

John Martin Fischer, Robert Kane, Derk Pereboom, Manuel Vargas, *Four Views on Free Will*, Blackwell Publishing, Oxford, 2007.

A Brief Introduction to Some Terms and Concepts

Basic Terms: Free Will, Moral Responsibility, and Determinism

Perhaps the three most important concepts in philosophical work on free will are *free will*, *moral responsibility*, and *determinism*.

The notion of freedom at stake in philosophical discussions is usually distinguished from a variety of other freedom concepts, including things like religious and political freedom. Usually, **free will** is also treated as distinct from several other concepts associated with human agency, such as autonomy and authenticity. As we will see in the chapters that follow, there are many different ways of thinking about the nature of free will, and there are serious disagreements about what would constitute an adequate theory of free will. Much of the tradition has taken “free will” to be a kind of power or ability to make decisions of the sort for which one can be morally responsible, but philosophers have also sometimes thought that free will might be required for a range of other things, including moral value, originality, and self-governance. Two other claims often made about free will are hotly disputed among philosophers; and authors of this volume will take different sides on these claims. One is the claim that free will requires “alternative possibilities” or the power to do otherwise, and the other is the claim that free will requires that we are the “ultimate sources” of our free actions or the ultimate sources of our wills to perform free actions.

Important to many discussions of free will is the idea of **moral responsibility**.

In the context of discussions of free will, moral responsibility is often understood as a kind of status connected to judgments and/or practices

of moral praise and blame. This meaning is distinct from another, perhaps more commonly used sense of responsibility: responsibilities as obligations (for example, when we talk about what responsibilities a parent has to a child). There are important connections between responsibility of the sort concerned with praise and blame and responsibility of the sort connected with obligations. However, philosophers writing on free will and moral responsibility are typically concerned with the former and not the latter.

Determinism is a third concept that is often important for philosophical discussions of free will. For present purposes, we can treat determinism as the thesis that at any time (at least right up to the very end) the universe has exactly one physically possible future. Something is deterministic if it has only one physically possible outcome.

It is important to bear in mind that a definition of determinism is just that – a characterization of what things would have to be like *if* things were deterministic. It does not follow that the universe is actually deterministic. Compare: “A creature is a gryphon if it has the hindquarters of a lion and the head and claws of an eagle.” Nothing about the definition of gryphon shows that there are such creatures in our universe. It simply tells us something about what sorts of things would count as gryphons. Similarly, to offer a definition of determinism does not show that the universe is deterministic. It only defines a term, and we may find that the term never properly applies to the world we live in.

When discussing these issues it is natural to wonder whether the world is deterministic. Most physicists and philosophers think that the answer is no, but the technical issues are extremely complex. Nevertheless, if we accept that the universe isn’t deterministic there are still good reasons to think about the compatibility of free will and determinism. First, it could turn out that future physicists conclude that the universe is deterministic, contrary to the contemporary consensus about at least quantum mechanics. It is notoriously difficult to predict how future science will turn out, and it might be useful

to have an answer to the question in advance of the scientific issues getting sorted out. Second, even if the universe were not fully deterministic, determinism might hold locally (either as a matter of how local spacetime is constructed, or as a matter of how the physics for non-quantum physical objects operates). Third, we could be interested in whether free will is compatible with a broadly scientific picture of the universe. Since some aspects of the universe seem deterministic and others do not, we might ask if free will is compatible with determinism as a first step to answering the more general question of whether free will is compatible with a broadly scientific picture of the universe. **Philosophical Options on the Free Will Problem**

One particularly important issue for contemporary philosophers thinking about free will is whether we could have free will in a deterministic universe. Call this issue – whether free will could exist if the universe were deterministic – **the compatibility issue**. There is a long-standing tradition of dividing up the conceptual terrain in light of the main answers to the compatibility issue. Traditionally, **incompatibilists** are those who think that free will is incompatible with the world being deterministic. **Compatibilists**, conveniently enough, are those who hold that free will is compatible with the universe being deterministic.

It is important to recognize that the compatibility issue is distinct from the issue of whether we have free will. You could be an incompatibilist, and maintain that we do have free will. Or you might be an incompatibilist and think that we lack free will. (You could even think that irrespective of how the compatibility issue is settled, there are threats to free will apart from determinism.)

In the philosophical literature, **libertarianism** is the view that we have free will and that free will is incompatible with determinism. “Libertarianism” as it is used in the context of free will is distinct from libertarianism in political philosophy. (Indeed, “libertarianism” in the free will sense is the original meaning – it was only later appropriated as the label for a view in political philosophy.) One might be a libertarian in both political and free

will senses, but you can be a libertarian about free will without being a libertarian in political philosophy. And, perhaps, you could also be a political libertarian without being a free will libertarian (although many political libertarians seem to also be free will libertarians).

Following Derk Pereboom, we will label as “**hard incompatibilism**” any view that holds that (1) incompatibilism is true and (2) we lack free will. Historically, most hard incompatibilists were what William James called **hard determinists**. (Indeed, Pereboom’s coining of the term “hard incompatibilism” reflects James’ older and narrower terminology.) Hard determinists think we lack free will *because the world is deterministic*. Contemporary hard determinists are few and far between. What is more common are views that hold that we have no free will irrespective of whether or not the world is deterministic, and views that hold that although freedom might be not be conceptually incompatible with determinism (or indeterminism, for that matter), we simply do not have it.

To summarize, then: A traditional way of dividing up the terrain concerns answers to the compatibility issue. The two main approaches are incompatibilism and compatibilism. We have been considering the incompatibilist fork, where the two main species of incompatibilism are libertarianism and hard incompatibilism. Both forms of incompatibilism have further species we have not discussed in this brief introduction.

The remaining fork of the compatibility debate is **compatibilism**. There are many varieties of compatibilism. Some compatibilists have emphasized a particular understanding of “can,” others have emphasized a kind of identification with one’s motives or values, and others emphasizing the role of responsiveness to reasons. One influential variation, however, is the view that holds that responsibility is compatible with determinism, combined with agnosticism about whether free will understood in some particular way might not be compatible with determinism. This view is **semicompatibilism**, and its most prominent defender is John Martin Fischer.

Lastly, there are views that do not neatly fit the traditional taxonomy of

incompatibilism and compatibilism. One such class of views is **revisionism**.

The core idea of revisionism is that the picture of free will and moral responsibility embedded in commonsense is in need of revision, but not abandonment.

That is, the revisionist holds that the correct account of free will and moral responsibility will depart from commonsense. As is the case with libertarianism, hard incompatibilism, and compatibilism, this view can take a variety of more specific forms.

For a different way to think about the relationship between the various views, see the grid below.

	Is commonsense thinking about free will and moral responsibility basically correct?	Is free will compatible with determinism?	Is moral responsibility compatible with determinism?	Do we have free will?
Libertarianism	Yes	No	No	Yes
Compatibilism	Yes	Yes (although semicompatibilists may say “no”)	Yes	Yes
Hard Incompatibilism	No	No	No	No
Revisionism	No	Yes, but only with revision to our self-image	Yes	Yes (or close enough)

Ilham Dilman, *Free Will - Historical and Philosophical Introduction*, Routledge, New York, 2009.

ST AUGUSTINE

86

5 Conclusion

This chapter has been confined to a discussion of the issues which exercised Augustine in his treatise *De Libero Arbitrio*. In it Augustine takes for granted that we have free will, which does not mean that we are always free in our choices, decisions and actions. Augustine thinks that it is when we are disunited in ourselves that we cannot be wholly behind what we will – that is behind our choices, decisions and actions. Thus, for instance, we may be tempted to do something which we know we should not do. As he puts it, ‘the mind commands the body and is obeyed forthwith; and the mind commands itself and is resisted’. This is something Augustine knew from personal experience. In his youth he was torn between his mother’s faith and his father’s paganism. He came to an inner unity, a unity of self and will, with his conversion to his mother’s faith: Christianity. This left him with the problem of resolving the conceptual conflicts which thinking within the framework of his faith posed for him.

At the heart of our capacity to own our intentions and be the author of our actions lies our capacity to do good and evil. So one question that exercised Augustine was how we come to evil, and how this differs from the way we come to goodness. He believed that we are free to choose between good and evil, but that we come to them in different ways. We come to evil by ‘neglecting eternal things’. We fall a prey to temptations which cloud our understanding; we act without wisdom. The problem here is this: if temptation clouds our judgement and understanding are we not acting in

slavery to our emotions and desires? Augustine argues that all the same we fall by our own will, and so are responsible for the state we are in and for what we do in such a state. We are the cause of our sins, not God. Augustine holds that by contrast we cannot rise to goodness of our own will and that we need God's grace to do so. But if so, why does God give this grace to some and not to others? It cannot be because those who have this grace are already good and so deserve it since they need His grace to rise to goodness. And if it is not based on their desert then God acts purely arbitrarily. In any case, if it is by God's grace that they rise to goodness does this not make them a by-stander to what they come to, in which case how can what they come to be goodness?

Augustine's answer is that God's grace is not a form of predestination and that it is available to all those who struggle to turn to God, including to sinners. Sinners, for instance, turn to God in the remorse they feel for their sins. Their part consists in accepting the pain of remorse, not trying to evade it, and in waiting – in being patient without any ulterior motive. This is a turning away from the self, a mourning of the evil which constitutes their sins. The time it takes to wait is important and the patience it takes to wait, and also the waiting not being directed to obtaining anything for the self, but being purely sustained by the pain of the evil done in the experience of remorse. The flip side of this same coin is love of goodness, for without it there is no pain in evil. But the person who is steeped in the pain of remorse is unaware of it, the love is hidden from him. In the course of this work of remorse which necessarily takes time and patient, selfless endurance, the love gradually becomes visible, establishes itself, takes root in the person's life, and transforms it. This is God entering his life in the form of goodness. God's grace is this transformation welcomed by the person in the humility of gratitude: 'I owe it to God. God has been good to me.' What the person receives from God is thus a gift which comes to him from a God to which he stands in a personal relation of faith.

This is the way I spell out Augustine's 'answer' to the philosophical

problem which God's grace raises for him in the way it challenges his belief in the reality of free will in human life, especially in its moral dimension. Let me point out in parenthesis that the form of predestination with which what Augustine calls the grace of God contrasts, is something which takes over a person's life if he takes the opposite attitude of patience and humility towards what befalls him in his life. This is in fact the attitude of hubris we have examined in the case of Oedipus in Sophocles' play. As Oedipus puts it in the play: 'Apollo ... laid this agony upon me, Not by his hand; I did it.' This is an instance of what Augustine means by 'we fall by our own will' and do so by 'neglecting eternal things'.

Another problem for Augustine concerns God's foreknowledge: how is it compatible with man's free will? If I know that something is the case then what I know to be the case must be the case. And if it were possible to know that something will be the case in the future in the way I know that something is the case in the present or was the case in the past then it must of necessity be the case in the future. If what is the case is an action of mine, then I cannot avoid it, in which case it cannot be the case that when the time comes I act of my own free will. The die is cast before I act. But of course no one can know the future in this way. However Christians believe that in His infinite wisdom God knows everything, sees everything. It is natural to take this 'everything' to include the future actions of men, as Augustine does, together with many other thinkers. Augustine's answer to the problem this poses for our 'belief' in the reality of free will in human life is this: 'God knows what we shall will before we have willed it. In that case what He thus knows in advance comes to pass by our willing it. Hence God's foreknowledge does not exclude human agency.' In other words he is suggesting that what God knows in advance in such cases is what we will freely choose.

I argue that this would be true only if God's foreknowledge in such cases is a prediction. But a prediction may not be fulfilled. That is if I know that p then p must be the case – of necessity, so that it cannot not be the case. Whereas, by contrast, a prediction that p does not entail its fulfilment. The

future cannot be foreclosed; what is in the future is what has not yet taken place.

‘A knows that p’ entails ‘p is true’. However that does not mean that A’s knowing p makes it true, in the sense that p is true independently of anybody knowing that p. So, Augustine would say, the fact that God knows what I shall decide in advance of my decision does not make it happen. It is I who will decide when the time comes. I argue that this does not meet Augustine’s problem with God’s foreknowledge of what I shall decide. Because if it were possible for God or anybody else to foreknow what I shall decide, when the time comes I should only go through the moves of deciding it. My decision would not be a genuine decision.

