

CHAPTER 6

Free Will and Determinism

Theodore Sider

The Problem

Suppose you are kidnaped and forced to commit a series of terrible murders. The kidnaper makes you shoot a first victim by forcing your finger to squeeze the trigger of a gun, hypnotizes you into poisoning a second, and then throws you from an airplane, causing you to squash a third. Miraculously, you survive the fall from the airplane. You stagger from the scene, relieved that the ordeal is over. But then, to your amazement, you are apprehended by the police, who handcuff you and charge you with murder. The parents of the victims scream obscenities at you as you are led away in disgrace.

Are the police and parents fair to blame you for the killings? Obviously not, for you have an unassailable excuse: you did not act of your own free will. You couldn't help what you did; you could not have done otherwise. And only those who act freely are morally responsible.

We all believe that we have free will. How could we not? Renouncing freedom would mean no longer planning for the

future, for why make plans if you are not free to change what will happen? It would mean renouncing morality, for only those who act freely deserve blame or punishment. Without freedom, we march along pre-determined paths, unable to control our destinies. Such a life is not worth living.

Yet freedom seems to conflict with a certain apparent fact. Incredibly, this fact is no secret; most people are fully aware of it. We uncritically accept free will only because we fail to put two and two together. The problem of free will is 'a time bomb hidden within our most deeply held beliefs.

Here is the fact: *every event has a cause*. This fact is known as *determinism*.

We all believe in causes. If scientists discovered debris in the upper stratosphere spelling out 'Ozzy Osbourne!', they would immediately go to work to discover the cause. Was the debris put there by a renegade division of NASA comprised of heavy-metal fans? Was it a science project from a school for adolescent geniuses? If these things were ruled out as causes, the scientists would start to consider stranger hypotheses. Perhaps aliens from another planet are playing a joke on us. Perhaps the debris is left over from a collision between comets, and the resemblance to the name of the heavy-metal singer is purely coincidental. Perhaps different bits of the debris each have different kinds of causes. Any of these hypotheses might be entertained. But the one thing the scientists would *not* contemplate is that there simply is no cause whatsoever. Causes can be hard to discover, or coincidental, or have many different parts, but they are always there.

It's not that uncaused events are utterly inconceivable. We can imagine what it would be like for an uncaused event to occur. For that matter, we can imagine what it would be like for all sorts of strange things to occur: pigs flying, monkeys making 10,000 feet tall statues from jello, and so on. But it is reasonable to believe that no such things *in fact* occur. Likewise, it is reasonable

to believe that there are in fact no uncaused events—that is, it is reasonable to believe in determinism.

Our belief in determinism is reasonable because we have all seen science succeed, again and again, in its search for the underlying causes of things. Technological innovations owe their existence to science: skyscrapers, vaccination, rocket ships, the internet. Science seems to explain everything we observe: the changing of the seasons, the movement of the planets, the inner workings of plants and animals. Given this track record, we reasonably expect the march of scientific progress to continue; we expect that science will eventually discover the causes of everything.

The threat to freedom comes when we realize that this march will eventually overtake us. From the scientific point of view, human choices and behavior are just another part of the natural world. Like the seasons, planets, plants, and animals, our actions are studyable, predictable, explainable, controllable. It is hard to say when, if ever, scientists will learn enough about what makes humans tick in order to predict everything we do. But regardless of when the causes of human behavior are *discovered*, determinism assures us that these causes *exist*.

It is hard to accept that one's own choices are subject to causes. Suppose you become sleepy and are tempted to put down this book. The causes are trying to put you to sleep. But you resist them! You are strong and continue reading anyway. Have you thwarted the causes and refuted determinism? Of course not. Continuing to read has its own cause. Perhaps your love of metaphysics overcomes your drowsiness. Perhaps your parents taught you to be disciplined. Or perhaps you are just stubborn. No matter what the reason, there was some cause.

