

Richard Taylor, Metaphysics 4th ed. (Prentice Hall, 1992)

Freedom and Determinism

If I consider the world or any part of it at any particular moment, it seems certain that it is perfectly determinate in every detail. There is no vagueness, looseness, or ambiguity. There is, indeed, vagueness, and even error, in my conceptions of reality, but not in reality itself. A lilac bush, which surely has a certain exact number of blossoms, appears to me only to have many blossoms, and I do not know how many. Things seen in the distance appear of indefinite form, and often of a color and size that in fact they are not. Things near the border of my visual field seem to me vague and amorphous, and I can never even say exactly where that border itself is, it is so indefinite and vague. But all such indeterminateness resides solely in my conceptions and ideas; the world itself shares none of it. The sea, at any exact time and place, has exactly a certain salinity and temperature, and every grain of sand on its shore is exactly disposed with respect to all the others. The wind at any point in space has at any moment a certain direction and force, not more nor less. It matters not whether these properties and relations are known to anyone. A field of wheat at any moment contains just an exact number of ripening grains, each having reached just the ripeness it exhibits, each presenting a determinate color and shade, an exact shape and mass. A person, too, at any given point in his life, is perfectly determinate to the minutest cells of his body. My own brain, nerves — even my thoughts, intentions, and feelings — are at any moment just what they then specifically are. These thoughts might, to be sure, be vague and even false as representations, but as thoughts they are not, and even a false idea is no less an exact and determinate idea than a true one.

Nothing seems more obvious. But if I now ask *why* the world and all its

larger or smaller parts are this moment just what they are, the answer comes to mind: because the world, the moment before, was precisely what it then was. Given exactly what went before, the world, it seems, could now be none other than it is. And what it was a moment before, in all its larger and minuter parts, was the consequence of what had gone just before then, and so on, back to the very beginning of the world, if it had a beginning, or through an infinite past time, in case it had not. In any case, the world as it now is, and every part of it, and every detail of every part, would seem to be the only world that now could be, given just what it has been.

DETERMINISM

Reflections such as this suggest that, in the case of everything that exists, there are antecedent conditions, known or unknown, which, because they are given, mean that things could not be other than they are. That is an exact statement of the metaphysical thesis of determinism. More loosely, it says that everything, including every cause, is the effect of some cause or causes; or that everything is not only determinate but causally determined. The statement, moreover, makes no allowance for time, for past, or for future. Hence, if true, it holds not only for all things that have existed but for all things that do or ever will exist.

Of course people rarely think of such a principle, and hardly one in a thousand will ever formulate it to himself in words. Yet all do seem to assume it in their daily affairs, so much so that some philosophers have declared it an *a priori* principle of the understanding, that is, something that is known independently of experience, while others have deemed it to be at least a part of the common sense of mankind. Thus, when I hear a noise I look up to see where it came from. I never suppose that it was just a noise that came from nowhere and had no cause. Everyone does the same — even animals, though they have never once thought about metaphysics or the principle of universal determinism. People believe, or at least act as though they believed, that things have causes, without exception. When a child or animal touches a hot stove for the first time, it unhesitatingly believes that the pain then felt was caused by that stove, and so firm and immediate is that belief that hot stoves are avoided ever after. We all use our metaphysical principles, whether we think of them or not, or are even capable of thinking of them. If I have a bodily or other disorder — a rash, for instance, or a fever or a phobia — I consult a physician for a diagnosis and explanation in the hope that the cause of it might be found and removed or moderated. I am never tempted to suppose that such things just have no causes, arising from nowhere, else I would take no steps to remove the causes. The principle of determinism is here, as in everything else, simply assumed, without being thought about.

DETERMINISM AND HUMAN BEHAVIOR

I am a part of the world. So is each of the cells and minute parts of which I am composed. The principle of determinism, then, in case it is true, applies to me and to each of those minute parts, no less than to the sand, wheat, winds, and waters of which we have spoken. There is no particular difficulty in thinking so, as long as I consider only what are sometimes called the "purely physiological" changes of my body, like growth, the pulse, glandular secretions, and the like. But what of my thoughts and ideas? And what of my behavior that is supposed to be deliberate, purposeful, and perhaps morally significant? These are all changes of my own being, changes that I undergo, and if these are all but the consequences of the conditions under which they occur, and these conditions are the only ones that could have obtained, given the state of the world just before and when they arose, what now becomes of my responsibility for my behavior and of the control over my conduct that I fancy myself to possess? What am I but a helpless product of nature, destined by her to do whatever I do and to become whatever I become?

There is no moral blame nor merit in anyone who cannot help what he does. It matters not whether the explanation for his behavior is found within him or without, whether it is expressed in terms of ordinary physical causes or allegedly "mental" ones, or whether the causes be proximate or remote. I am not responsible for being a man rather than a woman, nor for having the temperament and desires characteristic of that sex. I was never asked whether these should be given to me. The kleptomaniac, similarly, steals from compulsion, the alcoholic drinks from compulsion, and sometimes even the hero dies from compulsive courage. Though these causes are within them, they compel no less for that, and their victims never chose to have them inflicted upon themselves. To say they are compulsions is to say only that they compel. But to say that they compel is only to say that they cause; for the cause of a thing being given, the effect cannot fail to follow. By the thesis of determinism, however, everything whatever is caused, and not one single thing could ever be other than exactly what it is. Perhaps one thinks that the kleptomaniac and the drunkard did not have to become what they are, that they could have done better at another time and thereby ended up better than they are now, or that the hero could have done worse and then ended up a coward. But this shows only an unwillingness to understand what made them become as they are. Having found that their behavior is caused from within them, we can hardly avoid asking what caused these inner springs of action, and then asking what were the causes of these causes, and so on through the infinite past. We shall not, certainly, with our small understanding and our fragmentary knowledge of the past ever know why the world should at just this time and place have produced just this thief, this drunkard, and this hero, but the vagueness and smattered nature of our knowledge should not tempt us to imagine a similar vagueness in nature herself. Everything in nature is and always has been