The point is that the idea of such knowledge involves a contradiction and so even God cannot have it. It does not, however, follow from this that the Christian belief that ‘God knows everything’ is false or cannot be true. What follows is that Augustine’s account of it is at fault. I take three different beliefs that may be seen to be covered by the claim that God knows everything, namely ‘God sees what’s in your heart’, ‘God sees what is in store for each one of us’, and ‘What befalls each one of us is God’s will’. In the last case one could say that God knows it as part of his will. I argue that in none of these is a knowledge attributed to God that forecloses the future – as in the case of clairvoyance, such as is attributed to Mystic Meg before the lottery draw on television. But, in any case, the trouble with clairvoyance, even if it were intelligible, is that attributed to the God of Christianity it takes away some of His spiritual status and character. A belief in a God with the power of such clairvoyance would lack the spirituality of the belief in the Christian God as we find it in the Bible.

ST THOMAS AQUINAS

106

6 Conclusion

Aquinas, like Aristotle, is nearer to what is called ‘analytic philosophy’ these days than many of the other thinkers I discuss in this book. Accordingly I give a fair amount of space in this chapter to his analysis of the relationship between reason and the will in free choice.

He thinks of human beings as having a material and a spiritual side. He thinks that if we were purely bodily beings we would have been enslaved by the necessities of nature belonging to the matter of our bodies. But we are not, and in what constitutes our spiritual side we have the source of our freedom. The will is what moves us into action by virtue of its capacity to embody our desires. As such it is a form of appetite; it is moved by needs presented to it by our desires – as it is in animals too. But these ends are presented to it in animals by instincts which belong to their species and not to the individual animal, whereas in human beings it is presented by a judgement of reason. As such he characterizes the human will as a ‘rational appetite’. Thus a frightened animal runs away; but a frightened man does not have to. His convictions or loyalties in the particular circumstances can give him reason to stand his ground and in those reasons he has the possibility of resisting the urge to run away. This is not the outcome of the opposition between two conflicting urges or inclinations, one of which is the stronger of the two. It is not the urges to which he is subject which decide the outcome; it is he who decides what to do. He is not the plaything of inclinations to which he is subject, he does not have to be: he has free will. The first two sections of the chapter provide a critical discussion of the way Aquinas represents this. They contain comparisons and contrasts between Aquinas on the one hand and Plato and Aristotle and also Augustine on the other. He differs from Plato and Augustine in thinking that to will evil is irrational and so denies that a person can will evil except by error of

judgement. Plato's conception of moral knowledge is very different from Aquinas' intellectual conception of it. For Plato to know goodness is to love it. Thus turning to evil is for him a failure of love, not of the intellect.

For Aquinas the will is directed to an object only insofar as it takes it to be good in the sense of desirable. But he allows moral desires, so that goodness itself is a good, that is desirable to the will. On Aquinas' view, I pointed out, we can will evil only because we mistakenly think it desirable. But, I ask, what makes this thought mistaken? Why should not people, given their feelings of envy for instance, find evil attractive, desirable – genuinely so? I argue that only two answers are possible: (i) Because they will find ultimately that giving themselves to evil will have undesirable consequences. This is what Aquinas calls an inferior reason and names prudence. (ii) Because they will find that eventually it will lead to the frustration of their moral desires. They will find this, Aquinas would say, if they are wise. In their wisdom they will have a superior reason for avoiding evil. In the case of the first answer I point out to

Aquinas that the wicked may prosper, and indeed they often do. A wily and clever person can get away with evil without paying any price measured in terms of non-moral criteria. In the second case we have the instance of Archelaus as referred to by Polus in the *Gorgias*. Archelaus does not care for Socrates' values and so has no moral desires to be frustrated and no moral scruples to restrain him. Socrates' response is that he stands condemned from the perspective of these values. But in his indifference to them this response means nothing to him. Aquinas has no answer to these objections. Thus I contrast Aquinas' intellectualist conception of the will according to which we will evil because we mistakenly think it desirable with Plato's view that we will evil in moral error or ignorance and that this is not a matter of the intellect but of the heart.

I point out that drawing a moral conclusion in the form of a moral judgement or decision from moral considerations is a personal matter and brings in the relation in which the individual stands to the values in which he believes and so is committed to – as an individual. Whereas, by contrast, a

deductive conclusion, for instance, is one which anyone can see must be drawn from the premises. Here his relation to the logical norms or criteria is purely intellectual. He learns them as he learns to speak and so to reason. There is no question of where he stands with regard to them and whether or not he has made them his own. In the way Aquinas gives precedence to the intellect in its relation to the

will and thinks of desire as providing it with goals which the intellect embraces he abstracts our reasons for acting, as practical reasons, from what makes a person who he is in his actions, that is himself. He considers a person to be defined by his goals. I suggest, in contrast, that a person is who he is in his convictions and commitments, in his loves and loyalties, and that his reasons engage his will as affective reasons.

Aquinas is in agreement with Augustine that to say that we rise to goodness by God's grace is not to abdicate responsibility and take a passive attitude towards one's life. In my discussion of Augustine on this question I spoke of 'waiting', 'time' and 'patience'. Patience is not the same as passivity. It is trust and self-abnegation. Aquinas asks how what comes from us and what comes from God are intertwined here and how it is that God's grace does not exclude human freedom of choice. His view is a form of compatibilism which is reminiscent of Hume's view that human freedom is compatible with causality. But his arguments are highly metaphysical. I see little connection between the metaphysical idea of God as the first cause which activates both natural and voluntary causes and the religious idea of divine grace. He is nearer to the religious sense of grace when he asks what we can achieve morally by ourselves and what we cannot achieve without God's grace. Properly understood, I believe, our dependence on God's grace does not exclude human freedom and responsible action, nor is it a form of childish dependence, a form of affective immaturity. Indeed, if we were not free, if we did not possess free will, we could not depend on God's grace, and neither could we receive it.

In the last section of the paper I consider what Aquinas has to say about

the apparent incompatibility between God's foreknowledge and human free will. If God knows now that I shall commit a murder in the future it seems to follow either (i) that I cannot avoid committing a murder in the future whatever I will now or willed in the past, or at least (ii) that I cannot avoid willing to commit a murder in the future. In (i) my will does not engage what I do; in (ii) my will is determined in advance so that what I will is not what I will.

Aquinas takes a different line from the one Augustine takes. He starts by pointing out, rightly, that God's knowledge of what I shall do in the future is not really knowledge of the future. For God is not in time and so knows everything timelessly. All right; but the question is: what does this mean? Aquinas says, not being in time God, unlike us, does not predict what for us lies in the future. He sees it from a vantage point outside time. He tries to make sense of this in terms of a spatial analogy of a person not knowing that something is moving towards him because he cannot see it from where he is. Someone else, however, can know this because from where he is it is visible. I argue that this analogy fails to shed light on God's timeless knowledge.

I propose a different analogy of a timeless truth in mathematics: the roots of a quadratic equation. What are they? The answer for me awaits the solution of the equation and, therefore, lies in the future. But whatever the answer is, it is timelessly fixed in mathematics. On this analogy what God knows is not contingent, but logically necessary, and so timelessly true for the believer. But how can this be? How can a future event be logically necessary in God's scheme of things? Isn't it as absurd as Augustine's suggestion that God's knowledge of the future may be like our knowledge of the past we can remember? As I argued in the chapter on Augustine, what is necessary and timeless is God's judgement on our actions, past, present and future, and on our life when it is completed. That is what God sees. This is the sense in which He sees into the future. In the following short passage in the *Gorgias* Callicles misunderstands this very thing and Socrates puts him right.

CALLICLES: You seem to me, Socrates, as confident that none of these things

will happen to you as if you were living in another world and were not liable to be dragged into court, possibly by some scoundrel of the vilest character.

SOCRATES: I should be a fool, Callicles, if I didn't realize that in this state anything may happen to anybody. But this at least I am sure of, that, if I am brought to trial on a charge involving any of the penalties you mention, my prosecutor will be a villain, for no honest man would prosecute an innocent party. And [since my prosecutor would have to be a villain] it would not be at all surprising if I were executed (521). He is certain for this reason that whatever anybody does to him cannot harm him.

As for particular events in the future which affect our lives for good or ill, when they are seen as God's will they are seen as unquestionable. In Christian belief, God's will being eternal what God wills is inescapable and for the believer it is, therefore, to be accepted. That is in seeing what happens as what God has willed the believer sees it as willed not now, not in the present, but in eternity or outside time. That means that he sees it as to be accepted unconditionally: nothing that has happened or may happen in the future, nothing that can happen, can change that. The necessities in question belong to the connections in which things are seen and taken within the framework of Christian belief.

The Oxford Handbook of Free Will, ed. Robert Kane, Oxford University Press, Oxford, 2012.

CHAPTER 1

INTRODUCTION:

THE CONTOURS OF

CONTEMPORARY FREE

WILL DEBATES

ROBERT KANE

There is a disputation [that will continue] till mankind is raised from the dead, between the Necessitarians and the partisans of Free Will

Jalalu'ddin Rumi, twelfth-century Persian poet

THE problem of free will and necessity (or determinism) is "perhaps the most voluminously debated of all philosophical problems," according to a recent history of philosophy. I This situation has not changed at the end of the twentieth century and the beginning of a new millennium. Indeed, debates about free will have become more voluminous in the past century, especially in the latter half of it-so much so that it has become difficult to keep up with the latest developments. This handbook was compiled as a remedy in the form of a sourcebook or guide to current work on free will and related subjects for those who wish to keep up with the latest research.

The focus of the volume is on writings of the past thirty to forty years, an era of reborn interest in traditional issues regarding free will in the context of new developments in the sciences, philosophy, and humanistic studies.

INTRODUCTION

While references are frequent throughout this volume to major thinkers of the past who have discussed free will, the emphasis is on recent research.² Many of the writers of the following essays are long-time contributors to contemporary debates about free will; others are younger scholars who are beginning to make significant contributions.

By surveying and evaluating recent writings, the hope is that their essays will serve as a guide to the latest work and a resource for future research. What is often called "the free will issue" or "the problem of free will," when viewed in historical perspective, is related to a cluster of philosophical issues-all of them to be dealt with to some degree in this volume.³ These include issues about (1) moral agency and responsibility, dignity, desert, accountability, and blameworthiness in ethics; (2) the nature and limits of human freedom, autonomy, coercion, and control in social and political theory; issues about (3) compulsion, addiction, self-control, self-deception, and weakness of will in philosophical psychology; (4) criminal liability, responsibility, and punishment in legal theory; (5) the relation of mind to body, consciousness, the nature of action,⁴ and personhood in the philosophy of mind and the cognitive and neurosciences; (6) the nature of rationality and rational choice in philosophy and social theory; (7) questions about divine foreknowledge, predestination, evil, and human freedom in theology and philosophy of religion; and (8) general metaphysical issues about necessity and possibility, determinism, time and chance, quantum reality, laws of nature, causation, and explanation in philosophy and the sciences. Obviously, this volume does not discuss every aspect of these complex issues, but it does attempt to show how contemporary debates about free will are related to them.

In the remainder of this introduction, I describe the contours of contemporary free will debates, placing them-and the essays to follow-in historical and dialectical perspective.

