1. Introduction

This paper takes up an idea that has interested me for some years: the idea that, contrary to a very widely shared assumption, a radically indeterministic metaphysics does provide a way of understanding human freedom as a real and important feature of the world. However, whereas I used to think of this as a solution to the free will problem within the tradition of radical voluntarism, I now prefer to present it under the rubric of indeterminist compatibilism. In the most crucial respects this position remains true to the voluntarist tradition, but in its current incarnation it aims to capture the powerful intuitions that underlie compatibilist thinking.

The paper will begin with a brief summary of the indeterministic metaphysics to which I subscribe, and indicate some of the grounds for holding such a set of views. I shall then consider in a general way how human autonomy could fit into such a metaphysical framework, paying some attention to the question of the emergence of autonomy in the course of human ontogeny. In the final section of the paper I shall explain the contrast between my earlier voluntarist understanding of the issue and the indeterministic compatibilism I now prefer. This will include an explanation of how, contrary to a natural suspicion, this position is not oxymoronic. This section will show how indeterministic compatibilism captures both the commitment to genuine autonomy that has driven much of the voluntarist tradition, and at the same time the powerful intuitions that have made compatibilism so attractive to so many.

2. Indeterministic Metaphysics

I am a committed indeterminist. In one sense, almost everyone is an indeterminist, of course. We all know that certain physical processes, for ex-
ample radioactive decay, are indeterministic. While these processes happen with extremely precise probabilities, so that one can say with great precision how many atoms in a lump of radioactive material will decay in a particular time period, there appears to be in principle no way of predicting the decay of a particular atom. But without disputing this indeterminism of certain microscopic processes, most philosophers, I think, still believe that determinism reasserts itself at the macroscopic level, just as the behavior of the radioactive lump is deterministic despite problems with individual atoms.

The indeterminism to which I subscribe is much more radical, and I suppose that such a version remains an uncommon or even eccentric opinion. First, I think that indeterminism is rife even at the macroscopic level; indeed it seems to me the normal default state of things. Determinism is, I would argue, a rare achievement, occasionally approximated in the manufactures of highly skilled engineers, or in wonderfully simple systems dominated by single causal principles, such as the solar system. Moreover, I don’t think that this indeterminism is typically because the laws that govern the behavior of things are statistical rather than deterministic; rather, I doubt whether for most things or systems of things there are typically any laws of any kind telling them what to do. Order, in short, I see as an occasional and precious feature of the world, usually hard won by us or by the eons of evolutionary time.

I won’t, in this paper, attempt to provide a systematic defense of this general metaphysical outlook. The reader who finds it incredible could either look up where I have argued for it in more detail (Dupré 1993: part III; Cartwright 1999), continue out of curiosity about what might be done with such a strange view, or give up. My aim in this paper is to make a different point, namely that contrary to a widely held assumption, abandoning determinism does point the way to a solution of sorts to the traditional problem of free will. Perhaps if this argument convinces, it will provide a reason to reconsider the grounds for determinist convictions. I have defended a position of this kind for some time (Dupré 1996), but recently I have been led to reconceive this defense as a version not, as I used to think, of radical voluntarism, but rather as a version of compatibilism. This will be explained in due course. But first I will elaborate the general metaphysical view a bit further, and describe some consequences that are especially important for understanding the argument about free will.

1 A slightly revised version is presented in Dupré 2001, ch. 7.
The only major point I will make here in favor of radical indeterminism is to ask why we should suppose that the onus of proof should lie with the claim that some things happen without being mandated by universal laws. The claim that everything happens in accordance with law is a very strong one, and one might think that the onus should lie with the defender of such a strong, positive thesis. Certainly we must acknowledge as much causal order as is necessary to account for whatever regularities we find, and for the possibility of exploiting these to produce even greater regularity, as is evinced by our better machines and our most successful scientific experiments. But while these force the acknowledgement of some order, the great difficulty of extracting practically useful or theoretically illuminating order that is demonstrated by these enterprises surely speaks against a ready assumption of complete and universal order. Certainly something more does need to be said here about where whatever order there is in the world comes from if, as I have claimed, this is far from universal. To this I next turn.