You may reply: 'But I felt no compulsion to read or not to read; I simply decided to do one or the other. I sensed no cause.' It is true that many thoughts, feelings, and decisions do not feel caused. But this does not really threaten determinism. Some-

times the causes of our decisions aren't consciously detectable, but those causes still exist. Some causes of behavior are preconscious functions of the brain, as contemporary psychology teaches, or perhaps even subconscious desires, as Freud thought. Other causes of decisions may not even be mental. The brain is an incredibly complicated physical object, and might 'swerve' this way or that as a result of certain motions of its tiniest parts. Such purely physical causes cannot be detected merely by directing one's attention inward, no matter how long and hard and calmly one meditates. We can't expect to be able to detect all the causes of our decisions just by introspection.

So: determinism is true, even for human actions. But now, consider any allegedly free action. To illustrate how much is at stake here, let's consider an action that is horribly morally reprehensible: Hitler's invasion of Poland in 1939. We most certainly blame Hitler for this action. We thus consider him to have acted freely. But determinism seems to imply that Hitler was not free at all. Homer

To see why, we must first investigate the concepts of cause and effect. A cause is an earlier event that *makes* a later effect happen. Given the laws of nature,¹ once the cause has occurred, the effect *must* occur. Lightning causes thunder: the laws of nature governing electricity and sound guarantee that, when lightning strikes, thunder will follow. CIE

Determinism says that Hitler's invasion of Poland was caused by some earlier event. So far, there is little to threaten Hitler's freedom. The cause of the invasion might be something under Hitler's control, in which case the invasion would also be under his control. For instance, the cause might be a decision that Hitler made just before the invasion. If so, then it seems we can still blame Hitler for ordering the invasion.

¹ Chapter 9 discusses laws of nature.

But now consider this decision itself. It is just another event. So determinism implies that it too must have a cause. This new cause might be an even earlier decision Hitler made, or something his advisers told him, or something he ate, or, more likely, a combination of many factors. Whatever it is, call this cause of Hitler's decision to invade Poland ' c '. Notice that c also caused the invasion of Poland. For as we saw above, a cause is an earlier event that makes a later event happen. Once c occurred, Hitler's decision had to occur; and once that decision occurred, the invasion had to occur.

We can repeat this reasoning indefinitely. Determinism implies that c must have an earlier cause c_1 , which in turn must have an earlier cause c_2 , and so on. The resulting sequence of events stretches back in time:

... $c_2 \rightarrow c_1 \rightarrow c \rightarrow$ the decision \rightarrow the invasion

Each event in the sequence causes the invasion, since each event causes the event that occurs immediately after it, which then causes the next event occurring immediately after that one, and so on. The final few events in this sequence look like ones under Hitler's control. But the earlier ones do not, for as we move back in time, we eventually reach events before Hitler's birth.

This argument can be repeated for any human action, however momentous or trivial. Suppose an old man slips while crossing the street, and I laugh at him instead of helping him up. Using the above chain of reasoning, we can show that my laughter was caused by events before my birth.

Things now look very bad for freedom. Hitler no longer seems to have had a free choice about whether to invade Poland. I seem to have had no choice but to laugh at the old man. For these actions were all caused by things outside our control. But then what was morally wrong about what Hitler or I did? How can we blame Hitler for invading Poland if it was settled

before his birth that he would do it? How can we blame me for laughing? How can we blame anyone for anything?

We can restate the challenge to freedom in terms of physics. Any action or decision involves the motion of sub-atomic particles in one's body and brain. These sub-atomic particles move according to the laws of physics. Physics lets us calculate the future positions of particles from information about (i) the previous states of the particles, and (ii) the forces acting on the particles. So, in principle, one could have examined the sub-atomic particles one hundred years before the invasion of Poland, calculated exactly how those particles would be moving one hundred years later, and thereby calculated that Hitler would invade Poland. Such calculations are far too difficult to ever complete in practice, but that doesn't matter. Whether or not anyone could have completed the calculations, *the particles were there*, before Hitler's birth, and the fact that they were there, and arranged in the way that they were, made it *inevitable* that Hitler would invade Poland. Once again, we have found a cause for Hitler's invasion that already existed before Hitler was born. And the existence of such a cause seems to imply that Hitler's invasion of Poland was not a free action.