determinate, with no loose edges at all, and she was forever destined to bring forth just what she has produced, however slight may be our understanding of the origins of these works. Ultimate responsibility for anything that exists, and hence for any person and his deeds, can thus rest only with the first cause of all things, if there is such a cause, or nowhere at all, in case there is not. Such, at least, seems to be the unavoidable implication of determinism.

DETERMINISM AND MORALS

Some philosophers, faced with all this, which seems quite clear to the ordinary understanding, have tried to cling to determinism while modifying traditional conceptions of morals. They continue to *use* such words as *merit*, *blame*, *praise*, and *desert*, but they so divest them of their meanings as to finish by talking about things entirely different, sometimes without themselves realizing that they are no longer on the subject. An ordinary person will hardly understand that anyone can possess merit or vice and be deserving of moral praise or blame, as a result of traits that he has or of behavior arising from those traits, once it is well understood that he could never have avoided being just what he is and doing just what he does.

We are happily spared going into all this, however, for the question whether determinism is true of human nature is not a question of ethics at all but of metaphysics. There is accordingly no hope of answering it within the context of ethics. One can, to be sure, simply *assume* an answer to it — assume that determinism is true, for instance — and then see what are the implications of this answer for ethics; but that does not answer the question. Or one can *assume* some theory or other of ethics — assume some version of “the greatest happiness” principle, for instance — and then see whether that theory is consistent with determinism. But such confrontations of theories with theories likewise make us no wiser, so far as any fundamental question is concerned. We can suppose at once that determinism is consistent with some conceptions of morals, and inconsistent with others, and that the same holds for indeterminism. We shall still not know what theories are true; we shall only know which are consistent with another.

We shall, then, eschew all considerations of ethics as having no real bearing on our problem. We want to learn, if we can, whether determinism is true, and this is a question of metaphysics. It can, like all good questions of philosophy, be answered only on the basis of certain data; that is, by seeing whether or not it squares with certain things that everyone knows, or believes himself to know, or with things everyone is at least more sure about than the answer to the question at issue.

Now I could, of course, simply affirm that I am a morally responsible being, in the sense in which my responsibility for my behavior implies that I could have avoided that behavior. But this would take us into the nebulous realm of

ethics, and it is, in fact, far from obvious that I am responsible in that sense. Many have doubted that they are responsible in that sense, and it is in any case not difficult to doubt it, however strongly one might feel about it.

There are, however, two things about myself of which I feel quite certain and that have no necessary connection with morals. The first is that I sometimes deliberate, with the view to making a decision; a decision, namely, to do this thing or that. And the second is that whether or not I deliberate about what to do, it is sometimes up to me what I do. This might all be an illusion, of course; but so also might any philosophical theory, such as the theory of determinism, be false. The point remains that it is far more difficult for me to doubt that I sometimes deliberate, and that it is sometimes up to me what to do, than to doubt any philosophical theory whatever, including the theory of determinism. We must, accordingly, if we ever hope to be wiser, adjust our theories to our data and not try to adjust our data to our theories.

Let us, then, get these two data quite clearly before us so we can see what they are, what they presuppose, and what they do and do not entail.

DELIBERATION

Deliberation is an activity, or at least a kind of experience, that cannot be defined, or even described, without metaphors. We speak of weighing this and that in our minds, of trying to anticipate consequences of various possible courses of action, and so on, but such descriptions do not convey to us what deliberation is unless we already know.

Whenever I deliberate, however, I find that I make certain presuppositions, whether I actually think of them or not. That is, I assume that certain things are true, certain things which are such that, if I thought they were not true, it would be impossible for me to deliberate at all. Some of these can be listed as follows:

First, I find that I can deliberate only about my own behavior and never about the behavior of another. I can try to guess, speculate, or figure out what another person is going to do; I can read certain signs and sometimes infer what he will do; but I cannot deliberate about it. When I deliberate I try to decide something, to make up my mind, and this is as remote as anything could be from speculating, trying to guess, or inferring from signs. Sometimes one *does* speculate on what he is going to do, by trying to draw conclusions from certain signs or omens — he might infer that he is going to sneeze, for instance, or speculate that he is going to become a grandfather — but he is not then deliberating whether to do things or not. One does, to be sure, sometimes deliberate about whether another person will do a certain act, when that other person is subject to his command or otherwise under his control; but then he is not really deliberating about another person's acts at all, but about his own — namely, whether or not to have that other person carry out the order.

Second, I find that I can deliberate only about future things, never things past or present. I may not know what I did at a certain time in the past, in case I have forgotten, but I can no longer deliberate whether to do it then or not. I can, again, only speculate, guess, try to infer, or perhaps try to remember. Similarly, I cannot deliberate whether or not to be doing something now; I can only ascertain whether or not I am in fact doing it. If I am sitting I cannot deliberate about whether or not to be sitting. I can only deliberate about whether to remain sitting — and this has to do with the future.