Implicit in my very brief explanation of why I am not a determinist is something like a methodological principle: assume as much as is necessary to explain the empirical phenomena, in this case what regularity there is, but no more. And the empirical reality, as I see it, is that regularity is a fascinating but none too common phenomenon. I have mentioned the mechanisms that we build, of the set-ups we construct for scientific experiment, and the regularities of celestial mechanics. Perhaps the crucial case, since it is the one under which we as possibly free agents fall, is that of living things. If, as is currently very common, we think of life processes as essentially similar to machines, the realm of deterministic, machine-like causality is beginning to expand rapidly (Glennan 1996).

The minimal assumption needed to account for any of these regularities is that things have (ceteris paribus) dispositions, or capacities. A paradigm for a scientifically interesting capacity might be the disposition of massive objects to move at constant velocity in a straight line. This is of course a much weaker assumption than the universal reign of law; and it might immediately be noted that massive objects seldom do in fact move at constant velocity in a straight line. The exceptional case of determination, in a world containing such capacities, is a set of constraints such that only one kind of behavior, one way that the relevant capacity or capacities can be exercised, is possible. So suppose, for example, that an explosion is in

---

2 I have in mind here the ‘new mechanism’, for which the classic reference is Machamer, Darden, and Craver 2000.
general a largely indeterministic process in which a lot of energy is dissipated very rapidly somehow and somewhere. If, however, we confine the explosion to the interior of a sufficiently strong metal cylinder containing a piston capable of moving a certain distance along its length, then we can be certain that a large proportion of the energy of the explosion will be transmitted into the motion of the piston; very roughly speaking, there is nowhere else for it to go. This is a simple example of how we constrain capacities (in this case the capacity to generate a lot of energy in a very short time) to produce highly reliable behavior. A great many other such constraints need to be ingeniously engineered if the motion of the piston is going to be translated into the rotation of the wheel, and a great deal more still is needed to turn this into a usable vehicle. But I propose that the model of constraining the exercise of capacities will work well throughout this further elaboration of the workings of a familiar machine.

I should note in passing that I am quite happy to admit that the existence of capacities may be fully amenable to reductive explanation. I certainly don’t mean to assert that this is universally the case, and in fact I don’t believe that it is, but for present purposes I have no reason to deny it. What I do deny is that there could be a fully reductive account of the conditions under which the capacity will be exercised. As all the examples so far have illustrated, that will always or at least typically depend on additional circumstances beyond the confines of the entity to which the capacity is attributed. Of course one may attempt at that point to extend the range of the phenomena on which the reduction is proposed to be based. But this will bring in capacities of the additional parts of the system, the exercise of which will again typically depend on further properties beyond the extended system. To summarize the problem, a truly closed system may perhaps be amenable to fully reductive explanation, and this is why the process of making reliable machines can be seen as one of approximating in many respects the production of a closed system. However in nature all systems are open, and reductionist explanation is mainly limited to explaining how things have the capacities they do. It may possibly be that reductionism is true in the sense that what happens always supervenes on the total physical state of the universe; this is hardly, however, a possibility of much relevance to empirical science.

I have mentioned the solar system a couple of times as an example of an untypically deterministic system, and I should say a little more about the general topic of physical law to which this example has had such great historical significance. Let me introduce that discussion with a highly pertinent quote from Elizabeth Anscombe:
The high success of Newton’s astronomy was in one way an intellectual disaster: it produced an illusion from which we tend still to suffer. This illusion was created by the circumstance that Newton’s mechanics had a good model in the solar system. For this gave the impression that we had here an ideal of scientific explanation; whereas the truth was, it was mere obligingness on the part of the solar system, by having had so peaceful a history in recorded time, to provide such a model (Anscombe 1971: 20).