And yet, it *must* have been free, for how else can we *blame* him for this despicable act? The time bomb has exploded. Two of our most deeply held beliefs, our belief in science and our belief in freedom and morality, seem to contradict each other. We must resolve this conflict.

Hard Determinism

The simplest strategy for resolution is to reject one of the beliefs that produce the conflict. One could reject free will, or one could reject determinism.

The rejection of free will in the face of determinism is called **hard determinism**. Think of the hard determinist as a hard-nosed intel-

PHYSICS

lectual who tolerates no softies. Free will conflicts with science, so free will has got to go. Here is a typical hard determinist speech:

We must get used to the idea that no one is really responsible for anything. Belief in freedom and moral responsibility was a luxury of a pre-scientific age. Now that we have grown up, we must put aside childish ways and face the facts. Science has disproved the existence of freedom and morality.

Can we live with this depressing philosophy? Philosophers must seek the truth, however difficult it may be to accept. Maybe hard determinism is one of those difficult truths. Hard determinists might attempt 'damage control', arguing that life without freedom is not as bad as one might think. Society might still punish criminals, for instance. Hard determinists must deny that criminals *deserve* punishment, since the crimes were not committed freely. But they can say that there is still a use for punishment: punishing criminals keeps them off the streets and discourages future crimes. Still, accepting hard determinism is nearly unthinkable. Nor is it clear that one *could* stop believing in free will, even if one wanted to. If you find someone who claims to believe hard determinism, here's a little experiment to try. Punch him in the face, really hard. Then try to convince him not to blame you. After all, according to him, you had no choice but to punch him! I predict you will find it very difficult to convince him to practice what he preaches.

Hard determinism is a position of last resort. Let's see what the other options look like.

Libertarianism

If the hard determinist is the intellectually hard-nosed devotee of science, the libertarian² has the opposite mindset. Libertarians

² The use of the word 'libertarian' in politics is unrelated.

resolve the conflict between free will and determinism by rejecting determinism. Their guiding thought is that *people are special*. The march of science, subjugating observed phenomena to exceptionless law, is limited to the non-human realm. For libertarians, science is good as far as it goes, but it will never succeed in completely predicting human behavior. Humans, and humans alone, transcend the laws of nature: they are free.

What makes people so special? Some libertarians answer that we have souls, non-physical sources of consciousness, which make choices that are not controlled by laws of nature. Others say that humans are indeed purely physical systems, but that they are not subject to the natural laws that govern other physical systems. Either way, laws of nature do not wholly determine human behavior.

Although libertarians are clear on what freedom *isn't*—namely, determinism—they have a little more trouble telling us what freedom *is*. They do not want to say that freedom is merely uncaused action. Saying *that* would equate freedom with *randomness*, and libertarians don't want to do that. Here's why.

Suppose Mother Teresa discovers a hand-grenade in an orphanage in Calcutta. As you might expect, she picks up the hand-grenade in order to dispose of it safely. But now an utterly uncaused event occurs: to her horror, her hand suddenly pulls out the pin and throws the grenade into the heart of the orphanage. The grenade explodes, resulting in mayhem and destruction. When I say 'uncaused', I really mean that there is *no* cause, none whatsoever. As I am imagining the example, the action of pulling the pin and throwing the grenade was not caused by any decision on Mother Teresa's part; nor did it have an external physical cause. No dormant dark side of Mother Teresa's personality has finally come to light. She has no nervous tic. Her hand simply flew up from absolutely no cause whatsoever. This clearly is *not* a free action. We could not blame Mother Teresa; she is the victim of a cruel accident.

The alarming thing for libertarians is that Mother Teresa seems *unfree* precisely because her action was uncaused. Freedom now appears to *require* causation. This obviously threatens the fundamental libertarian claim that the key to the problem of freedom is indeterminism of human action. Libertarians must somehow distinguish between free undetermined action and randomness.