1. FREE WILL AND CONFLICTING VIEWS ABOUT PERSONS

The problem of free will arises when humans reach a certain higher stage of selfconsciousness about how profoundly the world may influence their behavior in ways of which they were unaware (Kane 1996: 95-6) . Various authors have described this stage of self-consciousness as the recognition of a conflict between two perspectives we may have on ourselves and our place in the universe

(P . F . Strawson 1962; Nagel 1986; Bok 1998; Blackburn 1999) . From a personal or practical standpoint, we see ourselves as free agents capable of influencing the world in various ways. Open alternatives seem to lie before us. We reason or deliberate among them and choose. We feel it is "up to us" what we choose and how we act; and this means that we could have chosen or acted otherwise-for, as Aristotle succinctly put it, "[When acting is up to us, so is not acting" (1915b: 1113b6) .This "up to us-ness" also suggests that the origins or sources of our actions are in us and not in something else over which we have no control-whether that something else is fate or God, the laws of nature, birth or upbringing, or other humans.⁵

These two features of the personal or practical standpoint are pivotal to what has traditionally been called free will: we believe we have free will when (a) it is "up to us" what we choose from an array of alternative possibilities and (b) the origin or source of our choices and actions is in us and not in anyone or anything else over which we have no control. Because of these features free will is frequently associated with other valued notions such as moral responsibility, autonomy, genuine creativity, self-control, personal worth or dignity, and genuine desert for our deeds or accomplishments (Anglin 1990; Kane 1996: ch. 6) . These two features of free will also lie behind various reactive attitudes that we naturally assume toward our behavior and that of others from a personal standpoint (P . F. Straws on 1962) . Gratitude, resentment, admiration, indignation, and other such reactive attitudes seem to depend upon the assumption that the acts for which we feel grateful, resentful, or admiring originated in the persons to whom we direct these attitudes. We believe that it was up to them whether they performed those acts or not (cf. Nathan: 1992: 46) .

But something happens to this familiar picture o f ourselves and other persons when we view ourselves from various impersonal, objective or theoretical perspectives (Nagel 1986: 110) . Perhaps we only seem to "move ourselves" in a primordial way when in fact our actions are caused by physical forces over which we have no control (Trusted 1984) . Perhaps our choices from among alternative possibilities are determined by unconscious motives and other psychological springs of

action of which we are unaware (Hospers 1958). These thoughts take many forms in human history, but in all their forms they threaten our self- image and cause a corresponding crisis in human thinking (Farrer 1967, Kenny 1978). Such is the case when we learn that much of our character and behavior is influenced by heredity or environment (Felt 1994) , or that our thoughts and behavior can be covertly influenced by social conditioning (Waller 1990; Double 1991) , or by subtle chemical imbalances of the neurotransmitters or hormones of our brains or bodies.

Free will becomes an issue when, by reflections such as these, humans realize how profoundly the world may influence them in ways previously unknown.

6 INTRODUCTION

The advent of doctrines of determinism or necessity in the history of ideas is an indication that this higher stage of awareness has been reached-which accounts for the importance of such doctrines in the long history of debates about free will (Woody 1998).⁶ Determinist or necessitarian threats to free will have taken many historical forms-fatalist, theological, physical or scientific, psychological, and logical-all of which are discussed in this volume. But a core notion runs through all these forms of determinism, which explains why these doctrines appear to threaten free will. Any event is determined, according to this core notion, just in case there are conditions (such as the decrees of fate, the foreordaining acts of God, antecedent physical causes plus laws of nature) whose joint occurrence is (logically) sufficient for the occurrence of the event: it must be the case that if these determining conditions jointly obtain, the determined event occurs.

Determination is thus a kind of conditional necessity that can be described in various ways. In the language of modal logicians, the determined event occurs in every logically possible world in which the determining conditions (e.g., antecedent physical causes plus laws of nature) obtain. In more familiar terms, the occurrence of the determined event is inevitable, given these determining conditions.

Historical doctrines of determinism refer to different kinds of determining conditions, but they all imply that every event (including every human choice or action) is determined in this general sense.⁷ One can understand as a consequence

why such doctrines pose a threat to free will. If one or another form of determinism were true, it seems that it would not be (a) "up to us" what we chose from an array of alternative possibilities, since only one alternative would be possible; and it seems that (b) the origin or source of our choices and actions would not be "in us" but in conditions, such as the decrees of fate, the foreordaining acts of God, or antecedent causes and laws, over which we had no control. But these apparent conflicts can only be the first word on a subject as difficult as this one. Many philosophers, especially in modern times, have argued that, despite intuitions to the contrary, determinism (in all of its guises) poses no threat to free will, or at least to any free will "worth wanting," as Daniel Dennett (1984) has put it.⁸

As a consequence, debates about free will in the modern era (since the seventeenth century) have been dominated by two questions, not one-the " Determinist Question": "Is determinism true?" and the "Compatibility Question": " I s free will compatible (or incompatible) with determinism?" Answers to these questions give rise to two of the major divisions in contemporary free will debates, that between determinists and indeterminists, on the one hand, and that between compatibilists and incompatibilists, on the other. Let us look at the two questions in turn.

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 7

2. THE DETERMINIST QUESTION AND MODERN SCIENCE

One may legitimately wonder why worries about determinism persist at all in the twentieth- first century, when the physical sciences-once the stronghold of determinist thinking-seem to have turned away from determinism. Modern quantum physics, according to its usual interpretations, has introduced indeterminism into the physical world, giving us a more sophisticated version of the Epicurean chance "swerve of the atoms" than the ancient philosophers could ever have conceived. We have come a long way since the beginning of the nineteenth century, when Pierre Simon, marquis de Laplace, could claim that discoveries in mechanics and astronomy unified by Newton's theory of gravitation have made it possible

to comprehend in the same analytical expressions the past and future states of the system of the world Given for an instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it-an intelligence sufficiently vast to submit these data to analysis-it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it nothing would be uncertain and the future, as the past, would be present to its eyes. (1951: 3-4)

Twentieth-century physics threatened this Laplacean or Newtonian determinist vision in several related ways. Quantum theory, according to its usual interpretations, denies that elementary particles composing the "system of the world" have exact positions and momenta that could be simultaneously known by any such intelligence (the Heisenberg Uncertainty Principle) ; and it implies that much of the behavior of elementary particles, from quantum jumps in atoms to radioactive decay, is not precisely predictable and can be explained only by probabilistic, not deterministic, laws. Moreover, the uncertainty and indeterminacy of the quantum world, according to the orthodox view of it, is not merely due to our limitations as knowers but to the nature of the physical world itself.

In the light of these indeterministic developments of twentieth-century physics, one may wonder why physical or natural determinism continues to be regarded as a serious threat to free will, as evident in many essays of the volume.⁹

Indeed, it is an important fact about the intellectual history of the twentieth century that, while universal determinism has been in retreat in the physical sciences, determinist (and compatibilist) views of human behavior have been thriving (while antideterminist and incompatibilist views of free will continue to be on the defensive) .

What accounts for these apparently paradoxical trends? There are four reasons, I believe, why indeterministic developments in modern physics have not

8 I N T R O D U C T I O N

disposed of determinist threats to free will, all of them on display in this volume.

First, there has been, and continues. to be, considerable debate about the conceptual

foundations of quantum physics and much disagreement about how it is to be interpreted. Orthodox interpretations of quantum phenomena are indeterministic, but they have not gone unchallenged. These issues about determinism and indeterminism in modern physics and related sciences, and their implications for the free will problem, are the subject of two essays of this volume, by David Hodgson and Robert Bishop.

Second, contemporary determinists about free will often concede that if modern physics is correct, the behavior of elementary particles is not always determined (see Honderich 1988; Weatherford 1991; Pereboom 1995). Yet they insist that this has little bearing on how we should think about human behavior, since quantum indeterminacy is comparatively negligible in macroscopic physical systems as large as the human brain and body. Since physical systems involving many particles and higher energies tend to be regular and predictable in their behavior for the most part, according to quantum physics itself, modern determinists argue that we can continue to regard human behavior as determined at the macroscopic level " for all practical purposes" (or "near-determined," as one of them has put it) even if microphysics should turn out to be indeterministic; and this is all that determinists need to affirm in free will debates. (For this line of argument, see the essays in this volume by Ted Honderich and Derk Pereboom; and for discussion of conflicting views about the role indeterminism might play in macroscopic systems, see the essays of Hodgson and Bishop.)

Third, one often hears the argument in contemporary free will debates that if quantum jumps or other undetermined events did sometimes have nonnegligible effects on the brain or behavior, this would be of no help to defenders of an incompatibilist free will. Such undetermined effects would be unpredictable and uncontrollable by the agents, like the unanticipated emergence of a thought or the uncontrolled jerking of an arm—just the opposite of the way we envision free and responsible actions (for example, Dennett 1984; G. Strawson 1986; Honderich 1988; Double 1991). This argument has been made in response to suggestions by prominent twentieth century scientists (such as Nobel laureates Louis De Broglie and A. H . Compton [1935] in physics and Sir John Eccles [1970] in neurophysiology)

that room might be made for free will in nature if undetermined events in the brain were somehow amplified to have large-scale effects on human choice and action. Unfortunately, such a modernized version of the Epicurean chance swerve of the atoms seems to be vulnerable to the same criticisms as its ancient counterpart. It seems that such undetermined events in the brain or body would occur spontaneously and would be more of a nuisance-or perhaps a curse, like epilepsy-than an enhancement of an agent's freedom. (For this line of argument and others about the limitations of indeterminist free will, see the essays THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 9 in this volume by Galen Strawson, Honderich, and Christopher Taylor and Daniel Dennett.)

The fourth and final reason why indeterministic developments of twentieth-century physics have not undermined determinist thinking about human behavior is perhaps the most important. While determinism has been in retreat in the physical sciences during the twentieth century, developments in sciences other than physics-in biology, neuroscience, psychology, psychiatry, social and behavior sciences-have been moving in the opposite direction. They have convinced many persons that more of our behavior is determined by causes unknown to us and beyond our control than previously believed. These scientific developments are many, but they clearly include a greatly enhanced knowledge of the influence of genetics and heredity upon human behavior (the recent mapping of the human genome is a symbolic as well as real indication of this influence, naturally arousing fears of future control of behavior by genetic manipulation) ; greater awareness of biochemical influences on the brain; the susceptibility of human moods and behavior to drugs; the advent of psychoanalysis and other theories of unconscious motivation; development of computers and intelligent machines that mimic aspects of human cognition in deterministic ways; comparative studies of animal and human behavior suggesting that much of our motivational and behavioral repertoire is a product of our evolutionary history; influences of psychological, social, and cultural conditioning upon upbringing and subsequent behavior, and so on. (The impact of such trends on contemporary free will debates is discussed

in essays by Taylor and Dennett, Paul Russell, Richard Double, Benjamin Libet, and Henrik Walter .)

In sum, there continues to be considerable debate about determinism and indeterminism in the physical world, and about the relationship of both to human behavior, while contemporary sciences other than physics provide continuing support for deterministic thinking about human behavior. Worries about determinism in human affairs therefore persist with good reason in contemporary debates about free will. 11

3. THE COMPATIBILITY QUESTION AND ARGUMENTS FOR INCOMPATIBILISM

These worries about determinism make the second pivotal question of modern free will debates, the Compatibility Question, all the more important. Is free will

10 INTRODUCTION

compatible (or incompatible) with determinism? The free will problem arose historically because it was assumed that there was some kind of conflict between free will and determinism. If it turns out, to the contrary, that determinism is no threat to free will because the two can be reconciled, then worries about determinism would be misplaced. The traditional problem of free will would not only be solved, but in a manner "dissolved," for the supposed conflict with determinism that gave rise to it in the first place would have been shown to be illusory.

Such a "dissolutionist" strategy has been the reigning strategy of modern compatibilists about free will since Thomas Hobbes in the seventeenth century.

The strategy led in the twentieth century to claims by logical positivists and others that the traditional problem of free will is a "pseudo-problem" or a "dead issue" and ought to be laid to rest. This has not happened, of course; debates about free will are more alive than ever today. But compatibilist views have had a powerful influence throughout the twentieth century. The idea that free will and determinism are compatible continues to be a majority view among philosophers and scientists because it seems to offer a simple resolution of the conflict between ordinary views of human behavior from a practical standpoint and theoretical images of human beings in the natural and social sciences. (Philosophers always

believe they have made progress when they discover something we don't have to worry about-or when they discover something we do have to worry about.)

So while the debate over the Compatibility Question has not ended, the burden of proof has shifted back to those who believe in a traditional free will that is incompatible with determinism. One cannot simply assume that if determinism is true, we would lack freedom or free will in an important sense. Arguments must now be provided to show this; and one of the interesting developments of the past thirty years is that new arguments for incompatibilism have indeed been proposed to meet the challenge. These incompatibilist arguments have in turn provoked more sophisticated compatibilist responses, and new theories on both sides of the Compatibility Question, as we shall now see.