Third, I cannot deliberate about what I shall do if I already know what I am going to do. If I were to say, for example, "I know that I am going to be married tomorrow and in the meantime I am going to deliberate about whether to get married," I would contradict myself. There are only two ways that I could know now what I am going to do tomorrow; namely, either by inferring this from certain signs and omens or by having already decided what I am going to do. But if I have inferred from signs and omens what I am going to do, I cannot deliberate about it — there is just nothing for me to decide; and similarly, if I have already decided. If, on the other hand, I can still deliberate about what I am going to do, to that extent I must regard the signs and omens as unreliable, and the inference uncertain, and I therefore do not know what I am going to do after all.

And finally, I cannot deliberate about what to do, even though I may not know what I am going to do, unless I believe that it is up to me what I am going to do. If I am within the power of another person, or at the mercy of circumstances over which I have no control, then, although I may have no idea what I am going to do, I cannot deliberate about it. I can only wait and see. If, for instance, I am a serviceman, and regulations regarding uniforms are posted each day by my commanding officer and are strictly enforced by him, then I shall not know what uniforms I shall be wearing from time to time, but I cannot deliberate about it. I can only wait and see what regulations are posted; it is not up to me. Similarly, a woman who is about to give birth to a child cannot deliberate whether to have a boy or a girl, even though she may not know. She can only wait and see; it is not up to her. Such examples can be generalized to cover any case wherein one does not know what he is going to do but believes that it is not up to him, and hence no matter for his decision and hence none for his deliberation.

IT IS UP TO ME"

I sometimes feel certain that it is, at least to some extent, up to me what I am going to do; indeed, I must believe this if I am to deliberate about what to do. But what does this mean? It is, again, hard to say, but the idea can be illustrated, and we can fairly easily see what it does *not* mean.

Let us consider the simplest possible sort of situation in which this belief

might be involved. At this moment, for instance, it seems quite certain to me that, holding my finger before me, I can move it either to the left or to the right, that each of these motions is possible for me. This does not mean merely that my finger can move either way, although it entails that, for this would be true in case nothing obstructed it, even if I had no control over it at all. I can say of a distant, fluttering leaf that it can move either way, but not that I can move it, since I have no control over it. How it moves is not up to me. Nor does it mean merely that my finger can be moved either way, although it entails this too. If the motions of my finger are under the control of some other person or of some machine, then it might be true that the finger can be moved either way, by that person or machine, though false that I can move it at all.

If I say, then, that it is up to me how I move my finger, I mean that I can move it in this way and I can move it in that way, and not merely that it can move or be moved in this way and that. I mean that the motion of my finger is within my direct control. If someone were to ask me to move it to the right, I could do that, and if he were to ask me to move it to the left, I could do that too. Further, I could do these simple acts without being asked at all, and having been asked, I could move it in a manner the exact opposite of what was requested, since I can ignore the request. There are, to be sure, some motions of my finger that I cannot make, so it is not *entirely* up to me how it moves. I cannot bend it backward, for instance, or bend it into a knot, for these motions are obstructed by the very anatomical construction of the finger itself; and to say that I can move my finger at all means at least that nothing obstructs such a motion, though it does not mean merely this. There is, however, at this moment, no obstruction, anatomical or otherwise, to my moving it to the right, and none to my moving it to the left.

This datum, it should be noted, is properly expressed as a conjunction and not as a disjunction. That is, my belief is that I can move my finger in one way *and* that I can also move it another way; and it does not do justice to this belief to say that I can move it one way *or* the other. It is fairly easy to see the truth of this, for the latter claim, that I can move it one way *or* the other, would be satisfied in case there were only one way I could move it, and *that* is not what I believe. Suppose, for instance, that my hand were strapped to a device in such a fashion that I could move my finger to the right but not to the left. Then it would still be entirely true that I could move it either to the left *or* to the right — since it would be true that I could move it to the right. But that is not what I now believe. My finger is not strapped to anything, and nothing obstructs its motion in either direction. And what I believe, in this situation, is that I can move it to the right *and* I can move it to the left.

We must note further that the belief expressed in our datum is not a belief in what is logically impossible. It is the belief that I now *can* move my finger in different ways but not that I can move it in different ways at once. What I believe is that I am now able to move my finger one way and that I am now equally able to move it another way, but I do not claim to be able now or at any

other time to move it both ways simultaneously. The situation here is analogous to one in which I might, for instance, be offered a choice of either of two apples but forbidden to take both. Each apple is such that I may select it, but neither is such that I may select it together with the other.

Now, are these two data — the belief that I do sometimes deliberate, and the belief that it is sometimes up to me what I do — consistent with the metaphysical theory of determinism? We do not know yet. We intend to find out. It is fairly clear, however, that they are going to present difficulties to that theory. But let us not, in any case, try to avoid those difficulties by just denying the data themselves. If we eventually deny the data, we shall do so for better reasons than this. Virtually everyone is convinced that beliefs such as are expressed in our data are sometimes true. They cannot be simply dismissed as false just because they might appear to conflict with a metaphysical theory that hardly anyone has ever really thought much about at all. Almost anyone, unless his fingers are paralyzed, bound, or otherwise incapable of movement, believes sometimes that the motions of his fingers are within his control, in exactly the sense expressed by our data. If consequences of considerable importance to him depend on how he moves his fingers, he sometimes deliberates before moving them, or at least he is convinced that he does or that he can. Philosophers might have different notions of just what things are implied by such data, but there is in any case no more, and in fact considerably less, reason for denying the data than for denying some philosophical theory.