Some libertarians address this problem by postulating a special kind of causation that only humans wield, called agent causation. Ordinary mechanistic causation, the kind studied in physics and the other hard sciences, obeys laws. Mechanistic causes are repeatable and predictable: if you repeat the same cause again and again, the very same effect is guaranteed to occur each time. Agent causation, on the other hand, does not obey laws. There is no saying which way a free human being will exercise her agent causation. The very same person in exactly similar circumstances might agent-cause different things. According to the theory of agent causation, you act freely when (i) your action is not caused in the ordinary, mechanistic way, but (ii) your action is caused by you—by agent causation. If you freely decide to eat Wheaties one morning rather than your usual helping of Apple Jacks, it would have been impossible to predict beforehand which cereal you would choose. Nevertheless, your choice was not a random occurrence, for you yourself caused it. You caused it by agent causation.

It is unclear whether agent causation really solves the problem of randomness. Consider what an agent-causation theorist would say about your freely making a difficult decision. There are two important factors in decision-making: what you desire, and what you believe is the best means to achieve that desire. If you are undecided whether to vote Democrat or Republican in a US presidential election, for instance, this is because some of your beliefs and desires favor a Democratic vote, and others favor a Republican vote. Suppose that, in the end, the set favoring a

Democratic vote wins out. A libertarian would say that mechanistic causes that occurred in the past did not determine this outcome. It was you yourself, via agent causation, that selected the Democratic vote. Your selection was subject to no laws; it was unpredictable. This activity of agent causation was not caused by your beliefs and desires. But now—and here is the problem—since the selection was not causally based in your beliefs and desires, it seems entirely *detached* from you. The selection did not emerge from what you know about the candidates and what sort of leader you want for your country. Your vote didn't arise from who you are. It just appeared in the world, as if by magic. Given this, it would be odd to praise or blame you for it. And this suggests that it was *unfree*.

Whether or not libertarianism relies on agent causation, its most worrisome feature is its clash with science. First, libertarians must reject the possibility of an all-encompassing psychology. Human behavior would be governed by the laws of such a science, and libertarians deny that human behavior is controlled by any laws. But the clash does not end there. Libertarians must also reject the possibility of an all-encompassing *physics*. The realms of psychology and physics cannot be neatly separated, for human bodies are physical objects, made up of subatomic particles. An all-encompassing physics could predict the future motions of *all* particles—even those in human bodies—based on the earlier states of particles. Since libertarians say that human behavior cannot be scientifically predicted, they must deny the possibility of such a physics. According to libertarians, if physicists turned their measuring instruments on the subatomic particles composing a free person, formerly observed patterns would break down.

This attitude toward science seems rash. Here in the twenty-first century, we have the benefit of hindsight on various disagreements between science, on the one hand, and religion and philosophy, on the other. Remember the Catholic Church's

decision to censor Copernicus and Galileo for saying that the Earth moves around the Sun. No one wants to repeat that mistake. And remember the dramatic successes of science, both theoretical and technological. Of course, science is not infallible. But a philosopher had better have *very* good reasons to declare that an existing science is just plain wrong, or that a certain kind of scientific progress will never happen. One's philosophy should avoid colliding with or limiting science.

Our choices look grim. On the one hand, there is the dismal philosophy of hard determinism, which robs life of all that is distinctly human and worthwhile. On the other hand, there is the radically anti-scientific philosophy of libertarianism—which, given the problem of randomness, may not even succeed in salvaging free will.

Interlude: Quantum Mechanics

Before moving on, we should investigate a side issue: whether quantum mechanics bears on the problem of freedom. Quantum mechanics is a theory about the behavior of tiny particles. This theory was developed in the early part of the twentieth century and continues to be accepted by physicists today. Quantum mechanics (or at least, a certain version of it) is a radically indeterministic theory. It does not predict with certainty what will occur; it only gives *probabilities* of outcomes. No matter how much information you have about a particle, you cannot predict with certainty where it will be later. All you can say is how likely it is that the particle will be found in various locations. And this is not a mere limitation on human knowledge. The particle's future position is simply not determined by the past, regardless of how much we know about it. Only the probabilities are determined.