Two features of free will were mentioned earlier that seem to imply its incompatibility with determinism- (a) it is "up to us" what we choose from an array of alternative possibilities and (b) the origin or source of our choices and actions is in us and not in anyone or anything else over which we have no control. Most modern arguments for the incompatibility of free will and determinism have proceeded from feature (a) -the requirement that an agent acted freely, or of his or her own free will, only if the agent had alternative possibilities, or could have done otherwise.¹² Let us refer to this requirement as the AP condition (for "alternative possibilities") or simply AP. (It is also sometimes called the "could have done otherwise" condition or the "avoidability" condition.)

The case for incompatibility from this AP (or "could have done otherwise ") condition has two premises:

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 11

1. The existence of alternative possibilities (or the agent's power to do otherwise) is a necessary condition for acting freely, or acting "of one's own free will."

2. Determinism is not compatible with alternative possibilities (it precludes the power to do otherwise) .

Since it follows immediately from these premises that determinism is not compatible with acting freely, or acting of one's own free will, the case for incompatibilism

(and the case against it) must focus on one or another of these premises.

In fact, there have been heated and labyrinthine debates in recent philosophy about both premises. Premise 1 is just the AP condition itself (free will requires alternative possibilities or the power to do otherwise) and it has been subjected to searching criticisms. But I shall begin with premise 2, which is usually regarded as the most crucial (and vulnerable) premise since it asserts the incompatibility of determinism with the power to do otherwise.

The most widely discussed argument in support of premise 2 in recent philosophy has been the so-called "Modal" or "Consequence" Argument for incompatibilism.

This argument was first formulated in varying ways by Carl Ginet (1966, 1980), David Wiggins (1973), Peter van Inwagen (1975, 1983), James Lamb (1977), and (in a theological form) by Nelson Pike (1965).¹⁴ Alternative formulations have since been proposed and defended by many others. Van Inwagen (1983), who offers three versions of the argument, regards these as versions of the same basic argument, which he calls the "Consequence Argument," and states informally as follows:

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born; and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us.
(p. 16)

To see the connection between this argument and premise 2, recall that something's being "up to us" implies our having the power to bring it about or not to bring it about. Given this assumption, the claims in the quote that "it is not up to us what went on before we were born" or "what the laws of nature are" imply that there is nothing we could have done to alter the past before we were born or the laws of nature. But if determinism is true, the past before we were born and the laws jointly entail our present actions. So it seems that there is nothing we can do to alter our present actions. We do not have the power to do otherwise and hence lack alternative possibilities.

This Modal or Consequence Argument for incompatibilism is the topic of

two essays in this volume, one by Tomis Kapitan, a compatibilist critic of the argument, the other by Peter van Inwagen, one of its best-known incompatibilist

12 INTRODUCTION

defenders. The first half of Kapitan's essay surveys various formulations of the Consequence Argument and criticisms made against it over the past three decades. Like many of the argument's critics, Kapitan believes its soundness depends upon how one interprets modal notions such as "power" or "ability" (to bring something about) and "could have done otherwise"; and the argument also depends on how one interprets related conditional statements about what would or might have happened if various abilities or powers had been exercised. Kapitan explores these topics in the latter part of his essay and, in the light of them, considers possible compatibilist strategies for answering the argument.

In the first half of his essay, van Inwagen restates and reaffirms a formal version of the Consequence Argument first presented in van Inwagen (1983) (which has become the most widely discussed version of the argument in the 1990s) and then defends this version against a recent objection by Thomas McKay and David Johnson (1996). In the second part of his essay, van Inwagen turns to a different topic that we will consider later-how one is to make sense of the incompatibilist or nondeterminist kind of freedom that the Consequence Argument seems to require. Van Inwagen believes that no one to date has been able to give an intelligible account of incompatibilist freedom; and he has doubts about the possibility of doing so. Yet because he also thinks the Consequence Argument is undeniably sound, he argues that we must continue to believe in an undetermined free will even if we do not know how to give an intelligible account of it.

4. CLASSICAL COMPATIBILISM:

INTERPRETATIONS OF "CAN," "POWER," AND "COULD HAVE DONE OTHERWISE"

Most compatibilists believe that the Consequence Argument and all arguments for incompatibilism can be defeated by giving a proper analysis of what it means to say that agents can (or have the power or ability to) do something; and consequently there has been much debate in recent philosophy about the meaning of

these notions. Traditionally, compatibilists themselves have defined freedom in terms of "can," "power," and "ability." To be free, most compatibilists have insisted, means in ordinary language (1) to have the power or ability to do what we will (desire or choose) to do, and this entails (2) an absence of constraints or impediments preventing us from doing what we will, desire, or choose. The constraints or impediments they have in mind include physical restraints, lack of

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 13

opportunity, duress or coercion, physical or mental impairment, and the like. You lack the freedom to meet a friend in a cafe across town if you are tied to a chair, are in a jail cell, lack transportation, someone is holding a gun to your head, or you are paralyzed. Compatibilists have typically insisted that (1) and (2) capture what freedom means in everyday life—that is, an absence of such constraints and hence the power (= ability plus opportunity) to do whatever you will or want to do.

A view that defines freedom in this way has been called "classical compatibilism" by Gary Watson (1975); and this is a useful designation. Classical compatibilists include well-known philosophers of the modern era such as Thomas Hobbes, David Hume, John Stuart Mill, and numerous twentieth-century figures (such as A. J. Ayer 1954, Moritz Schlick 1966 and Donald Davidson 1973).¹⁵ Despite differences in detail, we can say that what these classical compatibilists have in common is that they define the freedom to do something in terms of (1) and (2). What do they say about the freedom to do otherwise? It is also defined by classical compatibilists in terms of (1) and (2). You are free to do otherwise than meet your friend when you (1) have the power or ability to avoid meeting him, which entails in turn that (2) there are no constraints or impediments preventing you from avoiding the meeting (e.g., no one is forcing you at gunpoint to meet him). Of course, an absence of constraints and hence the freedom to do something does not mean you will actually do it. But for classical compatibilists it does mean that you would do it, if you wanted or desired to do it. Thus they hold that (1) and (2) entail a third feature of classical compatibilism, namely, that terms such as can, power, ability, and freedom should be given a conditional or

hypothetical analysis: (3) that an agent can (has the power, is able, is free, to) do something means that the agent would do it, if the agent wanted (or desired or chose) to do it.

Such conditional or hypothetical analyses of can, power, and freedom were not invented by compatibilists to thwart the Consequence Argument. They were invented long before that argument in an effort to represent ordinary notions of freedom. But, if conditional analyses are correct, they would effectively thwart the Consequence Argument and other arguments for the incompatibility of freedom and determinism that appeal to alternative possibilities, or the power to do otherwise. For if the power to do otherwise means only that you would have done otherwise if you had wanted or desired, it would be consistent with determinism. It might be true that you would have done otherwise if you had wanted, though it is determined that you did not in fact want otherwise. Likewise, if the power to do otherwise has only such a conditional meaning, it would not require changing the past or violating laws of nature. To say "you could have done otherwise" would only amount to the counterfactual claim that you would have done otherwise, if (contrary to fact) the past (or the laws) had been different in some way, for example, if you had wanted or desired or chosen otherwise.¹⁶

14 INTRODUCTION

So if conditional or hypothetical analyses of power and freedom favored by classical compatibilists are correct, the Consequence Argument would fail. But such analyses are themselves controversial and have also been subject to searching criticisms in contemporary philosophy. Current debates about conditional analyses- and generally about the meanings of can, power, ability and " could have done otherwise" -are surveyed and critically evaluated in another essay of this volume, by Bernard Berofsky. Berofsky's survey begins with G. E. Moore (the first important figure of the twentieth century to put forward a conditional analysis of can) but focuses primarily on the forty-year period since the publication of J. L. Austin's influential essay "Ifs and Cans" (1961) , which criticized Moore's view. Though Berofsky is a compatibilist himself, he has been among the critics of conditional or hypothetical analyses of power and ability. He thinks compatibilists

should look elsewhere if they wish to blunt the force of incompatibilist arguments and he explains his own compatibilist alternative at the end of his essay. Despite the difficulties with conditional analyses of can, power, and freedom chronicled in Berofsky's essay, conditional analyses are far from dead among contemporary philosophers. Many compatibilists continue to believe that the spirit, if not the letter, of conditional or hypothetical analyses of power and freedom can be salvaged by focusing on more sophisticated interpretations of modal and counterfactual claims about what might or might not occur in possible worlds that are similar to the actual world. (See, for example, Lehrer 1976; Lewis 1981; Falk 1981; Flint 1987; Vivhelin 1991; Audi 1993; M. White 1993; Peacocke 1999; and the essays in this volume by Kapitan, Paul Russell and Taylor and Dennett) . Yet conditional analyses of can, power, and freedom have frequently been on the defensive in recent philosophy; and many other compatibilists now try to avoid them altogether while seeking other ways to undermine arguments for incompatibilism. To these new alternative compatibilist views we now turn.

5. MORAL RESPONSIBILITY AND ALTERNATIVE POSSIBILITIES: FRANKFURT-STYLE EXAMPLES AND SEMI-COMPATIBILISM

Recall the two premises mentioned earlier, on which arguments for the incompatibility of free will and determinism have been based. (1) The existence of alternative possibilities (or the agent's power to do otherwise) is a necessary condition for acting freely, or acting "of one's own free will." (2) Determinism is not compatible with alternative possibilities (it precludes the power to do otherwise) . We have been focusing on arguments for and against premise 2. Premise 1 would appear to be less vulnerable. It simply states the AP, or alternative possibilities, condition-that free will requires alternative possibilities or could have done otherwise-a claim that persons on both sides of the compatibility debate (until recently) have tended to accept. Even most compatibilists in the past were willing to grant that freedom required the power to do otherwise. They assumed their

task was to give an analysis or interpretation of this power (conditional or otherwise) that would show it to be compatible with determinism.

This situation has changed dramatically in the past thirty years. Many contemporary compatibilists on free will and determinism-influenced in part by the difficulties of conditional analyses, but more by new arguments-would now deny that alternative possibilities or the power to do otherwise are needed for free will in the first place, in the sense in which free will is required for moral responsibility. If this denial is correct and premise 1 (the AP condition) is false, then any argument for the incompatibility of free will and determinism based on alternative possibilities-including the Consequence Argument--could not get off the ground.

Recent support for the denial of premise 1, or AP, has come from two distinct kinds of argument, which should be distinguished. The first kind of argument appeals to what David Shatz (1997) has called "character examples" 17; the second kind appeals to what have come to be known as " Frankfurt-style examples (or cases)" -named after Harry Frankfurt, who introduced the first example of this kind in an influential article "Alternate Possibilities and Moral Responsibility" (1969) .

One o f the best known arguments o f the first kind-appealing to character examples-comes from Daniel Dennett (1984) . Dennett argues that the "could have done otherwise" principle (our AP) is false, citing as a counterexample the case of Martin Luther. According to Dennett, when Luther said, " Here I stand. I can do no other," upon finally breaking with the church in Rome, he meant "that his conscience made it impossible for him to recant" (p . 133) . Suppose, says Dennett, that Luther was literally right about this: he could not then and there have done otherwise because his act was determined by his character and motives. This would not matter to Luther's free will or responsibility, Dennett argues, for "we simply do not exempt someone from blame or praise for an act because we think he could do no other" or because we think his act was determined by his character (ibid.) . In saying, "I can do no other," Luther was not renouncing free will or moral responsibility, but rather taking full responsibility fo r acting of his own fr ee

will. Dennett concludes that neither free will nor moral responsibility require "could have done otherwise"; and hence neither requires the falsity of determinism.