CAUSAL VERSUS LOGICAL NECESSITY

Philosophers have long since pointed out that causal connections involve no logical necessity, that the denial of a particular causal connection is never self-contradictory, and this is undoubtedly true. But neither does the assertion or the denial of determinism involve any concept of what is and what is not logically necessary. If determinism is true, then anything that happens is, given the conditions under which it occurs, the only thing possible, the thing that is necessitated by those conditions. But it is not the only thing that is logically possible, nor do those conditions logically necessitate it. Similarly, if one denies the thesis of determinism by asserting, for instance, that each of two bodily motions is possible for him under identical conditions, he is asserting much more than that each is logically possible, for that would be a trivial claim.

This distinction, between logical necessity and the sort of necessity involved in determinism, can be illustrated with examples. If, for instance, a man is beheaded, we can surely say that it is impossible for him to go on living, that his being beheaded necessitates his death, and so on; but there are no logical necessities or impossibilities involved here. It is not logically impossible for a man to live without his head. Yet no one will deny that a man cannot live under conditions that include his being headless, that such a state of affairs is

in a perfectly clear sense impossible. Similarly, if my finger is in a tight and fairly strong cast, then it is impossible for me to move it in any way at all, though this is not logically impossible. It is logically possible that I should be vastly stronger than I am, and that I should move my finger and, in doing so, break the cast, though this would ordinarily not be possible in the sense that concerns us. Again, it is not logically impossible that I should bend my finger backward, or into a knot, though it is, in fact, impossible for me to do either or, what means the same thing, necessary that I should do neither. Certain conditions prohibit my doing such things, though they impose no logical barrier. And finally, if someone — a physician, for example — should ask me whether I can move my finger, and I should reply truthfully that I can, I would not merely be telling her that it is logically possible for me to move it, for this she already knows. I would be telling her that I am able to move it, that it is within my power to do so, that there are no conditions, such as paralysis or whatnot, that prevent my moving it.

It follows that not all necessity is logical necessity, nor all impossibility logical impossibility, and that to say that something is possible is sometimes to say much more than that it is logically possible. The kind of necessity involved in the thesis of determinism is quite obviously the nonlogical kind, as is also the kind of possibility involved in its denial. If we needed a *name* for these nonlogical modalities, we could call them *causal* necessity, impossibility, and possibility, but the concepts are clear enough without making a great deal of the name.

FREEDOM

To say that it is, in a given instance, up to me what I do is to say that I am in that instance *free* with respect to what I then do. Thus, I am sometimes free to move my finger this way and that, but not, certainly, to bend it backward or into a knot. But what does this mean?

It means, first that there is no *obstacle* or *impediment* to my activity. Thus, there is sometimes no obstacle to my moving my finger this way and that, though there are obvious obstacles to my moving it backward or into a knot. Those things, accordingly, that pose obstacles to my motions limit my freedom. If my hand were strapped in such a way as to permit only a leftward motion of my finger, I would not then be free to move it to the right. If it were encased in a tight cast that permitted no motion, I would not be free to move it at all. Freedom of motion, then, is limited by obstacles.

Further, to say that it is, in a given instance, up to me what I do, means that nothing *constrains* or *forces* me to do one thing rather than another. Constraints are like obstacles, except that while the latter prevent, the former enforce. Thus, if my finger is being forcibly bent to the left — by a machine, for instance, or by another person, or by any force that I cannot overcome — then

I am not free to move it this way and that. I cannot, in fact, move it at all; I can only watch to see how it is moved, and perhaps vainly resist: Its motions are not up to me, or within my control, but in the control of some other thing or person.

Obstacles and constraints, then, both obviously limit my freedom. To say that I am free to perform some action thus means at least that there is no obstacle to my doing it, and that nothing constrains me to do otherwise.

Now if we rest content with this observation, as many have, and construe free activity simply as activity that is unimpeded and unconstrained, there is evidently no inconsistency between affirming both the thesis of determinism and the claim that I am sometimes free. For to say that some action of mine is neither impeded nor constrained does not by itself imply that it is not causally determined. The absence of obstacles and constraints is a mere negative condition, and does not by itself rule out the presence of positive causes. It might seem, then, that we can say of some of my actions that there are conditions antecedent to their performance so that no other actions were possible, and also that these actions were unobstructed and unconstrained. And to say that would logically entail that such actions were both causally determined, and free.

SOFT DETERMINISM

It is this kind of consideration that has led many philosophers to embrace what is sometimes called "soft determinism." All versions of this theory have in common three claims, by means of which, it is naively supposed, a reconciliation is achieved between determinism and freedom. Freedom being, furthermore, a condition of moral responsibility and the only condition that metaphysics seriously questions, it is supposed by the partisans of this view that determinism is perfectly compatible with such responsibility. This, no doubt, accounts for its great appeal and wide acceptance, even by some people of considerable learning.