In the previous sections I was ignoring quantum mechanics. For instance, I assumed that if a cause occurs, its effect *must*

occur, even though quantum mechanics says that causes merely make their effects probable. Why did I ignore quantum mechanics? Because randomness is not freedom. Let us try a little thought experiment. First pretend that quantum mechanics is incorrect and physics is truly deterministic. The threat to human freedom that this presents is what we have been talking about so far in this chapter. Next, in each person's brain, add a little lottery, which every so often randomly causes the person to swerve one way rather than another. This is like what quantum mechanics says really happens: there is an element of randomness to what events occur. Does the threat to freedom go away? Clearly not. If the original, wholly determined person had no free will, then the new, randomized person has no free will either; the lottery injects only randomness, not freedom or responsibility. And as we learned from the case of Mother Teresa, randomness does not mean freedom. If anything, randomness undermines freedom.

A libertarian might concede that quantum randomness is not *sufficient* for freedom, but nevertheless claim that quantum randomness *makes room for* freedom, because it makes room for agent causation. Imagine that it is 1939, and Hitler has not yet decided to invade Poland. He is trying to decide what to do among the following three options:

Invade Poland

Invade France

Stop being such an evil guy and become a ballet dancer

Quantum mechanics assigns probabilities to each of these possible decisions; it does not say which one Hitler will choose. Suppose, for the sake of argument, that the probabilities are as follows:

95.0% Invade Poland

4.9% Invade France

0.1% Become a ballet dancer

After assigning these probabilities, the work of quantum mechanics is complete. According to some libertarians, agent causation now steps in. After quantum mechanics sets the probabilities, Hitler himself chooses, by agent causation, which decision he will in fact make. Physics sets probabilities, but *people*, by agent causation, ultimately decide what occurs.

If this picture were correct, then my criticism of libertarianism as being anti-scientific would be rebutted: agent causation could peacefully coexist with quantum mechanics. In fact, though, the coexistence picture makes agent causation a slave to quantum-mechanical probabilities.

Imagine running the following interesting (if wildly unethical) experiment. First produce one million exact clones of Hitler as he was in 1939. Then, in one million separate laboratories, reproduce the exact conditions that Hitler faced before he decided to invade Poland. Put each clone in his own laboratory and deceive him into thinking that it is really 1939 and that he is in charge of Germany. Then sit back and watch. Record how many clones attempt to invade Poland, how many attempt to invade France, and how many attempt to become ballet dancers. The coexistence picture says that you will observe a distribution of behaviors that roughly matches the probabilities listed above, for the coexistence picture says that quantum mechanics correctly gives the probabilities of outcomes. Thus, you will observe around 950,000 clones trying to invade Poland, around 49,000 trying to invade France, and around 1,000 practicing ballet. If you repeat the procedure again and again, you will continue to observe outcomes in approximately the same ratios. (The more times you repeat the experiment, the closer the total ratios will match the probabilities, just as the more times one flips a coin, the closer the ratio of heads to tails approaches one-to-one.) If

you change the laboratory conditions faced by the clones, so that quantum mechanics predicts different probabilities, you will observe a new distribution of behaviors that fits the new probabilities. The distribution keeps following what quantum mechanics says.

What good then is agent causation? It seems to mindlessly follow the probabilities, having no effect of its own on the distribution of outcomes. This sort of agent causation is empty; it adds nothing to freedom or responsibility. Agent causation, if it is to be worth anything, must be capable of disrupting the probabilities given by quantum mechanics. There can be no peaceful coexistence: agent causation theorists must clash with science. Quantum mechanics does not help the agent-causation theorist. So I will go back to ignoring quantum mechanics.

We are back to the grim dilemma. Apparently, we must reject science or reject freedom. Yet neither option seems at all appealing.

Compatibilism

Many philosophers believe that there is a way out of this dilemma. Others think that this way out is a big mistake. You must decide for yourself.

The way out is called *compatibilism*. According to compatibilists, our discussion took a wrong turn all the way back when we said that the available options were rejecting freedom or rejecting determinism. Compatibilists say that this overlooks a third option. We can have our cake and eat it too: we can retain *both* freedom and determinism. That way we can preserve both our science and our humanity. The argument in the first section, which concluded that freedom and determinism are opposed to each other, was a mistake. Free will is in fact compatible with determinism. The alleged conflict is an illusion, based on a misunderstanding of the concept of free will. Our actions (or at

least their probabilities) are indeed caused by events before our births. But they are often free despite this.