It may even be that part of the legacy of this illusion is the idea that still underlies so many reductionist intuitions, the belief in the completeness of physics. Is this view a problem for my indeterministic, disordered, or dappled view of the casual order? Yes. If physics is the theory that applies to all the stuff in the world, for physics to be complete is for all that stuff to be law-governed; and assuming, as I do, that there is nothing in the world that is not made of physical stuff then under some description, at least, everything is subject to the laws of physics. These laws may perhaps be indeterministic, but here is also where it is often argued that the source of indeterminism in a complete physics, quantum mechanics, is increasingly negligible as we approach the macroscopic world, and determinism is at least indefinitely closely approximated. So should we believe that physics is complete? I am again happy to put a lot of weight on an onus of proof argument: why should we believe the claims of a theory to govern phenomena in principle in a realm (the macroscopic world generally) in which there is no expectation whatever that there could ever be evidence of their successful application? Of course the enthusiasts for physical law will say that the formalisms they deploy work perfectly wherever they can be applied in practice, and they apply in principle to everything. But this is a feeble argument. There are many features of the macroscopic entities and processes to which these laws cannot in practice be applied that are not true of those things to which they can. There are differences in size, complexity, and so on, which may perfectly well make the laws inapplicable in principle as well as in practice. The evidence-free hypothesis that a set of laws that has been found empirically applicable to one kind of phenomenon should be assumed to apply to a much wider range of phenomena for which there is no evidence of such applicability extant or foreseeable, is surely the sort of hypothesis that even the slightest reflection on the hazards of induction should lead us to treat with deep suspicion.

One last piece of philosophical background needs to be considered before I turn to the question of free will itself. This is the vexed topic of downward causation. Can the behavior of anything be caused by the wider system of which it is part? Given my view that causation is a scarce
and special phenomenon, I have generally thought that this should be no problem. So, anticipating the case in point, if humans are complex systems with exceptional capacities to impose causal order, then it should be no surprise that this causal order should encompass the parts of the system: when I raise my arm, I, the whole system, cause a part, my arm, to rise. However, I am increasingly sympathetic to the suggestion that it is better not to describe this issue in causal terms at all. Carl Craver and William Bechtel (Craver and Bechtel 2007) provide a convincing argument that causation works neither upwards – smaller things causing the behavior of the wholes of which they are parts – or downwards, but applies properly only to interactions at a particular level of structural complexity. The movements of parts or wholes consequent on the behavior of wholes or parts, respectively, they call “mechanistically mediated effects.”

Mechanistically mediated effects are hybrids of constitutive and causal relations in a mechanism, where the constitutive relations are interlevel, and the causal relations are exclusively intralevel. Appeal to top-down causation seems spooky or incoherent when it cannot be explicated in terms of mechanistically mediated effects (Craver and Bechtel 2007: 547).

In essence, the point, as I understand it, is just that we should separate the truism that parts and wholes cannot help but coordinate their actions, from the substantive question whether there can be autonomous systems at higher levels, as opposed to merely mechanistically mediated effects of lower level systems. Again, if I am right that causality, or the rule of law, is far from universal, there should be no problem in admitting such autonomous higher level systems. Craver and Bechtel’s analysis, it seems to me, makes it clear that this is at any rate a possibility that should not be ruled out merely by reflection on the nature of causation or the intuition of spookiness that some people find so strongly associated with downward causation.

3. The Metaphysics of Free Will

Voluntarists have generally felt the need to insist on there being some kind of gaps in the surrounding causal order that leave space for genuinely free actions, and the metaphysical views I have just sketched plainly have no problem meeting this need. They have not, however, generally wanted to think of the voluntary acts that take advantage of these gaps as being indeterministic in the sense exemplified by phenomena such as radioactive
Decay. Thinking of free human acts as merely random is probably even less attractive to ideals of human autonomy than is thinking of them as deterministically caused by some phenomena far beyond the control or even knowledge of the agent, such as the transactions of microphysical particles. What the voluntarist wants is some kind of determination not only of a different kind from that in the surrounding causal nexus, but also somehow independent of it. Attempts to provide such a thing have generally involved some kind of agent causation, a way of seeing agents as capable of initiating causal chains in some way independent of the causal nexus in which they are embedded. My past discussions of this topic (Dupré 1996) have tended to follow this line, and indeed I do find compelling the idea that human agents, by virtue of the remarkable organization of their parts, can impose causal order on their more chaotic surroundings. However, I have recently come to see that a better way of presenting my views, and one that avoids the somewhat opaque concept of agent causation, is in terms of indeterministic compatibilism. The appearance of oxymoron in this phrase will be dispelled, I hope, in what follows. But before I explain this any further, I want to talk a little about the development of autonomy from a biological perspective.