The three claims of soft determinism are (1) that the thesis of determinism is true, and that accordingly all human behavior, voluntary or other, like the behavior of all other things, arises from antecedent conditions, given which no other behavior is possible — in short, that all human behavior is caused and determined; (2) that voluntary behavior is nonetheless free to the extent that it is not externally constrained or impeded; and (3) that, in the absence of such obstacles and constraints, the causes of voluntary behavior are certain states, events, or conditions within the agent himself; namely, his own acts of will or volitions, choices, decisions, desires, and so on.

Thus, on this view, I am free, and therefore sometimes responsible for what I do, provided nothing prevents me from acting according to my own choice, desire, or volition, or constrains me to act otherwise. There may, to be sure, be

other conditions for my responsibility — such as, for example, an understanding of the probable consequences of my behavior, and that sort of thing — but absence of constraint or impediment is, at least, one such condition. And, it is claimed, it is a condition that is compatible with the supposition that my behavior is caused — for it is, by hypothesis, caused by my own inner choices, desires, and volitions.

THE REFUTATION OF THIS

The theory of soft determinism looks good at first — so good that it has for generations been solemnly taught from innumerable philosophical chairs and implanted in the minds of students as sound philosophy — but no great acumen is needed to discover that far from solving any problem, it only camouflages it.

My free actions are those unimpeded and unconstrained motions that arise from my own inner desires, choices, and volitions; let us grant this provisionally. But now, whence arise those inner states that determine what my body shall do? Are they within my control or not? Having made my choice or decision and acted upon it, could I have chosen otherwise or not?

Here the determinist, hoping to surrender nothing and yet to avoid the problem implied in that question, bids us not to ask it; the question itself, he announces, is without meaning. For to say that I could have done otherwise, he says, means only that I *would* have done otherwise, *if* those inner states that determined my action had been different; if, that is, I had decided or chosen differently. To ask, accordingly, whether I could have chosen or decided differently is only to ask whether, had I decided to decide differently or chosen to choose differently, or willed to will differently, I *would* have decided or chosen or willed differently. And this, of course, is unintelligible nonsense.

But it is not nonsense to ask whether the causes of my actions — my own inner choices, decisions, and desires — are themselves caused. And of course they are, if determinism is true, for on that thesis everything is caused and determined. And if they are, then we cannot avoid concluding that, given the causal conditions of those inner states, I could not have decided, willed, chosen, or desired other than I, in fact, did, for this is a logical consequence of the very definition of determinism. Of course we can still say that, *if* the causes of those inner states, whatever they were, had been different, then their effects, those inner states themselves, would have been different, and that in this hypothetical sense I could have decided, chosen, willed, or desired differently — but that only pushes our problem back still another step. For we will then want to know whether the causes of those inner states were within my control, and so on *ad infinitum*. We are, at each step, permitted to say "could have been otherwise" only in a provisional sense — provided, that is, that something else had been different — but must then retract it and replace it with "could not

46 have been otherwise" as soon as we discover, as we must at each step, that whatever would have to have been different could not have been different.

EXAMPLES

Such is the dialectic of the problem. The easiest way to see the shadowy quality of soft determinism, however, is by means of examples.

Let us suppose that my body is moving in various ways, that these motions are not externally constrained or impeded, and that they are all exactly in accordance with my own desires, choices, or acts of will and whatnot. When I will that my arm should move in a certain way, I find it moving in that way, unobstructed and unconstrained. When I will to speak, my lips and tongue move, unobstructed and unconstrained, in a manner suitable to the formation of the words I choose to utter. Now, given that this is a correct description of my behavior, namely, that it consists of the unconstrained and unimpeded motions of my body in response to my own volitions, then it follows that my behavior is free, on the soft determinist's definition of "free." It follows further that I am responsible for that behavior; or at least, that if I am not, it is not from any lack of freedom on my part.

But if the fulfillment of these conditions renders my behavior free — that is to say, if my behavior satisfies the conditions of free action set forth in the theory of soft determinism — then my behavior will be no less free if we assume further conditions that are perfectly consistent with those already satisfied.

We suppose further, accordingly, that while my behavior is entirely in accordance with my own volitions, and thus "free" in terms of the conception of freedom we are examining, my volitions themselves are caused. To make this graphic, we can suppose that an ingenious physiologist can induce in me any volition he pleases, simply by pushing various buttons on an instrument to which, let us suppose, I am attached by numerous wires. All the volitions I have in that situation are, accordingly, precisely the ones he gives me. By pushing one button, he evokes in me the volition to raise my hand; and my hand, being unimpeded, rises in response to that volition. By pushing another, he induces the volition in me to kick, and my foot, being unimpeded, kicks in response to that volition. We can even suppose that the physiologist puts a rifle in my hands, aims it at some passerby, and then, by pushing the proper button, evokes in me the volition to squeeze my finger against the trigger, whereupon the passerby falls dead of a bullet wound.

This is the description of a man who is acting in accordance with his inner volitions, a man whose body is unimpeded and unconstrained in its motions, these motions being the effects of those inner states. It is hardly the description of a free and responsible agent. It is the perfect description of a puppet. To render someone your puppet, it is not necessary forcibly to constrain the motions of his limbs, after the fashion that real puppets are moved. A subtler but

no less effective means of making a person your puppet would be to gain complete control of his inner states, and ensuring, as the theory of soft determinism does ensure, that his body will move in accordance with them.