To explain what compatibilists are up to, let's first consider some examples. Imagine a very young boy with a serious misunderstanding of the concept of a *man*. This boy thinks it is part of the definition of the word 'man' that men never cry. As far as he knows, the men in his family never cry, the men on television never cry, and so on. He believes that his father is a man, of course, but one day he sees his father crying. The boy becomes very confused. Two of his beliefs now conflict: his belief that his father is a man and his belief that his father is crying. Which should he give up? Should he decide that his father is not a man after all? Or should he decide that his father was not really crying—that he was only cutting up onions, say? Obviously, he should do neither. Instead, he should clear up his conceptual confusion about the nature of manhood. Then he will see that his beliefs about his father's manhood and about his father's crying are compatible after all.

Here is a second example. How would you define the word 'contact', as in 'Barry Bonds' bat made contact with the baseball? If you are like most people, your first answer is probably something like this: *things are in 'contact' when there is no empty space between them*. But now remember your high-school science. Baseballs and bats are made up of atoms. These atoms consist of nuclei and surrounding clouds of electrons. When one atom approaches another, the electrons of the atoms repel one another with electromagnetic forces. The closer together the atoms get, the stronger the forces become. Eventually the forces become so strong that they push the atoms away from each other. This occurs when the atoms get very close to each other, but before their clouds of electrons start to overlap. Thus, as Bonds' bat closed in on the baseball, the outermost atoms of the bat began to repel the outermost atoms of the ball, until eventually the ball came to a halt and flew in the opposite direction. At every

moment there was some space between the bat and the ball. In fact, there is *never* absolutely zero space between bats and balls, nor between fists and jaws, fingers and computer keyboards, or any other things we consider to be in contact. Yet we all believe that contact regularly occurs. So we have another apparent conflict, this time between our belief in high-school science and our belief that things are regularly in contact. Should we renounce one of these beliefs? Obviously not. We should instead reject the proposed definition of 'contact'. Those who accept that definition are in a sense conceptually confused. For things can be in contact even when there is a small amount of space in between them. (What then is the correct definition of contact? Tough question! What about: *things are in contact when there is no visible space in between?* This is only a start.)

The compatibilist makes a similar claim about free will. Determinism seems to conflict with freedom only because we misunderstand the concept of freedom. If 'free' meant 'uncaused', then the conflict would be real. But that's not what 'free' means. (Remember Mother Teresa.) Once we clear up our conceptual confusion, the conflict will vanish. Then we can believe in *both* free will and determinism. Properly understood, they were never really opposed.

So far so good. But if 'free' doesn't mean 'uncaused', what *does* it mean? The compatibilist wants to say, roughly, that a *free action is one that is caused in the right way*. When you were kidnaped and forced to commit murders, your actions were unfree because they were caused in the *wrong* way. Free actions, such as Hitler's invasion of Poland, my writing of this chapter, and your reading it, also have causes, but they are caused in the *right* way. All actions have causes, but having a cause doesn't settle whether an action is free. Whether it is free is settled by what kind of cause it has. If free actions are those that are caused in the right way, as this definition says, then an action can be *both free and* caused.

Thus, given this definition, freedom and determinism do not conflict.

Hard determinists and libertarians may object that all causes should be treated alike. So long as my choice is caused by events before my birth, it is unfree; it does not matter *how* it is caused. But for some purposes, compatibilists can reply, it is clear that causes are not all alike. Causing a running back to fall by tackling him is legal football; causing him to fall by shooting him with a crossbow is not. The rules of football treat some causes differently from others. According to compatibilists, we can think of freedom and morality in an analogous way. Morality, like football, has rules. These rules treat some causes differently from others. If an action is caused in a certain way—the right way—then the rules of morality count that action as free. But if an action is caused in the wrong way, then the rules count that action as unfree.

It is admittedly strange that my actions can be free even though they were caused by events that occurred before I was born. Some philosophers reject compatibilism on this basis. But given the implausibility of hard determinism and libertarianism, compatibilism at least deserves a fair hearing.