It is still often supposed that the development of a human from zygote to maturity and beyond is a kind of unfolding of a program encoded in the genome. But it is now quite clear that this is not the case (e.g. Barnes and Dupré 2008), and that human development is a flexible process requiring access to a diverse range of resources including, as well as the genome, a wide array of other chemicals and structures in the maternal egg, much that is provided by the uterine environment, and subsequent to birth a great deal of nurture and cognitive enrichment, provided by parents, hospitals, schools, etc. If there is such a thing as autonomy, it emerges at some point in this developmental process, the life cycle: fetuses surely don’t have it; well-educated adults do, if anything does. It would be nice to know when this happens, but perhaps more fundamental is the question, what it is for this to happen.

As any organism develops, it acquires a range of capacities: capacities to capture and process resources from the environment, to reproduce, and so on. In the human case, the range of these capacities can be extremely large. Humans acquire abilities to read, write and reason; to interact with...
one another in a range of playful and serious ways; to speak languages, operate machines, play musical instruments, etc. In the case of relatively simple organisms these capacities can give rise to good deal of causal regularity. When a caterpillar emerges from the egg on a suitable vegetable substrate, equipped with the proper array of mouth parts, digestive equipment, and so on to feed, it will very predictably munch away at the poor plant until it has acquired and processed enough chemical resources to pupate. On the other hand when a domestic cat has acquired the ability to catch mice, there is little reason to think any deterministic cause forces it to pursue mice on any particular cue. My cat, at any rate, is often too lazy to go out and catch anything, especially if the weather is bad. In other moods, on the contrary, he will in a remarkably short time bring in a fairly substantial cross section of the small vertebrates that live in our vicinity. If this is the acting out of a causally determined routine it is far from obviously so. Of course, extremely well-ordered capacities are deployed in the process of slaughtering these unfortunate creatures; it is just that whether they will be deployed at all, or whether they will be used for a long period of amusement at the victim’s expense or merely for a quick kill and a snack are highly variable from instance to instance.

In the human case, at any rate, the vision of the exercise of one or some among the repertoire of capacities being elicited deterministically in response to particular features of a situation seems incredible. The relevant kind of causal regularity supposes a systematic response to an environmental cue. But more complex organisms, and especially humans, respond to their environments not as reflexive reactions to experienced conditions, but teleologically, with a view to the production of some among a range of desired outcomes. (Even for my cat, desired outcomes include, at least, amusement and nutrition.) Humans not only acquire large numbers of capacities, abilities to intervene in the world in highly diverse ways, but can also increase the numbers of their goals. Although the contemporary office worker might, at one level of abstraction, claim the sole goal of emptying her email inbox, the experience of confronting a typical inbox is of confronting a range of different problems not all of which can be solved at the same time. Particular events – the arrival of a new email, for instance – may provide particular opportunities to resolve one among these various problems. Humans, whether working at computers, making works of art, tending a smallholding, or whatever, will typically have diverse goals and diverse possible means of achieving those goals. Any of such means will typically depend on the exercise of several or many of the capacities that they have acquired over decades of development. Humans thus take flexi-
bility and the possibility of choice between actions to an unprecedented level. And human intelligence (itself an exceedingly complex set of capacities) makes possible both the formulation of goals and the discernment of routes by which they can be achieved.

All these choices, between ends and between means, could in principle be conceived on a quasi-hydraulic model. Different ends would have different pressures, modified by the assessment of the probability that particular means would achieve them; the greatest pressure would determine action. A first response to this, or no doubt a more sophisticated possible variant, is that only the commitment to the universal presence of determining causes, a commitment I have explicitly disavowed, could draw one to such a picture. And this is because surely the picture is psychologically quite implausible. Consider the exercise of will in a case where akrasia is a pressing option. One might think of this as simply a matter of the composition of competing forces, but though talk of strength of will and weakness of will might encourage such a model, it is surely too simple. When a principle (I have decided to become a non-smoker, say) confronts an occasion for akrasia (a friend lights a cigarette while we are enjoying a drink together), this is certainly not experienced as a conflict between desires. Desire is all on one side; the motivation to act in accordance with a principle, though it may be entirely decisive, is something different. Although Kant’s conception of free action as action in accordance with moral duty seems excessively severe, a slightly less demanding idea of acting in accordance with commitment to principles or goals even despite immediate desires to the contrary surely captures something central to the conception of autonomy. I shall try to make clearer what this might be in the final section of this paper.