16 INTRODUCTION

"Character examples" of this kind have their source in David Hume's well-known observation that we cannot be held responsible (or be said to act from "our own" free wills) unless our actions are to a considerable degree determined by our characters or motives in regular ways. The obvious truth of this claim gives character examples, like Dennett's Luther example, an undeniable force. Yet there is reason to think such character examples do not provide conclusive evidence that free will does not require alternative possibilities and hence that free will is compatible with determinism. For the following response to such examples is available to incompatibilists and has been made by some of them in recent debates (see Kane 1985, 1996; van Inwagen, 1989; Shatz 1997; Ekstrom 2000). It may be true that Luther's "Here I stand" might have been a morally responsible act done "of his own free will," even if he could not have done otherwise at the time he performed it and even if his act was determined by his then-existing character and motives. But this would be true only to the extent that one could assume other things about the background of Luther's action that made him responsible or accountable for it-namely, that he was responsible by virtue of earlier choices and actions for making himself into the kind of person he now was, with this character and these motives, and that he could have done otherwise with respect to at least some of those earlier acts. If this were not so, one might argue, there would have been nothing he could have ever done to make himself different than he was-a consequence that is difficult to reconcile with the claim that he is morally responsible for being what he is. IS (The implications of this line of reasoning for issues about moral responsibility, desert, and freedom are discussed from differing points of view-in several essays of this volume, including, those of Paul Russell, Galen Strawson, Laura Ekstrom, and Robert Kane.) This argument, even if correct, would not show of course that character examples like that of Luther lack significance-far from it. Such examples seem to show that alternative possibilities need not be required for every morally responsible

act done of our own free will; and, if correct, this would be a significant implication. Yet the preceding argument also seems to show that, if we take a broader view of an agent's life history, rather than focusing on individual acts in isolation, it does not necessarily follow that free will and moral responsibility do not require alternative possibilities at all, that is, at any times, in the course of an agent's life.¹⁹ A stronger argument would be needed to show this; character examples alone do not suffice.

This leads to examples of the second kind mentioned earlier-Frankfurt-style examples or Frankfurt-style cases-which, according to many philosophers, do provide the stronger argument needed to show that alternative possibilities are not required at all for free will or moral responsibility. Examples of this kind were originally introduced by Frankfurt (1969) with the intent of undermining what he called the "Principle of Alternative Possibilities" (PAP) : " [A] person is morally responsible for what he has done only if he could have done otherwise" (p . 829) . This is our AP condition applied to "morally responsible" acts rather than to acts done "of our own free will." (The AP variant would be "a person acts of his own free will only if he could have done otherwise.") The two principles (PAP and AP) would be equivalent, if the moral responsibility at issue (in PAP) were precisely the kind that free will (in AP) is supposed to confer; and this assumption has been commonly made in free will debates. But we shall see that this assumption (linking moral responsibility and free will) has also come into question in contemporary free will debates, specifically in connection with examples of the Frankfurt type.

Frankfurt-type examples typically involve a controller who can make an agent do whatever the controller wants (perhaps by direct control over the agent's brain) . The controller will not intervene, however, if the agent is going to do on his own what the controller wants. Frankfurt argues that if the controller does not intervene because the agent performs the desired action entirely on his own, the agent can then be morally responsible (having acted on his own) -even though the agent literally could not have done otherwise (because the controller would not

have let him) . If this is so, the Principle of Alternative Possibilities, PAP, would be false: the agent would be morally responsible, though he could not in fact have done otherwise. If PAP and AP turn out to be equivalent, AP would be false as well. Neither moral responsibility nor free will (in the moral responsibilityentailing sense) would require alternative possibilities; and arguments for incompatibilism, such as the Consequence Argument, would be thwarted.

Frankfurt-style examples of this sort involving "pre-emptive" (o r "counterfactual") control have proliferated since they were first introduced. The literature on them is now enormous and has had a significant impact on contemporary free will debates. Note that Frankfurt-style examples provide extra leverage against PAP and AP that character examples do not provide. For one might go on to imagine a "global" Frankfurt controller hovering over agents throughout their entire lifetimes, so that the agents never could have done otherwise; and yet the controller never in fact intervenes because the agents always do on their own what the controller wants. Such a global controller would be a mere observer of events, never actually intervening in the agents' affairs (a mere " counterfactual intervener," in John Fischer's words) . It seems that the agents would act "on their own" throughout their entire lifetimes and would be responsible for many of their actions even though they never could have done otherwise and never had any alternative possibilities.²⁰

Contemporary debates about the implications of Frankfurt-style examples are the subject of three essays of this volume, by John Martin Fischer, Laura Ekstrom, and David Widerker. Fischer, whose prior writings have contributed as much as any contemporary philosopher to our understanding of the implications of these

18 INT RODUCTION

examples, provides a comprehensive survey of arguments about Frankfurt-style examples over the past thirty years. He considers various strategies by which critics of these examples have tried to rescue the Principle of Alternative Possibilities, PAP (or variations of it) from arguments based on such examples; and various responses to these strategies. Fischer himself is a defender of Frankfurt-style examples, who believes that moral responsibility does not require alternative possibilities

(that is, he denies PAP) . But, surprisingly, he is also an advocate of his own version of the Consequence Argument (Fischer 1994) and believes that freedom does imply alternative possibilities (that is, he affirms AP) . This view, which Fischer calls semi-compatibilism, is defended by him and also by Mark Ravizza in a number of recent writings (Fischer 1994; Ravizza 1994; Fischer and Ravizza 1998) .²¹ It amounts to the claim that moral responsibility is compatible with determinism (since it does not require the power to do otherwise) , while freedom (which does require the power to do otherwise) is not compatible with determinism. Y Laura Ekstrom and David Widerker look at Frankfurt-style examples from an opposing incompatibilist or libertarian perspective. Ekstrom, along with other incompatibilist critics of these examples (for example, van Inwagen 1978, 1983; Kane 1985, 1996a; Lamb 1993; Widerker 1995a, and b, Ginet 1996; Copp 1997; Wyma 1997) , has argued that Frankfurt-style examples do not refute every relevant form of PAP and do not show that moral responsibility is compatible with determinism (cf. Ekstrom 2000: ch. 6) . In her essay for this volume, she defends a number of objections to Frankfurt-style examples, arguing that intuitions to the effect that agents are morally responsible in such examples beg the question against those who believe free will is incompatible with determinism. Ekstrom also discusses a new Frankfurt-style example put forward in an influential article by Alfred Mele and David Robb (1998) , which was designed to answer objections to earlier Frankfurt-style cases by incompatibilist critics such as Widerker (1995a and b) and Kane (1985, 1996a) .

Widerker is the author of several articles (notably, 1995a and b) that have had a significant impact on recent debates about Frankfurt-style examples. In his essay for this volume, he defends the main theses of these articles by responding to new Frankfurt-style examples put forward in the past decade to answer his and other incompatibilist objections. Since Widerker has discussed the Mele/Robb example elsewhere (2000a) , and it is dealt with by Fischer and Ekstrom in this volume, he focuses on other recent Frankfurt-style examples designed to answer incompatibilist objections-examples suggested by Eleonore Stump (1996a, 1999b) , David Hunt (1996a) , and others. Widerker's essay also discusses some theological implications

of Frankfurt-style cases. He concludes with a general argument (called the "W-defense") designed to show that it would be unreasonable to hold an agent morally blameworthy for an action if the agent could not have avoided performing the action. In the light of this defense, Widerker argues for a version of PAP for THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 19 at least one kind of moral responsibility-moral blameworthiness-against supporters of Frankfurt, such as Fischer (1994) .

6 . BEYOND CLASSICAL COMPATIBILISM: NEW COMPATIBILIST APPROACHES TO FREEDOM AND RESPONSIBILITY

In addition to semi-compatibilism, a host of other new compatibilist views of both freedom and responsibility have been introduced in the past forty years in the attempt to answer objections to classical compatibilism. These new compatibilist theories are described and critically evaluated in two further essays of this volume, by Ishtiyaque Haji and Paul Russell.

Haji's essay deals with two broad categories of contemporary compatibilist views of freedom and responsibility, which he calls reactive attitude theories and mesh theories. Mesh theories are further divided into "hierarchical theories," "valuational theories," "reason theories," and others. Compatibilist theories of the first category-reactive attitude theories-have their roots in another seminal essay of modern free will debates, P. F. Strawson's "Freedom and Resentment" (1962) . Strawson argues that free will issues focus pivotally on the conditions required to hold persons responsible for their actions; and he argues that responsibility is constituted by persons adopting certain reactive attitudes toward themselves and others-attitudes such as resentment, admiration, gratitude, indignation, guilt, and the like. To be responsible, according to Strawson, is to be a fit subject of such attitudes. It is to be enmeshed in a "form of life" (to use Ludwig Wittgenstein's apt expression for this view) in which such reactive attitudes play a constitutive role.

Moreover, this form of life of which the reactive attitudes are constitutive is such that, according to Strawson, we would not give it up even if we found that

determinism was true, because we could not give up assessing ourselves and others in terms of the reactive attitudes if we continued to live a human form of life. So Strawson contends that the freedom and responsibility required to live a human life (whatever else they may involve) must be compatible with determinism. Freedom and responsibility do not require some mysterious indeterminist or "contra-causal" free will, as incompatibilists claim. This Strawsonian reactive attitude view has inspired considerable debate since the 1960s, which is documented in Haji's essay. It has also gained new adherents in the 1990s, one of whom is R. Jay Wallace, 20 INTRODUCTION

whose book *Responsibility and the Moral Sentiments* (1994) is the most thoroughly developed Strawsonian view in the recent literature. Wallace's view is also critically evaluated in Haji's essay.

Mesh theories, which form another influential class of new compatibilist theories, insist that the freedom required for responsibility is a function of the appropriate "mesh" or connection between agents' choices or actions, on the one hand, and their reasons or motives for acting, on the other. The most widely discussed of mesh theories are the hierarchical theories of motivation of Gerald Dworkin (1970), Harry Frankfurt (1971), Wright Neely (1974), and others. In his seminal essay "Freedom of the Will and the Concept of a Person" (1971), Frankfurt argued that persons, unlike similar animals, "have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires" (p. 7). Desires to have or not to have various first-order desires. Free will and responsibility require that we assess our first-order desires or motives and form "second-order volitions" about which of our first-order desires should move us to action.

Our "wills" - the first-order desires that move us to action - are free, according to Frankfurt, when they conform with our second-order volitions, so that we have the will (first-order desires) we want (second-order desires) to have and in that sense we "identify" with our will.

Classical compatibilism is deficient, according to hierarchical theorists such as Frankfurt, because it gives us only a theory of freedom of action (being able to do what we will) without a theory of freedom of will in terms of the conformity

of first-order motives to higher-order motives (being able, so to speak, to will what we will) . Hierarchical theories remain compatibilist, however, since they define free will in terms of a conformity (or "mesh ") between desires at different levels without requiring that desires at any level be undetermined. It does not matter, as Frankfurt puts it, how we came to have the wills we want to have, whether by a deterministic process or not. What matters is that we have the wills we want and the power to realize them in action. That is what makes us free.²³

Hierarchical views are an improvement in many ways over classical compatibilism since they provide a compatibilist account of free will as well as of free action and a richer picture of the human person. But hierarchical views are not without problems; and they have also been subjected to searching criticisms, which Haji considers in his essay. Some of the criticisms of hierarchical views have given rise to other "mesh theories" that depart from the hierarchical model in various ways. Among further mesh theories discussed in Haji's essay are "valuational" theories, such as that of Gary Watson (1975) , and Susan Wolf 's " reason view" (1990) . For Watson, the relevant mesh required for free agency is not necessarily between higher-and lower-order desires, but between an agent's "valuational system" (beliefs about what is good or ought to be done) , which has its source in THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 21 the agent's reason, and the "motivational system" (desires and other motives) , which has its source i n appetite. Watson thus revives the ancient Platonic opposition between reason and desire-arguing that freedom consists in a certain conformity of desire to reason.

Susan Wolf 's " reason view" takes this approach in yet another direction that also has ancient roots. She argues that freedom consists in being able to do the right thing for the right reasons, which requires in turn the ability to appreciate "the True and the Good." Wolf's theory thus has a stronger normative component than many other mesh theories. According to her, you are free only when you are doing the righ t thing for the right reasons. Another recent theory that shares this particular kind of normative component and has affinities to Wolf 's, though

it is original, is put forward by Phillip Pettit and Michael Smith 1996. Normative theories of freedom of somewhat different kinds have also been defended by Michael Slote (1980) and Paul Benson (1987) , among others.²⁴ Wolf 's reason view has some unusual and controversial implications that Haj i also evaluates in his essay. For example, her view contains an "asymmetry thesis" according to which we act freely when we do the right thing for the right reasons, but do not act freely when we act wrongly or otherwise fail to do the right thing for the right reasons. Finally, Haj i's essay discusses yet another compatibilist mesh theory of recent vintage put forward by Hilary Bok (1998) .