Compatibilists must refine their theory, though. When they say that free actions must be caused 'in the right way', what exactly does that mean? Examples were given: Hitler's invasion was caused in the right way; murders coerced by your kidnaper were caused in the wrong way. But examples are not good enough. We need a definition.

Here is a first stab: a *free action is one that is caused by the person's beliefs and desires*. This checks out with some of the examples. When kidnaped, your beliefs and desires did not cause you to shoot the first victim or to fall from the airplane onto the third. You did not want to do these things; your actions were caused by the beliefs and desires of your kidnaper. So the proposed definition correctly counts your behavior in those cases as *not* being

free. It also correctly counts Hitler's invasion as being free, since the invasion was caused by Hitler's sinister beliefs and desires. Likewise, since my beliefs and desires caused me to write this chapter, and yours caused you to read it, these actions are also free, according to this definition.

But the definition's success does not last. Recall the second victim, whom you poisoned while you were hypnotized. If your kidnaper hypnotized you into *wanting* to poison the victim, then the poisoning *was* caused by your beliefs and desires. So the definition says that you were free. Yet you obviously were *not* free. So the definition is wrong. The compatibilist needs a better definition.

When you were hypnotized, you acquired beliefs and desires against your will. So maybe we should change the definition to say: *a free action is one that is caused by the person's beliefs and desires, provided that the person has freely chosen those beliefs and desires*. But this definition is circular: the word 'free' is used in its own definition. If circular definitions were kosher, we could have used a much simpler one: *a free action is one that is free*. But this is clearly unhelpful. Circular definitions are unacceptable.

(Circularity aside, it's not even clear that the modified definition is correct. I have freely decided to continue to work on this chapter. My decision was caused by my desire to complete this book. Is it *really* true that I have freely chosen this desire? I doubt it. I want to complete the book simply because that's the kind of guy I am. I didn't choose to have this desire; I just find myself having it. But this doesn't seem to undermine the fact that my decision to continue working is free.)

What about this then: *a free action is one that is caused by the person's beliefs and desires, provided that the person was not compelled by another person to have those beliefs and desires?* This new definition raises as many questions as it answers. What does the word 'compelled' mean here? (Philosophers always ask questions like this.) When you think about it, 'compelled' in its ordinary sense

means something like: 'caused so as to destroy freedom'. But then it is circular to define 'free' in terms of 'compelled', for 'compelled' is itself defined in terms of 'free'. The circularity is not so blatant as when the word 'free' itself was used in the definition, but it is circularity all the same. So the compatibilist had better not be using 'compelled' in its ordinary sense.

The definition would not be circular if 'compelled' just meant 'caused'. But then the definition wouldn't work. Recall my free decision to continue to work on this chapter. The definition requires that this decision is caused by my beliefs and desires, and it is—by my desire to complete the book. The definition further requires that this desire is not caused by any other person. But one of the causes of this desire *does* involve other people: my parents instilled diligence and a love of learning in me. So if causal involvement by another person renders a desire compelled, then my desire to continue working is compelled. We all believe and desire as we do in part because of our causal interactions with others; no one is an island. So if 'compelled' meant 'caused', the definition would imply that no one ever does anything freely. That's not what the compatibilist intends.

Another problem with the definition is that not all compulsion is by another person. A kleptomaniac compulsively desires to steal, and so steals. But he is not free; he cannot help his compulsive desires. Yet the definition counts him as free. For his stealing is caused by his beliefs and desires, and he is not compelled by *another person* to have those beliefs and desires. We could just delete 'by another person'. The definition would then read: *a free action is one that is caused by the person's beliefs and desires, provided that the person was not compelled to have those beliefs and desires*. But the problem of the meaning of 'compelled' remains. It cannot mean 'caused' (given determinism, all beliefs and desires are caused). It cannot mean 'caused so as to not destroy freedom' (that would be circular).