4. Voluntarism vs. Compatibilism

Let me recapitulate the story so far. I see causal order as scarce rather than omnipresent, but as exceptionally densely realized by organisms and especially humans. Humans develop multiple and extremely sophisticated capacities throughout their ontogeny and become the densest and most diverse concentrations of causal capacity in our experience. Humans can therefore impose order on their surroundings by acting in ways intended precisely to do so. Unless you are reading these words in the midst of a wilderness hike, the chances are that a casual glance at your immediate surroundings will provide compelling evidence of the ordering powers of
humans. This does not yet amount to autonomy; fully deterministic mechanisms could perhaps impose order on their environment, though perhaps as an empirical fact they could only do this if they were thereby realizing the intentions of a designer. But we at least have more promising materials for an account of autonomy than are provided by traditional determinisms.

In my earlier work, I presented this picture as the basis of a kind of voluntarism, which assumed, though did not really explain in much detail, some kind of agent causation. Humans, on such a view, are a kind of miniature first cause, sending ripples of effect into the world as they pass through it. I don’t think this is wrong, but it is rather unhelpful. What determines what effects the passing human launches on the environment? Is this just a matter of whim? In which case we are little better off than with the account of free action as pure randomness.

Though I don’t mean to suggest that I have discovered a fatal flaw in the kind of incompatibilist voluntarism just sketched, recent work by John Perry (Perry 2010)4 convinced me that a much better solution would be an indeterministic compatibilism, as I shall now explain. The first thing that is needed is to dispel the air of oxymoron. For Perry, compatibilism is not the compatibility of voluntary action with a necessarily deterministic world order, but just with the normal causal order. Recalling Hume’s classic statement of compatibilism, we might immediately note that, although there are passages that suggest that Hume may very well have been a determinist, this is not the main crux of his argument either. Certainly Hume stresses that in specific cases human behavior may be as predictable as anything else. Consider, for example, the reflection that: “A prisoner who has neither money nor interest, discovers the impossibility of his escape, as well when he considers the obstinacy of the gaoler, as the walls and bars with which he is surrounded (Hume 1748: section VIII, part I).” As described in my general remarks about human capacities, the gaoler has acquired not only the necessary capacities to ensure my incarceration, but even the firm commitment to the goal of doing so. He can be relied upon to adopt whatever plan of action seems best suited to pursuing this end. But the situation Hume describes is as readily compatible with the metaphysics presented in this paper as with a deterministic one. Gaolers are

---

4 Perry does not, I should stress, call himself a compatibilist indeterminist, or indeed any other kind of indeterminist. His work does, rather, make clear how the important benefits of compatibilism that have been stressed since David Hume do not require a deterministic framework.
The sorts of people who are very unlikely to take pity on their charges and just let them go.

Perry develops the crucial idea for the present discussion by considering the phrase, notorious in discussions of the topic, “I could have done otherwise”. Generally this phrase is taken to encapsulate a problem for the compatibilist: if I acted freely, it is said, surely I could have done otherwise. But if determinism is true, then it is never the case that I could have done otherwise. So freedom is not compatible with determinism; it is rather the “wretched subterfuge” of Perry’s title (the phrase is borrowed from Kant).

Perry proposes that the anti-compatibilist argument trades on an ambiguity in the interpretation of ‘I could’. On the one hand I may mean, (1) “My capacities do not preclude my having done otherwise. I am causally competent to act or not act.” This is true of the gaoler. He is quite capable of operating locks, opening doors, and so on, and could if he chose, release the prisoner. In fact he is probably better placed to do so than anyone else. But the incompatibilist assumes a quite different interpretation, something like, (2) “I might have done otherwise even though it would have been entirely contrary to my ends, principles, etc.” Note the innocent shift from ‘could’ to ‘might’. If we replaced ‘might’ with ‘could’ in (2) it seems unexceptionable; it is just another way of saying what is expressed by (1). But ‘might’ here suggests that there is some definite probability of my doing something contrary to my ends, principles, etc. Why should we accept this?