Paul Russell's essay considers a number o f other contemporary compatibilist views, some of which fit into Haji's categories, but most of which are not easily classified. Russell organizes his essay around themes from Daniel Dennett's influential compatibilist work, *Elbow Room* (1984) , and in the light of these themes considers other compatibilist views along with Dennett's, including those of Paul Benson (1987) , Martha Klein (1990) , John Fischer and Mark Ravizza (1998) ,²⁵ Robert Audi (1993) , and Kevin Magill (1997) . Through these authors and a number of other authors cited in his essay, Russell discusses a variety of topics that have been of concern to contemporary compatibilists, such as control, reflexivity, responsiveness to reasons, " moral luck," the place of character in moral evaluation, ultimacy, blameworthiness, and normative elements of freedom.

As noted, Russell's discussion of Dennett focuses on the latter's earlier *Elbow Room* (1984) . Dennett's more recent views may be seen in the essay he himself has contributed to this volume in collaboration with Christopher Taylor. Taylor and Dennett argue in defense of compatibilism that objections to compatibilist accounts of free agency are based on a flawed understanding of the relation of such notions as possibility and causation to freedom and agency; and they undertake an analysis of the relevant notions of possibility and causation to show this.

Taylor and Dennett also employ analogies to the functioning of sophisticated computers to argue that the flexibility, reflexivity, and creativity that free will requires are consistent with the hypothesis that the behavior of humans, like that

of intelligent machines, is determined. As their essay illustrates, appeals to intelligent machines and computer simulations of human cognition and behavior have come to play an increasingly important role in modern debates about free will.

7 . LIBERTARIAN OR INCOMPATIBILIST

THEORY OF FREE WILL :

THE INTELLIGIBILITY QUESTION

Let us now turn from compatibilist theories to contemporary incompatibilist views of free will. Those who believe that free will is incompatible with determinism and who also affirm that free will exists (thus denying the truth of determinism) are usually referred to as libertarians in twentieth-century writings on free will. Libertarianism in this sense is not to be confused with the political doctrine of the same name. In free will contexts, libertarians are those who believe in the existence of a traditional antedeterminist (or incompatibilist) free will, which does not necessarily commit them to political beliefs about freedom associated with political libertarianism. To avoid confusion, it would be more accurate to speak in free will debates of "free will libertarians" or "libertarians about free will" or "defenders of an incompatibilist free will" -making clear that the designation libertarian and its cognates is an abbreviation for these longer expressions-as it is assumed to be throughout this volume.

Contemporary free will libertarians must not only answer the Determinist and Compatibility questions by denying determinism and denying the compatibility of free will and determinism. They face an even more daunting task of answering a third pivotal question that has been at the heart of modern debates about free will and may be called the Intelligibility Question. Can one make sense of a freedom or free will that is incompatible with determinism? Is such an incompatibilist freedom coherent or intelligible, or is it, as many critics contend, essentially mysterious and terminally obscure?

The threat to free will from the perspective of this Intelligibility Question does not come from determinism, but from its opposite, indeterminism. If free will is not compatible with determinism, it does not seem to be compatible with indeterminism either. (One might say that the Compatibility Question is about the

first half of this ancient dilemma, while the Intelligibility Question is about the second half.) An event that is undetermined might occur or not occur, given the entire past. Thus, whether or not it actually occurs, given its past, would seem to

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 23

be a matter of chance. But chance events are not under the control of anything, hence not under the control of the agent. How then could they be free and responsible actions? Reflections such as these have led to charges that undetermined choices or actions would be "arbitrary," "capricious," "random," "irrational," "uncontrolled," "inexplicable," or merely "matters of luck or chance," not really free and responsible actions at all. It appears that the indeterminism that libertarians demand for free will would not in fact enhance freedom but would undermine it. One of the significant features of recent free will debates is that an increasing amount of attention has been given to this Intelligibility Question concerning libertarian free will. It is one thing for libertarians to put forth arguments for incompatibilism or to point out flaws in compatibilist accounts of free agency (as they have often done) ; it is quite another to give a positive account of the libertarian free agency that will show how such a free will can be reconciled with indeterminism and how it is to be related to modern views of human behavior in the natural and human sciences. Recent efforts to give positive accounts of incompatibilist or libertarian free agency are discussed in four essays of this volume, by Timothy O'Connor, Randolph Clarke, Carl Ginet, and Robert Kane.

It is instructive in reading these essays to sort recent libertarian theories into two broad categories- (1) agent-causal (or AC) theories and (2) teleological intelligibility (or TI) theories (see Kane 1989) . Agent-causal or AC theories, in O'Connor's words (1999a: 7) , posit "a sui generis form of causation by an agent that is irreducible (ontologically as well as conceptually) to event-causal processes within the agent." (I shall follow a common practice in recent writings on free will of hyphenating expressions such as "agent-cause" and "agent-causation" when talking about AC theories to indicate that a special kind of relation is intended.) AC theories have taken many historical forms. Indeed they have been the most common kind of libertarian theory until recent times. But, despite differences,

they are usually motivated by a common line of reasoning. Since undetermined free acts might occur or not occur, given all the same prior events or states of affairs involving the agents, if we to avoid saying such acts merely happen by chance, we must posit some additional causal factor over and above (and not reducible to) prior events or state of affairs to account for their occurrence. This additional causal factor would be the agent itself, which cannot be caused in turn by prior events because it is not an event and therefore not of the right type to be the effect of any cause. The agent-cause, to use Roderick Chisholm's notable expression (1982b: 30), must be a kind of "prime mover unmoved."

Libertarian theories of the teleological intelligibility (or TI) variety, by contrast, attempt to make undetermined free actions intelligible in terms of reasons or motives and intentions or purposes (hence teleologically intelligible), without postulating sui generis kinds of agency or causation that cannot be spelled out in terms of events or states of affairs involving the agent. TI theories in turn fall into two categories, depending on how they interpret the relation of reasons and in

INTRODUCTION

tentions to actions: (1) simple indeterminist (or noncausalist) TI theories maintain, again from O'Connor (1995: 7), "that free agency doesn't require there to be any sort of causal connection (even of an indeterministic variety) between the agent ['s reasons and his free actions"; while (2) causal indeterminist (or event-causal) TI theories maintain that agents cause their "free actions via [their reasons for doing so, but indeterministically" (ibid.).

As a consequence of these distinctions, recent positive libertarian theories have often been sorted into three categories: agent-causal, simple indeterminist, and causal indeterminist theories (see O'Connor 1993, 1995a; Clarke 1995, Ekstrom 1999), the latter two being TI theories. But it is interesting to note that current agent-causal or AC theories can also be divided into two categories, depending on whether they interpret the relation of reasons to actions noncausally or causally. Some AC theorists maintain that reasons for acting play an essential (probabilistic) causal role in agent-causation (for example, Clarke 1993, 1996a), while other agent-cause theorists question this (see, e.g., O'Connor 1995a, 2000).

So there are really at least four types of libertarian theory in the contemporary literature: AC theories of noncausalist and causalist kinds and TI theories of noncausalist (simple indeterminist) and causalist (causal indeterminist) kinds. All four types are discussed in the four essays of this volume on current libertarian views. Indeed, the authors of the four essays represent each of the theories: O'Connor and Clarke are AC theorists of the noncausalist and causalist kinds, respectively; while Ginet and Kane are TI theorists of the noncausalist and causalist kinds, respectively.

O'Connor's essay provides an overview of recent agent-causal (or AC) theories, explaining what motivates them through a discussion of mechanism, teleology and agency. He considers different accounts of the agent-causal relation by, among others, libertarians such as C. A. Campbell (1967) , Roderick Chisholm (1966, 1976a) , Richard Taylor (1967) , John Thorp (1980) , Michael Zimmerman (1984) , Richard Swinburne (1986) ,²⁶ Godfrey Vesey (in Flew and Vesey 1987) , Alan Donagan (1987) , William Rowe (1991) ,²⁷ Randolph Clarke, (1993, 1996a) and O'Connor himself (1995a, 2000) . O'Connor also poses the question whether agentcausal theories require a substance dualism of mind and body-as many philosophers have suspected they must, since they posit a sui generis causal relation between an agent and action that is irreducible to ordinary modes of causation. O'Connor argues that AC theories do not necessarily require substance dualism but may require some sort of strong emergence of mind from matter. He also discusses some contemporary dualist accounts of free agency in the light of this question (for example, those of John Eccles and Karl Popper 1977 and Richard Swinburne 1986, among others ²⁸) .

Clarke's essay critically examines TI theories of both the simple indeterminist and causal indeterminist kinds. (He prefers to call the former "noncausal" theories and the latter "event-causal" theories-two alternative designations that are also common in contemporary discussions .) With regard to simple indeterminist or noncausal TI theories, Clarke critically examines two representative views, those of Carl Ginet (1990) and Hugh McCann (1998) (while also citing other views of

this sort, for example, of Storrs McCall 1994 and Stewart Goetz 1997) . Clarke poses questions about how well these theories account for the relation of reasons to action and how well they are able to account for the causal role of agents in the control of free actions. His discussion of the second kind of TI theory, causal indeterminist or event-causal theories, is more complicated in that there are two distinct forms such theories have taken. The general possibility of causal indeterminist libertarian theories (as alternatives to both agent -causation and simple indeterminism) was first suggested, though not worked out, by David Wiggins (1973) , Daniel Dennett (1978) , Richard Sorabji (1980) , and Robert Nozick (1981) . Attempts to develop such theories in more detail have taken two forms in the 1980s and 1990s.

In one form, the causal indeterminism is placed earlier in the deliberative process, in the coming to mind of considerations for choice or in the formation of preferences. It is these processes that are said to be undetermined. Dennett (1978) and Kane (1985) first suggested views of this sort (which have been called "Valerian" libertarianisms²⁹), but neither unqualifiedly endorsed them.³⁰ (Indeed, Dennett, a compatibilist, suggested such a view only to criticize it.) Clarke critically examines two more recent versions of this kind of causal indeterminist view, those of Alfred Mele (1995) and Laura Ekstrom (2000) . ³¹ For causal indeterminist theories of the second kind, the indeterminism is not only placed earlier in the deliberative process, but also later, in choices themselves and in efforts of will preceding choice and action. Clarke also critically examines the most developed version of this second form of causal indeterminist theory, the view of Robert Kane (1985, 1996a) . With regard to both versions of causal indeterminism, Clarke (as he did with simple indeterminist theories) considers questions about how well they are able to account for the rationality and control agents are supposed to exercise over their free actions. He concludes his essay with an assessment of the evidence for the indeterminism in nature which these and other libertarian theories would require.

Ginet's essay focuses on the issue that distinguishes causalist from noncausalist forms of both AC or TI theories-that is, the issue of how explanations of actions

in terms of reasons or motives (beliefs, desires, intentions, and other motivating psychological attitudes) are related to causal explanations of behavior. This has been a central issue in the philosophy of action generally for the past forty years; and Ginet has been a major contributor to the debates about it. Causalists hold that reasons explanations are a form of causal explanation and require that there be a causal connection between the agent's reasons or motives and the actions

26 INTRODUCTION

they explain. Noncausalists deny that reasons explanation are a form of causal explanation and deny that such explanations require a causal connection between reasons and the actions they explain.

Ginet is a noncausalist and he defends a noncausalist account of reasons explanation in his essay, as he has in other writings (for example, Ginet 1990).

He notes that the issue has an obvious bearing on the free will problem, for if reasons can explain actions noncausally, then actions could be explained without the supposition that they are either caused or determined. But noncausalist accounts of reasons explanations have been controversial and have been criticized by compatibilists and many incompatibilists as well, ever since the publication of Donald Davidson's seminal article on the subject, "Actions, Reasons and Causes," in 1963. Ginet surveys these debates since Davidson's essay and undertakes a defense of noncausalism against its critics. In the process, he also criticizes Clarke's (and O'Connor's) agent-causal views, which Ginet, as a TI theorist and simple indeterminist, also rejects.

