

6

Free Will

Thomas Nagel

Suppose you're going through a cafeteria line and when you come to the desserts, you hesitate between a peach and a big wedge of chocolate cake with creamy icing. The cake looks good, but you know it's fattening. Still, you take it and eat it with pleasure. The next day you look in the mirror or get on the scale and think, "I wish I hadn't eaten that chocolate cake. I could have had a peach instead."

"I could have had a peach instead." What does that mean, and is it true?

Peaches were available when you went through the cafeteria line: you had the *opportunity* to take a peach instead. But that isn't all you mean. You mean you could have *taken* the peach instead of the cake. You could have *done* some-

What Does It All Mean?

thing different from what you actually did. Before you made up your mind, it was open whether you would take fruit or cake, and it was only your choice that decided which it would be.

Is that it? When you say, "I could have had a peach instead," do you mean that it depended only on your choice? You chose chocolate cake, so that's what you had, but *if* you had chosen the peach, you would have had that.

This still doesn't seem to be enough. You don't mean only that *if* you had chosen the peach, you would have had it. When you say, "I could have had a peach instead," you also mean that you *could have chosen* it—no "ifs" about it. But what does that mean?

It can't be explained by pointing out other occasions when you *have* chosen fruit. And it can't be explained by saying that if you had thought about it harder, or if a friend had been with you who eats like a bird, you *would* have chosen it. What you are saying is that you could have chosen a peach instead of chocolate cake *just then, as things actually were*. You think you could have chosen a peach even if everything else had been exactly the same as it was up to the point when you in fact chose chocolate cake. The only difference would have been that instead of thinking, "Oh well," and reaching for the cake, you would have thought, "Better not," and reached for the peach.

Free Will

This is an idea of “can” or “could have” which we apply only to people (and maybe some animals). When we say, “The car could have climbed to the top of the hill,” we mean the car had enough power to reach the top of the hill *if* someone drove it there. We don’t mean that on an occasion when it was parked at the bottom of the hill, the car could have just taken off and climbed to the top, instead of continuing to sit there. Something else would have had to happen differently first, like a person getting in and starting the motor. But when it comes to people, we seem to think that they can do various things they don’t actually do, *just like that*, without anything else happening differently first. What does this mean?

Part of what it means may be this: Nothing up to the point at which you choose determines irrevocably what your choice will be. It remains an *open possibility* that you will choose a peach until the moment when you actually choose chocolate cake. It isn’t determined in advance.

Some things that happen *are* determined in advance. For instance, it seems to be determined in advance that the sun will rise tomorrow at a certain hour. It is not an open possibility that tomorrow the sun won’t rise and night will just continue. That is not possible because it could happen only if the earth stopped rotating, or the sun stopped existing, and there is nothing going

What Does It All Mean?

on in our galaxy which might make either of those things happen. The earth will continue rotating unless it is stopped, and tomorrow morning its rotation will bring us back around to face inward in the solar system, toward the sun, instead of outward, away from it. If there is no possibility that the earth will stop or that the sun won't be there, there is no possibility that the sun won't rise tomorrow.

When you say you could have had a peach instead of chocolate cake, part of what you mean may be that it wasn't determined in advance what you would do, as it *is* determined in advance that the sun will rise tomorrow. There were no processes or forces at work before you made your choice that made it inevitable that you would choose chocolate cake.

That may not be all you mean, but it seems to be at least part of what you mean. For if it was really determined in advance that you would choose cake, how could it also be true that you could have chosen fruit? It would be true that nothing would have prevented you from having a peach if you had chosen it instead of cake. But these *ifs* are not the same as saying you could have chosen a peach, period. You couldn't have chosen it unless the possibility remained open until you closed it off by choosing cake.

Some people have thought that it is never pos-

Free Will

sible for us to do anything different from what we actually do, in this absolute sense. They acknowledge that what we do depends on our choices, decisions, and wants, and that we make different choices in different circumstances: we're not like the earth rotating on its axis with monotonous regularity. But the claim is that, in each case, the circumstances that exist before we act determine our actions and make them inevitable. The sum total of a person's experiences, desires and knowledge, his hereditary constitution, the social circumstances and the nature of the choice facing him, together with other factors that we may not know about, all combine to make a particular action in the circumstances inevitable.

This view is called determinism. The idea is not that we can know all the laws of the universe and use them to *predict* what will happen. First of all, we can't know all the complex circumstances that affect a human choice. Secondly, even when we do learn something about the circumstances, and try to make a prediction, that is itself a *change* in the circumstances, which may change the predicted result. But predictability isn't the point. The hypothesis is that there *are* laws of nature, like those that govern the movement of the planets, which govern everything that happens in the world—and that in accor-

What Does It All Mean?

dance with those laws, the circumstances before an action determine that it will happen, and rule out any other possibility.

If that is true, then even while you were making up your mind about dessert, it was already determined by the many factors working on you and in you that you would choose cake. You *couldn't* have chosen the peach, even though you thought you could: the process of decision is just the working out of the determined result inside your mind.

If determinism is true for everything that happens, it was already determined before you were born that you would choose cake. Your choice was determined by the situation immediately before, and *that* situation was determined by the situation before *it*, and so on as far back as you want to go.

Even if determinism isn't true for everything that happens—even if some things just happen without being determined by causes that were there in advance—it would still be very significant if everything *we did* were determined before we did it. However free you might feel when choosing between fruit and cake, or between two candidates in an election, you would really be able to make only one choice in those circumstances—though if the circumstances or your desires had been different, you would have chosen differently.