Doing something contrary to my goals, principles, etc. is what is generally referred to as akrasia, or weakness of will. In a broadly Kantian vein, it is generally considered exactly the antithesis of autonomy. Strength of will, the causal tendency to do what is necessary to achieve the goals to which one is committed, is on the other hand a condition of genuine autonomy. So the possibility of not doing what my ends dictate, of exhibiting akrasia, sometimes licenses statements such as (2), but surely not always. The gaoler may well have no tendency at all to deviate from his professional responsibility, and it is just not true that he might have done otherwise. Perry illus-

5 There is a great deal of interesting discussion of this question, which the present paper overlooks at some serious risk. In particular, Eleonore Stump, who has provided an interpretation of Aquinas that is in many ways very congenial to the position developed here, argues that voluntarism does not need to be committed to the possibility of the agent acting otherwise (see Stump 1996). I think Stump would understand the position defended here as a version of voluntarism.
trates this point with a less weighty example. He invites us to imagine (or if we have the good fortune to hear the spoken version of his paper, to observe) that he is confronted with some Brussels sprouts, a vegetable which, he informs us, he detests. Might he nevertheless decide to eat them? Without bringing in some entirely new motivation, perhaps some kind of duress, or extreme hunger, the answer is surely no. (The reader is invited to substitute some food, or indeed non-food, that they find appropriately repulsive.) Nonetheless, statement (1) is surely true in this case. To have the capacity to eat sprouts in this sense just means having the ability to convey the things to one’s mouth, having teeth capable of reducing them to pieces small enough to swallow, and so on. He could have done otherwise – eaten the sprouts – but the chance of this happening was zero.

The point can be summarized by saying that what compatibilism insists on is that human action, certainly free human action, is compatible with the normal causal order. If the normal causal order is deterministic then, of course, what happens could not but have happened. Even in this case, it is still true that I could have done otherwise in the sense that I had all the necessary causal capacities to do so; it is not the case that I might have done otherwise, not only because I had no reason to (if I didn’t), but also because in a deterministic world it is never the case that anything might have happened except what actually did. Determinists had better not try to claim that things might have gone differently, and if freedom requires this possibility, they had better admit that there isn’t any. But they can, at least, allow a very plausible sense in which agents may claim that they could have done otherwise.

But indeterminism, or anyhow the kind of indeterminism I advocate, can now do even better. On my view, the normal causal order is a patchy one. Sometimes it tightly constrains what happens, sometimes it is thin and mutable. So John Perry’s aversion to sprouts or the obstinate dedication to duty of Hume’s gaoler constitute causally well-organized features of the world. On the other hand sometimes I may not be much committed to any one course of action. My causal properties may not determine whether I choose an orange or an apple; which is just to say my preferences for fruit are not a firmly established part of the normal causal order. Most interesting, perhaps, are cases that are not yet part of the causal order, but may become so as a result of deliberation. When Caesar first worries whether to cross the Rubicon, it may genuinely not be determined whether he will do so. His consideration of the issue, and his eventual decision, convert the case from one like my indifferent attitude to various fruit, to one more similar to Perry’s stronger views on the subject of veg-
etables. Thus deliberation and decision may change the causal order; the idea, contrary to the assumption of determinism, that the causal order is incomplete makes it possible for such cogitative processes to make a difference to the world. And of course, as in the case of Caesar, these differences may be highly significant.

As is characteristic of this topic, the question can always be pushed back a step. Given Caesar’s particular history, patrician education, and so on, is there any way he could have made any other decision in 49 b.c.e. than to cross the Rubicon? I shall remain agnostic on this question. The point I do want to insist on is that the fixing of the causal order that led to Caesar and his legions crossing the Rubicon was a process that had a history, culminating in whatever deliberations Caesar may have engaged in immediately prior to the seminal event. It may be that the relevant causal order was essentially fixed long prior to this time, or it may have been in the balance until the end. These are details for historians to ponder. The defender of free will should not want to insist that the agent might have done otherwise until the moment of action, even if, in the sense Perry clarifies, they still could have done otherwise. This insistence would lead us towards the view of action as random that advocates of compatibilism have so effectively warned against. But the picture of causal determination crystallizing out against an as yet undetermined background makes it possible that human thought, commitment, moral reflection, and so on, can play an essential role in the determination of the future. And that, surely, is everything that the defender of free will should want.

References