Kane agrees with Ginet that a special kind of agent- or nonevent causation is not needed to account for libertarian free will. As a TI theorist, he too rejects sui generis forms of causation, such as AC theories postulate. But, as a causal indeterminist, Kane disagrees with Ginet on the causalist issue: he does not think libertarians should deny that reasons explanation are a kind of causal explanation; nor need they deny that there are causal relations between reasons and actions in order to give an adequate account of incompatibilist free will. What matters is that the relevant causal relations involving reasons not always be deterministic (they may sometimes be nondeterministic or probabilistic) . In short, "undetermined"

need not mean "uncaused; and reasons, like other causes, may "incline without necessitating." If "undetermined" did mean "uncaused," one could see why libertarians might be tempted to posit (as AC theorists do) some extra kind of causation or agency to account for how free actions can be caused or produced, given that they are undetermined by events. Libertarian freedom, Kane contends, must be indeterminist, but it need not be "contra-causal."³²

But it is one thing to make such claims, and another to give an adequate account of libertarian free agency without appealing either to some special kind of agent-or nonevent causation or to the claim that reasons explanation are not causal. Kane undertakes this task in his essay and attempts to show how a TI account of free will of such a kind might be reconciled with modern conceptions of human beings in the natural and human sciences. His essay also addresses the Compatibility Question in addition to the Intelligibility Question, suggesting a novel route to incompatibilism that avoids direct appeals to alternative possibilities (AP) and the Consequence Argument, relying instead on a notion of "ultimate responsibility" (UR) .

THE CONTOURS OF CONTEMPORARY FREE W I L L DEBATES 27

8 . HARD DETERMINISM, SUCCESS OR VIEWS , AND OTHER NONSTANDARD THEORIES

Not all of those who believe that free will is incompatible with determinism affirm the existence of free will (as libertarians do) . Some incompatibilists also believe that determinism is true and so are committed to denying the existence of free will. Those who take such a stand are commonly referred to as hard determinists. The designation originated with the American philosopher and psychologist William James (1907) , who distinguished "soft" from "hard" determinists. Both groups believe that all human behavior is determined. But soft determinists are compatibilists who insist that determinism does not undermine any free will or responsibility worth having, while hard determinists are incompatibilists who take a harder line: since determinism is true, free will does not exist in a sense required for genuine responsibility, accountability, blameworthiness, or desert. Few thinkers have been willing to embrace such a hard determinist position

unqualifiedly, since it would require wholesale changes in the way we think about human relations and attitudes, how we treat criminals and criminal behavior, and so on. This has not prevented hard determinism from being unequivocally endorsed by some (for example, Baron d'Holbach in the eighteenth century and Clarence Darrow and Paul Edwards 1958 in the twentieth) , but unequivocal endorsement has been rare. The principle at work seems to have been that of the Victorian lady who exclaimed when she first heard of Darwin's theory. "Descended from the apes," she said: "let's hope it isn't true. But if it is, let's hope it does not become generally known."

Nonetheless, a core or kernel of the traditional hard determinist position has persisted into the twentieth century and continues to play a significant role in contemporary free will debates. This kernel persists in what might be called Successor Views to classical hard determinism. These Successor Views cannot strictly speaking be called hard determinist in the classical sense; and those who hold them would disown that title. But these views do contain a kernel of the classical hard determinist position that-detached from its traditional moorings-presents a powerful challenge to both compatibilist and libertarian views of free will. This kernel is defended in one form or another by four contributors to this volume, Galen Strawson, Ted Honderich, Derk Pereboom, and Saul Smilansky. None of them would accept the label of hard determinist with its classical connotations³³; and their views differ from each other. But each accepts an important kernel of traditional hard determinism; and each accordingly presents a challenging alternative to both contemporary compatibilist and libertarian views.

The kernel may be identified as follows. Classical hard determinism consists of three theses: (1) free will (in the strong sense required for ultimate responsibility

28 INTRODUCTION

and desert) is not compatible with determinism; (2) there is no free will in this strong sense because (3) all events are determined by natural causes (that is, determinism is true) . What I am calling Successor Views to classical hard determinism accept (1) and (2) , but remain noncommittal about (3) -whether universal determinism is true. Aware of developments in twentieth- century physics,

advocates of Successor Views are less confident than classical hard determinists about the truth of universal determinism; and they prefer to leave that question to the scientists. They remain convinced, however, that (1) free will-in what Galen Strawson (1986) calls the "true- responsibility entailing" sense-is incompatible with determinism, and that (2) there is, no such incompatibilist or libertarian free will.

This is the kernel-theses (1) and (2) . Though this kernel is clearly not hard " determinism" in the classical sense, because it does not unqualifiedly affirm universal determinism (3) , it still represents a pretty "hard" view since it rejects free will in the "true-responsibility entailing" sense. According to it, persons cannot be responsible or deserving for what they do in the ultimate sense assumed by believers in traditional free will. But what is especially interesting about this kernel is that it puts advocates of Successor Views who hold it at odds with both contemporary libertarian and compatibilist views. For, anyone holding the kernel holds (against compatibilism) that free will in the true- responsibility entailing sense is not compatible with determinism-thesis (1)-and (against libertarianism) that incompatibilist or libertarian free will does not exist-thesis (2) .

But why do advocates of Successor Views think libertarian free will does not exist, if they remain noncommittal about the truth of universal determinism? One answer lies in the dilemma of free will mentioned earlier: if free will is not compatible with determinism, it does not seem to be compatible with indeterminism either. In response to this dilemma, compatibilists try to refute the first horn (by arguing that free will can be reconciled with determinism) while libertarians try to refute the second horn (arguing that free will can be reconciled with indeterminism) . Many Successor Views (Pereboom's is an exception) reject both reconciliation projects and accept both horns of the dilemma. Thus, for these Successor Views, another thesis tends to play the role played by the thesis of universal determinism (3) in classical hard determinism, namely (3') : free will is not compatible with determinism and it is not compatible with indeterminism either. The first part of (3 ') is just (1) and the second half implies (2) ; so the kernel is reached in a different way without assuming the truth of determinism.

But this is not quite the whole story about Successor Views either. While they remain noncommittal about whether determinism holds universally in nature, most advocates of Successor Views do in fact believe that human behavior is regular and determined for the most part and that if indeterminism did exist in the microphysical world, its macroscopic effects on human behavior would be negligible and of no significance for free will. If the effects of indeterminism on the human brain and behavior are insignificant, as most advocates of Successor Views believe, then the traditional problems posed by determinism for free will remain unchanged, whatever one's view about the microphysical world. More important, most advocates of Successor Views insist that if indeterminism at the micro-level did sometimes have macroscopic effects on human behavior, it would be of "no help" to believers in free will, since such indeterminism would not enhance, but would only diminish, freedom and responsibility.

As noted, theses (1) and (2) put advocates of Successor Views at odds with both libertarians and compatibilists. We should not be surprised, therefore, to find the four contributors to this volume who defend Successor Views—Strawson, Honderich, Pereboom, and Smilansky—arguing against both existing libertarian and compatibilist solutions to the free will problem. In his essay, for example, Galen Strawson defends his influential argument that libertarian free will is an impossibility, whether determinism is true or not (Strawson 1986, 1994a, 2000). Strawson argues that either libertarian free will requires an impossible infinite series of backtracking free choices by which we formed ourselves or it terminates in mere chance events over which we have no control. Strawson has also been a critic of positive accounts of libertarian free agency, like those considered in the previous section. And he has criticized compatibilist views as well (including the "reactive attitude" view of his father, P. F. Strawson) for failing to give us all that we want in the way of true responsibility-entailing freedom (1986: ch. 5).

In the process of revisiting arguments against incompatibilist or libertarian accounts of freedom in this volume, Strawson focuses attention on a different line of argument for incompatibilism from those we have so far considered. Recall the

two features of traditional free will that seem to imply its incompatibility with determinism: (a) it must be "up to us" what we choose from an array of alternative possibilities and (b) the origin or source of our choices and actions must be "in us" and not in anyone or anything else over which we have no control. As noted, most arguments for incompatibilism (of which the Consequence Argument is the prime example) have proceeded from feature (a) . But another noteworthy feature of recent free will debates, especially of the past two decades, is that more attention has been directed toward the comparatively neglected condition (b) , the idea that agents must be the "ultimate sources" of their own wills and actions in a sense that entails that they, and they alone, are ultimately responsible for being the kinds of persons they are.³⁴ In his essay, Strawson subjects this so-called "ultimacy" or "ultimate responsibility" condition to searching examination. Compatibilists are criticized for failing to recognize the importance of ultimacy for free will, while libertarians are criticized for thinking it is a realizable condition. (It is instructive in this regard to read Strawson's essay in conjunction with the preceding essay in this volume by Kane. Both philosophers argue that notions of ultimate

30 INTRODUCTION

responsibility and desert are crucial for free will, but they differ about whether the freedom required for ultimate responsibility and desert is intelligible and can be realized in the actual world.³⁵)

Ted Honderich also argues for theses (1) and (2) of the kernel of classical hard determinism, as he has done in other influential writings (1988, 1993) . Honderich thinks that the traditional notion of free will requires a power of ultimate "origination" of choices or actions that is incompatible with determinism (thesis [1]) .³⁶

But he also argues that no such power of ultimate origination is possible or could exist in the real world (thesis [2]) . Of all the Successor Views canvassed in this volume, Honderich's comes closest to affirming the whole package of classical hard determinism, but even he does not quite affirm it all. The first clause of the title of his essay, "Determinism as true," suggests that he does, but it requires interpretation. The determinism Honderich affirms is a qualified "macro-determinism" of human behavior that, he says, is consistent with micro-indeterminism of the kind

that standard quantum theories postulate. Honderich's sympathies do clearly lie on the side of universal determinism, since he also questions whether quantum physics should be interpreted indeterministically and whether it will turn out to be the last word about the physical world. But he does not take a final stand on these issues or on the truth of universal determinism, as classical hard determinists do.

Honderich's main concern in the latter part of his essay (which he calls "the real problem" of free will) is how humans should react to the realization that their behavior is mostly determined and libertarian free will impossible. Some important "life-hopes" must be abandoned, he believes, if determinism is true, but many other life-hopes that matter to us can be retained.³⁷

Derk Pereboom also defends both theses of the kernel of classical hard determinism—that (1) genuine free will is incompatible with determinism and (2) libertarian free will does not exist, offering his own arguments for each thesis.

Pereboom calls the resulting position—which amounts to a rejection of both compatibilism and libertarianism—"hard incompatibilism" (see Pereboom 2001).

He candidly admits that accepting such a position involves "relinquishing our ordinary view of ourselves as blameworthy for immoral actions and praiseworthy for actions that are morally exemplary."³⁸ The question that chiefly concerns Pereboom in his essay is whether affirming such a view would have the dire consequences many people fear for a host of everyday concerns that matter to us—for example, for moral reform and education, for crime prevention, interpersonal relations, our reactive attitudes of indignation, guilt, gratitude, love and repentance, the ways we treat others, including children, and generally, for our form of life. Pereboom discusses each of these topics, arguing that the consequences of hard incompatibilism would not be as destructive as many people believe and would be compensated by benefits in the form of more humane treatment of others.

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 31

Saul Smilansky is another philosopher who holds the kernel of hard determinism without being a hard determinist in the traditional sense. But Smilansky's view is unusual among contemporary views of free will, including the other Successor

Views we have been discussing. His view is defined by two radical theses, both of which he defends in his essay for this volume and in a recent work (Smilansky 2000). The first thesis, Fundamental Dualism, states that we can and should be both incompatibilists and compatibilists about freedom and responsibility. There is no reason, Smilansky argues, why it should not be the case that certain forms of moral responsibility, desert, and blame require libertarian free will, while other forms can be sustained without it. Thus, if libertarian free will is impossible (as he believes) , there is no reason why we have to choose between hard determinism or compatibilism. We can hold a mixed view that embraces what is true in both hard determinism and compatibilism, while denying that either has the whole truth.