Free Will

If you believed that about yourself and other people, it would probably change the way you felt about things. For instance, could you blame yourself for giving in to temptation and having the cake? Would it make sense to say, "I really should have had a peach instead," if you *couldn't* have chosen a peach instead? It certainly wouldn't make sense to say it if there *was* no fruit. So how can it make sense if there *was* fruit, but you couldn't have chosen it because it was determined in advance that you would choose cake?

This seems to have serious consequences. Besides not being able sensibly to blame yourself for having had cake, you probably wouldn't be able sensibly to blame anyone at all for doing something bad, or praise them for doing something good. If it was determined in advance that they would do it, it was inevitable: they couldn't have done anything else, given the circumstances as they were. So how can we hold them responsible?

You may be very mad at someone who comes to a party at your house and steals all your Glenn Gould records, but suppose you believed that his action was determined in advance by his nature and the situation. Suppose you believed that everything he did, including the earlier actions that had contributed to the formation of his character, was determined in advance by ear-

What Does It All Mean?

lier circumstances. Could you still hold him responsible for such low-grade behavior? Or would it be more reasonable to regard him as a kind of natural disaster—as if your records had been eaten by termites?

People disagree about this. Some think that if determinism is true, no one can reasonably be praised or blamed for anything, any more than the rain can be praised or blamed for falling. Others think that it still makes sense to praise good actions and condemn bad ones, even if they were inevitable. After all, the fact that someone was determined in advance to behave badly doesn't mean that he *didn't* behave badly. If he steals your records, that shows inconsiderateness and dishonesty, whether it was determined or not. Furthermore, if we don't blame him, or perhaps even punish him, he'll probably do it again.

On the other hand, if we think that what he did was determined in advance, this seems more like punishing a dog for chewing on the rug. It doesn't mean we hold him responsible for what he did: we're just trying to influence his behavior in the future. I myself don't think it makes sense to blame someone for doing what it was impossible for him not to do. (Though of course determinism implies that it was determined in advance that I would think this.)

Free Will

These are the problems we must face if determinism is true. But perhaps it isn't true. Many scientists now believe that it isn't true for the basic particles of matter—that in a given situation, there's more than one thing that an electron may do. Perhaps if determinism isn't true for human actions, either, this leaves room for free will and responsibility. What if human actions, or at least some of them, are not determined in advance? What if, up to the moment when you choose, it's an open possibility that you will choose either chocolate cake or a peach? Then, so far as what has happened before is concerned, you *could* choose either one. Even if you actually choose cake, you could have chosen a peach.

But is even this enough for free will? Is this all you mean when you say, "I could have chosen fruit instead?"—that the choice wasn't determined in advance? No, you believe something more. You believe that *you* determined what you would do, by *doing* it. It wasn't determined in advance, but it didn't *just happen*, either. *You did it*, and you could have done the opposite. But what does that mean?

This is a funny question: we all know what it means to *do* something. But the problem is, if the act wasn't determined in advance, by your desires, beliefs, and personality, among other

What Does It All Mean?

things, it seems to be something that just happened, without any explanation. And in that case, how was it *your doing*?

One possible reply would be that there is no answer to that question. Free action is just a basic feature of the world, and it can't be analyzed. There's a difference between something just happening without a cause and an action just being *done* without a cause. It's a difference we all understand, even if we can't explain it.

Some people would leave it at that. But others find it suspicious that we must appeal to this unexplained idea to explain the sense in which you could have chosen fruit instead of cake. Up to now it has seemed that determinism is the big threat to responsibility. But now it seems that even if our choices are not determined in advance, it is still hard to understand in what way we *can* do what we don't do. Either of two choices may be possible in advance, but unless I determine which of them occurs, it is no more my responsibility than if it was determined by causes beyond my control. And how can *I* determine it if *nothing* determines it?

This raises the alarming possibility that we're not responsible for our actions whether determinism is true *or* whether it's false. If determinism is true, antecedent circumstances are re-

Free Will

sponsible. If determinism is false, nothing is responsible. That would really be a dead end.

There is another possible view, completely opposite to most of what we've been saying. Some people think responsibility for our actions *requires* that our actions be determined, rather than requiring that they not be. The claim is that for an action to be something you have done, it has to be produced by certain kinds of causes in you. For instance, when you chose the chocolate cake, that was something you did, rather than something that just happened, because you wanted chocolate cake more than you wanted a peach. Because your appetite for cake was stronger at the time than your desire to avoid gaining weight, it resulted in your choosing the cake. In other cases of action, the psychological explanation will be more complex, but there will always be one—otherwise the action wouldn't be yours. This explanation seems to mean that what you did was determined in advance after all. If it wasn't determined by anything, it was just an unexplained event, something that happened out of the blue rather than something that you did.

According to this position, causal determination by itself does not threaten freedom—only a certain *kind* of cause does that. If you grabbed

What Does It All Mean?

the cake because someone else pushed you into it, then it wouldn't be a free choice. But free action doesn't require that there be no determining cause at all: it means that the cause has to be of a familiar psychological type.

I myself can't accept this solution. If I thought that everything I did was determined by my circumstances and my psychological condition, I would feel trapped. And if I thought the same about everybody else, I would feel that they were like a lot of puppets. It wouldn't make sense to hold them responsible for their actions any more than you hold a dog or a cat or even an elevator responsible.

On the other hand, I'm not sure I understand how responsibility for our choices makes sense if they are *not* determined. It's not clear what it means to say *I* determine the choice, if nothing about me determines it. So perhaps the feeling that you could have chosen a peach instead of a piece of cake is a philosophical illusion, and couldn't be right whatever was the case.

To avoid this conclusion, you would have to explain (a) what you *mean* if you say you could have done something other than what you did, and (b) what you and the world would have to be like for this to be true.