inference from a statistically positive effect of a treatment, to the generally beneficial effect of that treatment. Recognition of the complexity of human behaviour, by contrast, should lead to the expectation that a pharmacological intervention will have effects that will be positive in some cases and negative in others. Refusal to reduce the patient to a physiological problem will reinforce the necessity of attending to the complex particularity of the individual case.

Medicine, indeed, though constantly provided with additional resources by science, remains an art. It must surely remain an art

because its object, human health, is both complex and normative. Full recognition of this necessity would surely lead to some significant reevaluation of the goals and methods of medical education. And I think that pluralism has profound consequences for education more generally. It is a cliché that we live in an age of exponentially increasing information, and it is often assumed that this necessitates ever-increasing specialization as the task of mastering more than an infinitesimal fraction of this information becomes more and more daunting. But a proper distinction between information and knowledge or even wisdom might lead to quite the opposite conclusion. If a subject matter can only be understood from simultaneous attention to a variety of perspectives, then knowledge of a subject matter will require access to diverse bodies of information. And perhaps part of what amounts to wisdom is the ability to know what kinds of information or knowledge are needed in application to a particular case.

Happily, the explosion of information has coincided with the growth of information technology. And surely much of the benefit of information technology is that we don't need to accumulate vast quantities of information in our brains, but need only learn to gain access to the information we want as it is stored outside our bodies. The combination of these two developments—the realization, first, that the most important problems we face have many aspects and no simple solutions, and second, that the accumulation of large quantities of information inside the human mind has become largely redundant—should make possible a quite radical reconception of education, especially, perhaps, tertiary education. The correct balance between breadth and depth of education is a very difficult matter. Notoriously, science education is thought to require ever-increasing depth and specificity of focus. Recognizing the limits to scientific methodologies should encourage us to shift the balance towards breadth and, perhaps more important, the skills necessary to integrate insights from a variety of perspectives. Medical education again provides a paradigm. As technical information about diagnosis, therapy, and prognosis becomes increasingly accessible from external sources, we may hope that trainee physicians will be able to devote more energy to the very difficult task of learning to appreciate and promote the total well-being of the patient.

The suggestions of the last paragraphs have, of course, been speculative. I indulge these speculations because I do want to em-

phasize that the thesis presented in this book is a radical one. The triumphant achievements of science in the last few centuries have been extraordinary, and it is hardly surprising that they have to some extent distorted our conception of knowledge as a whole. It is time, nonetheless, to take a more balanced look at what we can expect from science, and at what role may remain for very different approaches to the acquisition of knowledge. Part of the work of achieving this reevaluation is the recognition, which it was the aim of my earlier book (1993a) to accomplish, that the idea of a uniform scientific project gradually spreading its light across the full range of our interests, is a myth. The present book aims to reinforce this message by considering in some detail a domain of enquiry, human nature, for which this myth is particularly inappropriate and unfortunate. Much of the importance of this project is the negative one of immunizing ourselves against the worst excesses of scientism. But in these concluding speculative thoughts I want to suggest that there are more exciting intellectual vistas to explore beyond scientism.