Let's take one final crack at a definition: *a free action is one that is caused by the person's beliefs and desires, provided that those beliefs and desires flow from 'who the person is'*. The idea of 'who the person is' needs to be explained. As a human being moves toward adulthood, she gradually develops her character, her moral beliefs and habits, her self-conception, and other qualities that give her 'an identity'. It is these qualities, which make her distinctive from a personal and moral point of view, that I am referring to when I speak of who a person is. Who an adult person is is partly a matter of upbringing and circumstance, but also partly a matter of choice. As we mature we shape ourselves; and even after reaching adulthood we continue to reflect on ourselves, and try to change if we aren't living up to our ideals. So when the definition says that the beliefs and desires must flow from who the person is, this means that the beliefs and desires must be 'in character' for that person: they must fit with the character, moral beliefs and habits, and self-conception that the person has shaped for herself over time (and continues to fine-tune). In the example at the beginning of the chapter, after you snap out of your hypnotized state, you will be inclined to protest that poisoning the second victim does not result from 'who you are'. It is out of character for you. Even though you desired to poison him at the time (because of the hypnosis), that desire conflicts with the values by which you have always lived. The case of the kleptomaniac is trickier, but here too we can say that even though her thievery is caused by her beliefs and desires, it may not be free. For suppose that even though she has always found herself desiring to steal, this desire has always been unwelcome to her. She has always tried to resist the desires—sometimes successfully, but unfortunately, sometimes not. Further, suppose that she believes that stealing is morally wrong. Given all these facts about who she is—her moral beliefs, her desire not to desire to steal, and her pattern of resisting her desires to steal—the desire

to steal does not flow from 'who she is'. The definition therefore says that her stealing is not free.

This last definition may be on the right track, but there is still work to be done. First, the definition says that your desires under hypnosis do not flow from 'who you are' because they do not match the desires you usually have; they are uncharacteristic. But many perfectly ordinary free actions are caused by uncharacteristic desires. Though I am generally a nice person, a couple of times in my life I have irritably snapped at someone. Despite being uncharacteristic for me, my snapping was obviously a free action. So my desire to snap had better count as flowing from 'who I am'. Somehow, the definition must treat my desire to snap differently from your hypnotized desire to poison—even though each desire is out of character.

Second, compare two ways of *changing* 'who one is'. Way one: someone permanently brainwashes me into becoming a horrible person. The brainwashing is so thorough that for the rest of my life I want nothing more than to harm people. At first, my actions seem out of character. But soon everyone forgets my former good qualities and regards me as a monster. Are my subsequent actions free? The question is hard, but it seems that they are at least partially unfree, since the new, evil 'who I am' results from brainwashing. Way two: I undergo *moral transformation*. After recognizing that my life is going badly and in need of reform, I change 'who I am', perhaps with the help of a spiritual leader, therapist, or other moral guide. (Moral transformation can also go from better to worse: we have all heard stories of promising young people who make the wrong decisions, fall in with the wrong crowd, and become self-destructive and immoral. The members of the 'wrong crowd' serve as negative moral 'guides'.) Unlike brainwashing, moral transformation does

not destroy free will. But in each case, one acts in accordance with 'who one is', though that has changed under the influence of other people. Somehow, the definition must treat these cases differently.

Coming up with a good compatibilist definition of freedom is no piece of cake. Then again, who ever said it should be easy? Defining anything interesting is hard. (A few paragraphs ago, we couldn't even define a measly word like 'contact'.) And look at the alternatives to compatibilism: libertarianism ('I know from my armchair that physics is incomplete!') and hard determinism ('I reject everything good about humanity!'). If our first attempts to give a compatibilist definition of freedom don't succeed, we should just keep trying.

FURTHER READING

Gary Watson's anthology *Free Will* (Oxford University Press, 1982) contains a number of interesting papers on free will. See especially the papers by Roderick Chisholm, Peter van Inwagen, A. J. Ayer, and Susan Wolf. Chisholm defends libertarianism, van Inwagen gives a careful argument against compatibilism, Ayer defends a simple form of compatibilism, and Wolf defends a sophisticated form of compatibilism and also discusses compatibilist definitions of freedom like the final one discussed in the chapter.

Timothy O'Connor, *Persons and Causes* (Oxford University Press, 2000) defends libertarianism.