The example is somewhat unusual, but it is no worse for that. It is perfectly intelligible, and it does appear to refute the soft determinist's conception of freedom. One might think that, in such a case, the agent should not have allowed himself to be so rigged in the first place, but this is irrelevant; we can suppose that he was not aware that he was and was hence unaware of the source of those inner states that prompted his bodily motions. The example can, moreover, be modified in perfectly realistic ways, so as to coincide with actual and familiar cases. One can, for instance, be given a compulsive desire for certain drugs, simply by having them administered over a course of time. Suppose, then, that I do, with neither my knowledge nor consent, thus become a victim of such a desire and act upon it. Do I act freely, merely by virtue of the fact that I am unimpeded in my quest for drugs? In a sense I do, surely, but I am hardly free with respect to whether or not I shall use drugs. I never chose to have the desire for them inflicted upon me.

Nor does it, of course, matter whether the inner states that allegedly prompt all my "free" activity are evoked in me by another agent or by perfectly impersonal forces. Whether a desire that causes my body to behave in a certain way is inflicted upon me by another person, for instance, or derived from hereditary factors, or indeed from anything at all, matters not the least. In any case, if it is in fact the cause of my bodily behavior, I cannot help but act in accordance with it. Wherever it came from, whether from personal or impersonal origins, it was entirely caused or determined, and not within my control. Indeed, if determinism is true, as the theory of soft determinism holds it to be, all those inner states that cause my body to behave in whatever ways it behaves must arise from circumstances that existed before I was born; for the chain of causes and effects is infinite, and none could have been the least different, given those that preceded.

SIMPLE INDETERMINISM

We might at first now seem warranted in simply denying determinism, and saying that, insofar as they are free, my actions are not caused; or that, if they are caused by my own inner states — my own desires, impulses, choices, volitions, and whatnot — then these, in any case, are not caused. This is a perfectly clear sense in which a person's action, assuming that it was free, could have been otherwise. If it was uncaused, then, even given the conditions under which it occurred and all that preceded, some other act was nonetheless possible, and he did not have to do what he did. Or if his action was the inevitable consequence of his own inner states, and could not have been otherwise, given these, we can nevertheless say that these inner states, being uncaused,

could have been otherwise, and could thereby have produced different actions.

Only the slightest consideration will show, however, that this simple denial of determinism has not the slightest plausibility. For let us suppose it is true, and that some of my bodily motions — namely, those that I regard as my free acts — are not caused at all or, if caused by my own inner states, that these are not caused. We shall thereby avoid picturing a puppet, to be sure — but only by substituting something even less like a human being; for the conception that now emerges is not that of a free person, but of an erratic and jerking phantom, without any rhyme or reason at all.

Suppose that my right arm is free, according to this conception; that is, that its motions are uncaused. It moves this way and that from time to time, but nothing causes these motions. Sometimes it moves forth vigorously, sometimes up, sometimes down, sometimes it just drifts vaguely about — these motions all being wholly free and uncaused. Manifestly I have nothing to do with them at all; they just happen, and neither I nor anyone can ever tell what this arm will be doing next. It might seize a club and lay it on the head of the nearest bystander, no less to my astonishment than his. There will never be any point in asking why these motions occur, or in seeking any explanation of them, for under the conditions assumed there is no explanation. They just happen, from no causes at all.

This is no description of free, voluntary, or responsible behavior. Indeed, so far as the motions of my body or its parts are entirely uncaused, such motions cannot even be ascribed to me as my behavior in the first place, since I have nothing to do with them. The behavior of my arm is just the random motion of a foreign object. Behavior that is mine must be behavior that is within my control, but motions that occur from no causes are beyond the control of anyone. I can have no more to do with, and no more control over, the uncaused motions of my limbs than a gambler has over the motions of an honest roulette wheel. I can only, like him, idly wait to see what happens.

Nor does it improve things to suppose that my bodily motions are caused by my own inner states, so long as we suppose these to be wholly uncaused. The result will be the same as before. My arm, for example, will move this way and that, sometimes up and sometimes down, sometimes vigorously and sometimes just drifting about, always in response to certain inner states, to be sure. But since these are supposed to be wholly uncaused, it follows that I have no control over them and hence none over their effects. If my hand lays a club forcefully on the nearest bystander, we can indeed say that this motion resulted from an inner club-wielding desire of mine; but we must add that I had nothing to do with that desire, and that it arose, to be followed by its inevitable effect, no less to my astonishment than to his. Things like this do, alas, sometimes happen. We are all sometimes seized by compulsive impulses that arise we know not whence, and we do sometimes act upon these. But because they are far from being examples of free, voluntary, and responsible behavior, we

need only to learn that the behavior was of this sort to conclude that it was not free, voluntary, or responsible. It was erratic, impulsive, and irresponsible.

DETERMINISM AND SIMPLE INDETERMINISM AS THEORIES

Both determinism and simple indeterminism are loaded with difficulties, and no one who has thought much on them can affirm either of them without some embarrassment. Simple indeterminism has nothing whatever to be said for it, except that it appears to remove the grossest difficulties of determinism, only, however, to imply perfect absurdities of its own. Determinism, on the other hand, is at least initially plausible. People seem to have a natural inclination to believe in it; it is, indeed, almost required for the very exercise of practical intelligence. And beyond this, our experience appears always to confirm it, so long as we are dealing with everyday facts of common experience, as distinguished from the esoteric researches of theoretical physics. But determinism, as applied to human behavior, has implications that few can casually accept, and they appear to be implications that no modification of the theory can efface.

Both theories, moreover, appear logically irreconcilable to the two items of data that we set forth at the outset; namely, (1) that my behavior is sometimes the outcome of my deliberation, and (2) that in these and other cases it is sometimes up to me what I do. Because these were our data, it is important to see, as must already be quite clear, that these theories cannot be reconciled to them.

I can deliberate only about my own future actions, and then only if I do not already know what I am going to do. If a certain nasal tickle warns me that I am about to sneeze, for instance, then I cannot deliberate whether to sneeze or not; I can only prepare for the impending convulsion. But if determinism is true, then there are always conditions existing antecedently to everything I do, sufficient for my doing just that, and such as to render it inevitable. If I can know what those conditions are and what behavior they are sufficient to produce, then I can in every such case know what I am going to do and cannot then deliberate about it.

By itself this only shows, of course, that I can deliberate only in ignorance of the causal conditions of my behavior; it does not show that such conditions cannot exist. It is odd, however, to suppose that deliberation should be a mere substitute for clear knowledge. Ignorance is a condition of speculation, inference, and guesswork, which have nothing whatever to do with deliberation. A prisoner awaiting execution may not know when he is going to die, and he may even entertain the hope of reprieve, but he cannot deliberate about this. He can only speculate, guess — and wait.

Worse yet, however, it now becomes clear that I cannot deliberate about what I am going to do, if it is even *possible* for me to find out in advance,

whether I do in fact find out in advance or not. I can deliberate only with the view to deciding what to do, to making up my mind; and this is impossible if I believe that it could be inferred what I am going to do from conditions already existing, even though I have not made that inference myself. If I believe that what I am going to do has been rendered inevitable by conditions already existing, and could be inferred by anyone having the requisite sagacity, then I cannot try to decide whether to do it or not, for there is simply nothing left to decide. I can at best only guess or try to figure it out myself or, all prognostics failing, I can wait and see; but I cannot deliberate. I deliberate in order to *decide what to do*, not to *discover what it is that I am going to do*. But if determinism is true, then there are always antecedent conditions sufficient for everything that I do, and this can always be inferred by anyone having the requisite sagacity; that is, by anyone having a knowledge of what those conditions are and what behavior they are sufficient to produce.

This suggests what in fact seems quite clear, that determinism cannot be reconciled with our second datum either, to the effect that it is sometimes up to me what I am going to do. For if it is ever really up to me whether to do this thing or that, then, as we have seen, each alternative course of action must be such that I can do it; not that I can do it in some abstruse or hypothetical sense of "can"; not that I could do it if only something were true that is not true; but in the sense that it is then and there within my power to do it. But this is never so, if determinism is true, for on the very formulation of that theory whatever happens at any time is the only thing that can then happen, given all that precedes it. It is simply a logical consequence of this that whatever I do at any time is the only thing I can then do, given the conditions that precede my doing it. Nor does it help in the least to interpose, among the causal antecedents of my behavior, my own inner states, such as my desires, choices, acts of will, and so on. For even supposing these to be always involved in voluntary behavior — which is highly doubtful in itself — it is a consequence of determinism that these, whatever they are at any time, can never be other than what they then are. Every chain of causes and effects, if determinism is true, is infinite. This is why it is not now up to me whether I shall a moment hence be male or female. The conditions determining my sex have existed through my whole life, and even prior to my life. But if determinism is true, the same holds of anything that I ever am, ever become, or ever do. It matters not whether we are speaking of the most patent facts of my being, such as my sex; or the most subtle, such as my feelings, thoughts, desires, or choices. Nothing could be other than it is, given what was; and while we may indeed say, quite idly, that something — some inner state of mind, for instance — *could* have been different, had only something *else* been different, any consolation of this thought evaporates as soon as we add that whatever would have to have been different could not have been different.

It is even more obvious that our data cannot be reconciled to the theory of simple indeterminism. I can deliberate only about my own actions; this is

obvious. But the random, uncaused motion of any body whatever, whether it be a part of my body or not, is no action of mine and nothing that is within my power. I might try to guess what these motions will be, just as I might try to guess how a roulette wheel will behave, but I cannot deliberate about them or try to decide what they shall be, simply because these things are not up to me. Whatever is not caused by anything is not caused by me, and nothing could be more plainly inconsistent with saying that it is nevertheless up to me what it shall be.

THE THEORY OF AGENCY

The only conception of action that accords with our data is one according to which people — and perhaps some other things too — are sometimes, but of course not always, self-determining beings; that is, beings that are sometimes the causes of their own behavior. In the case of an action that is free, it must not only be such that it is caused by the agent who performs it, but also such that no antecedent conditions were sufficient for his performing just that action. In the case of an action that is both free and rational, it must be such that the agent who performed it did so for some reason, but this reason cannot have been the cause of it.

Now, this conception fits what people take themselves to be; namely, beings who act, or who are agents, rather than beings that are merely acted upon, and whose behavior is simply the causal consequence of conditions that they have not wrought. When I believe that I have done something, I do believe that it was I who caused it to be done, I who made something happen, and not merely something within me, such as one of my own subjective states, which is not identical with myself. If I believe that something not identical with myself was the cause of my behavior — some event wholly external to myself, for instance, or even one internal to myself, such as a nerve impulse, volition, or whatnot — then I cannot regard that behavior as being an act of mine, unless I further believe that I was the cause of that external or internal event. My pulse, for example, is caused and regulated by certain conditions existing within me, and not by myself. I do not, accordingly, regard this activity of my body as my action, and would be no more tempted to do so if I became suddenly conscious within myself of those conditions or impulses that produce it. This is behavior with which I have nothing to do, behavior that is not within my immediate control, behavior that is not only not free activity, but not even the activity of an agent to begin with; it is nothing but a mechanical reflex. Had I never learned that my very life depends on this pulse beat, I would regard it with complete indifference, as something foreign to me, like the oscillations of a clock pendulum that I idly contemplate.

Now this conception of activity, and of an agent who is the cause of it, involves two rather strange metaphysical notions that are never applied

elsewhere in nature. The first is that of a *self* or *person* — for example, a man — who is not merely a collection of things or events, but a self-moving being. For on this view it is a person, and not merely some part of him or something within him, that is the cause of his own activity. Now, we certainly do not know that a human being is anything more than an assemblage of physical things and processes that act in accordance with those laws that describe the behavior of all other physical things and processes. Even though he is a living being, of enormous complexity, there is nothing, apart from the requirements of this theory, to suggest that his behavior is so radically different in its origin from that of other physical objects, or that an understanding of it must be sought in some metaphysical realm wholly different from that appropriate to the understanding of nonliving things.

Second, this conception of activity involves an extraordinary conception of causation according to which an agent, which is a substance and not an event, can nevertheless be the cause of an event. Indeed, if he is a free agent then he can, on this conception, cause an event to occur — namely, some act of his own — without anything else causing him to do so. This means that an agent is sometimes a cause, without being an antecedent sufficient condition; for if I affirm that I am the cause of some act of mine, then I am plainly not saying that my very existence is sufficient for its occurrence, which would be absurd. If I say that my hand causes my pencil to move, then I am saying that the motion of my hand is, under the other conditions then prevailing, sufficient for the motion of the pencil. But if I then say that I cause my hand to move, I am not saying anything remotely like this, and surely not that the motion of my self is sufficient for the motion of my arm and hand, since these are the only things about me that are moving.

This conception of the causation of events by things that are not events is, in fact, so different from the usual philosophical conception of a cause that it should not even bear the same name, for “being a cause” ordinarily just means “being an antecedent sufficient condition or set of conditions.” Instead, then, of speaking of agents as *causing* their own acts, it would perhaps be better to use another word entirely, and say, for instance, that they *originate* them, *initiate* them, or simply that they *perform* them.

Now this is, on the face of it, a dubious conception of what a person is. Yet it is consistent with our data, reflecting the presuppositions of deliberation, and appears to be the only conception that is consistent with them, as determinism and simple indeterminism are not. The theory of agency avoids the absurdities of simple indeterminism by conceding that human behavior is caused, while at the same time avoiding the difficulties of determinism by denying that every chain of causes and effects is infinite. Some such causal chains, on this view, have beginnings, and they begin with agents themselves. Moreover, if we are to suppose that it is sometimes up to me what I do, and understand this in a sense that is not consistent with determinism, we must suppose that I am an agent or a being who initiates his own actions, sometimes under conditions

that do not determine what action I shall perform. Deliberation becomes, on this view, something that is not only possible but quite rational, for it does make sense to deliberate about activity that is truly my own and that depends in its outcome upon me as its author, and not merely upon something more or less esoteric that is supposed to be intimately associated with me, such as my thoughts, volitions, choices or whatnot.

One can hardly affirm such a theory of agency with complete comfort, however, and not wholly without embarrassment, for the conception of agents and their powers which is involved in it is strange indeed, if not positively mysterious. In fact, one can hardly be blamed here for simply denying our data outright, rather than embracing this theory to which they do most certainly point. Our data — to the effect that we do sometimes deliberate before acting, and that, when we do, we presuppose among other things that it is up to us what we are going to do — rest upon nothing more than fairly common consent. These data might simply be illusions. It might, in fact, be that no one ever deliberates but only imagines that he does, that from pure conceit he supposes himself to be the master of his behavior and the author of his acts. Spinoza has suggested that if a stone, having been thrown into the air, were suddenly to become conscious, it would suppose itself to be the source of its own motion, being then conscious of what it was doing but not aware of the real cause of its behavior. Certainly we are *sometimes* mistaken in believing that we are behaving as a result of choice deliberately arrived at. A man might, for example, easily imagine that his embarking upon matrimony is the result of the most careful and rational deliberation, when in fact the causes, perfectly sufficient for that behavior, might be of an entirely physiological, unconscious origin. If it is sometimes false that we deliberate and then act as the result of a decision deliberately arrived at, even when we suppose it to be true, it might always be false. No one seems able, as we have noted, to describe deliberation without metaphors, and the conception of a thing's being “within one's power” or “up to him” seems to defy analysis or definition altogether, if taken in a sense that the theory of agency appears to require.

These are, then, dubitable conceptions, despite their being so well implanted in common sense. Indeed, when we turn to the theory of fatalism, we shall find formidable metaphysical considerations that appear to rule them out altogether. Perhaps here, as elsewhere in metaphysics, we should be content with discovering difficulties, with seeing what is and what is not consistent with such convictions as we happen to have, and then drawing such satisfaction as we can from the realization that, no matter where we begin, the world is mysterious and that we who try to understand it are even more so. This realization can, with some justification, make one feel wise, even in the full realization of his ignorance.