Smilansky's second thesis, Illusionism, is even more radical. In contrast to both Honderich and Pereboom, Smilansky thinks the consequences for humanity of widespread belief that we lack libertarian free will would be dire and destructive. Illusion on free will is therefore morally necessary, he argues (this is the thesis of Illusionism) . It is not that Smilansky thinks we need to induce illusory beliefs in people-in Brave New World fashion-but rather that such beliefs are already "in place" (for example, most people either don't question whether they have libertarian free will or, if they are compatibilists, assume they have all the freedom and responsibility they need) ; and these illusory beliefs play a largely positive social and moral role, he thinks. Recognizing that this thesis of Illusionism is likely to meet with considerable resistance (to put it mildly) , Smilansky offers a series of arguments in the latter part of his essay to show the necessity of illusion by explaining the difficulties that would prevail without it. In this connection, he considers issues concerning guilt and innocence, value and worth, remorse and integrity, and related issues.

Richard Double defends yet another nonstandard view on free will which falls entirely outside the category of Successor Views we have been considering. Double calls his view free will subjectivism. The goal of his essay is to show how such a view and the free will problem in general are related to metaethical questions about the objectivity and subjectivity of value. In the process, he also considers

how free will debates are influenced by differences in metaphilosophy-by differing views about the nature of philosophy (see Double, 1991, 1996a) . Double thinks that people generally designate by "free will" the amount and kind of freedom that is required for moral responsibility. But ascriptions of moral responsibility, he also believes, do not attribute objective properties to persons; rather they express our subjective moral and evaluative attitudes toward people and their behavior. Consequently, ascribing free will to persons is also not a matter of ascrib -

32 INTRODUCTION

ing some objective property they may or may not possess, but of expressing subjective attitudes toward them and signaling how they will be treated-for example, through reactive attitudes, verbal recrimination, praise and blame, retributive punishment or reward, and so on.

Double relates this free will subjectivism to metaethical views about value in the first part of his essay and argues that such a view best explains the persistent and seemingly irresolvable disagreements that have characterized debates about free will. In the second part of his essay, he compares and contrasts his view with those of other prominent contributors to contemporary free will debates, including several contributors to this volume: B. F. Skinner (1948, 1971) , Daniel Dennett (1984) , Bruce Waller (1990) , Galen Strawson (1986) , P. F. Strawson (1962) , Thomas Nagel (1986) , Ted Honderich (1993) , and Peter Unger (1984) .

Finally, Alfred Mele defends another nonstandard position on free will that does not fit into any of these categories. In his essay for this volume, Mele arrives at this position by a discussion of three topics that have been intertwined with contemporary debates about free will: autonomy, self-control, and weakness of will. Mele follows Aristotle in taking self-control (*enkrateia*) and weakness of will (*akrasia*) to be contraries-weakness of will being a deficiency of self-control. His essay begins with a survey of recent debates about the nature of self-control and weakness of will-debates to which Mele himself has been a significant contributor (Mele 1987, 1995) .³⁹ This survey leads him to an account of what he takes to be an "ideally self-controlled agent."

Mele then poses the question whether an ideally self- controlled agent so

conceived would necessarily also be autonomous (that is, self-governing or selflegislating) in a manner that many contemporary philosophers would associate with free will. He discusses this question in the context of the growing corpus of recent philosophical writing on the topic of autonomy (including works by Feinberg 1986; Dworkin 1988; Benn 1988; Lehrer 1997; Haworth 1986; Lindley 1986; Christman 1991; Berofsky 1995; among others) . Mele argues that autonomy- and hence also free will-requires more than self-control, including ideal self-control, and he considers the additional conditions required, showing how contemporary discussions of autonomy are intertwined with debates about free will (compare Mele 1995) . Do these further conditions for genuine autonomy require that we choose between compatibilist and incompatibilist accounts of autonomy (and hence free will) ? Mele thinks not, because one can give a "robust, satisfiable" set of adequate conditions for both compatibilist autonomy and incompatibilist autonomy. One can thus remain agnostic on the Compatibility Question regarding autonomy and free will without giving up the belief that there are autonomous human beings. He calls this view "agnostic autonomy-ism."

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 33

9 . NEUROSCIENCE AND FREE WILL

Contemporary debates about determinism in human behavior have by no means been confined to the implications of modern physics. As noted in section 2, while determinism was in retreat in the physical sciences during the twentieth century, developments in sciences other than physics-in biology, neuroscience, psychology, psychiatry, social and behavior sciences-have convinced many persons that more of their behavior is determined by causes unknown to them and beyond their control than previously believed. Of particular significance among these scientific developments is the growing knowledge of genetics and physiology, of biochemical influences on the brain, including the susceptibility of human moods and behavior to drugs and biochemical sources of psychiatric disorders. All this, coupled with advances in understanding human cognition and neural networks, has led to a growing interest during the past two decades in the implications of

the cognitive sciences and neurosciences for traditional issues about free will. This interest is reflected in two essays of this volume, by Benjamin Libet and Henrik Walter.

Benjamin Libet, professor of psychology at the University of California, San Francisco, has been a pioneer in the neurophysiological study of volition and willed action. In 1958 he began a series of groundbreaking experimental studies in human subjects (in collaboration with neurosurgeon Bertram Feinstein) relating brain activities to the appearance or production of conscious experience and willed action that have been much discussed by philosophers as well as scientists. In his essay for this volume, Libet discusses the implications of these experiments for traditional debates about free will, with special reference to the role that consciousness plays in free voluntary action.

Libet and his collaborators found that voluntary acts are preceded by a specific electrical charge in the brain (the "readiness potential") which begins several hundred milliseconds before the human subjects become consciously aware of their intention to act. This suggests that the volitional process is initiated unconsciously. While some philosophers and scientists have been tempted to conclude from Libet's findings that willed actions are determined by unconscious forces and hence that our awareness of conscious control is illusory, Libet himself has a more nuanced view of the results, which he presents in this essay. He believes there is still a role for consciousness in controlling the outcome of willed actions, since consciousness can veto the act once underway. Thus, free will is not necessarily excluded, though novel neuroscientific findings place constraints on how free will could operate and how we are to make sense of it in terms of current research on the brain.

Henrik Walter is a neuropsychiatrist and philosopher who has also written perceptively about the implications for neuroscientific research for issues about

34 INTRODUCTION

free will. His contribution to this volume consists of excerpts from his *Neurophilosophie des Willensfreiheit*, published in Germany in 1996 and in English as *Neurophilosophy of Free Will* in 2001. In the excerpts, Walter discusses the role of

the frontal cortex (and particularly the prefrontal association cortex) of the brain in the planning of actions, the selection from various options, and the organizing of behavior over time. He discusses neurological evidence for and against the claim that this region of the brain is the seat of the "will" in human beings, which some scientists have suggested (for example, Crick 1994) . Walter's view is that the functions of willing (deliberation, planning, and the like) are distributed throughout the brain, though the prefrontal areas play a pivotal role since they provide a link between the cortical regions involved in higher cognitive functioning and other parts of the brain that are the sources of emotions, feelings, and motor reactions. Walter also discusses fascinating neurological evidence suggesting that interruption of circuits involving the frontal cortex and related parts of the brain due to lesions or other deficiencies gives rise to disturbances in the feeling of agency. These include, among other examples, "alien hand syndrome" (where the patient's hand seems to have a "will of its own") , obsessive-compulsive disorders, and the "self-disorder" of schizophrenics, where patients feel that certain experiences and mental actions no longer belong to themselves or are produced outside of themselves. In the book from which these excerpts are taken, Walter defends a compatibilist view of free will. He sides with those who believe that libertarian free will cannot be made intelligible. But he thinks that justice can be done to many libertarian intuitions, if we take a neurophilosophical approach to the notions of autonomy and free will.

For further references to research in the neurosciences relevant to current debates about free will, readers should consult the bibliographical references in the essays of Libet and Walter as well as two noteworthy recent anthologies, *The Volitional Brain: Towards a Neuroscience of Free Will* (1999), edited by Libet, Anthony Freeman, and Keith Sutherland, and *Neurobiology of Decision Making* (1996), edited by A. Damasio, H. Damasio, and Y. Crick.

10 . THEOLOGICAL DETERMINISM AND FATALISM

While determinist threats to free will from the natural and human sciences have taken center stage in modern free will debates, the scientific challenges have not

been the only ones of importance to current debates. Other historically important
THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 35
determinist threats to free will of continuing interest are dealt with in two further
essays of this volume-theological determinism (in Linda Zagzebski's "Recent
Work on Divine Foreknowledge and Free Will") and fatalism, or logical determinism
(in Mark Bernstein's "Fatalism") .

The theological implications of the free will problem have been a central
preoccupation of many religious traditions, Jewish, Christian, and Muslim, among
others, as well as a central preoccupation of Western intellectual history in general,
especially since St. Augustine's seminal work, *On the Free Choice of the Will*. In
his classic poem *Paradise Lost*, John Milton describes the angels debating about
predestination and free will-wondering how they could have freely chosen to
serve or reject God, given that God had made them what they were and had
complete foreknowledge of what they would do. Milton tells us that the angels
debating this issue were in "Endless Mazes lost, " not a comforting thought for us
mortals.⁴⁰

Many theologians through the centuries have believed that God's power, omniscience,
and providence would be unacceptably compromised if one did not
affirm that all events in the universe, including human choices and actions, were
foreordained and foreknown by God. But many other theologians argued, with
equal force, that if God did in fact foreordain or foreknow all human choices and
actions, then no one could have chosen or acted differently, making it hard to see
how humans could have ultimate control over their actions in a manner that
would justify divine rewards or punishments. In such a case, the ultimate responsibility
for good or bad deeds, and hence responsibility for evil, would devolve
to God-an unacceptable consequence for traditional theists.

In the past thirty years, there has been renewed interest among philosophers
and theologians in these issues of theological determinism or theological fatalism;
and contemporary debates about them have surpassed even medieval discussions
in labyrinthine complexity. Most of the recent literature on this topic has focused
on the relation of divine foreknowledge and divine providence to human freedom

and the implications of these topics to such things as divine omniscience and power, prophecy, petitionary prayer, the relation of time and eternity, and numerous other religious concerns. Linda Zagzebski's essay is a comprehensive and illuminating guide to the contemporary literature on theological determinism by a philosopher who has herself contributed significantly to current discussions of the religious implications of free will issues.

Mark Bernstein's essay deals with yet another historically important source of determinist thinking. The earliest of determinist or necessitarian doctrines that posed a threat to free will involved fate, conceived either as an impersonal cosmic force, or in the words of the ancient philosopher Empedocles, "an oracle of Necessity, an ancient decree of the gods, eternal, sealed fast with oaths."⁴¹ Ancient concerns about fatalism were taken one step farther early in the development of Western philosophy when ancient Greek thinkers of the Megarian and (later) the

36 INTRODUCTION

Stoic schools conceived the idea that laws of logic alone might imply that human wills are fated and not free. If every proposition must be true or false (as the logical law of bivalence requires), and if this is the case for propositions about the future as well, then it seems that every future event would be fated either to occur or not to occur. If the proposition "a sea fight will occur tomorrow" were true today (to use an example made famous by Aristotle), then a sea fight could not but occur tomorrow. If the proposition were false, then a sea fight could not occur. Either way, the outcome would be necessitated by the past, together with the requirement that every proposition be true or false. This esoteric doctrine of logical fatalism or logical determinism has exercised thinkers for centuries and continues to be discussed in contemporary philosophy. Recent discussions of it are the subject of Mark Bernstein's essay.

11. PLAN OF THE VOLUME

This is a sourcebook. Each essay can be read on its own and the references within each essay direct the reader to further writings in that topic area. While the essays can be read in any order, some naturally go together and are grouped into sections, guided by the three central questions discussed in this introduction.

The Determinist Question. The first four essays consider various determinist threats to free will and the contemporary debates they have generated, from theological and fatalist doctrines that posed the earliest threats (part I: essays of Zagzebski and Bernstein) to considerations of determinism and indeterminism in the modern physical sciences (part II: essays by Hodgson and Bishop) .

The Compatibility Question. The Modal or Consequence Argument for incompatibilism, the most widely discussed recent argument for the incompatibility of free will and determinism, is discussed in part III (essays by Kapitan and van Inwagen) . Part IV considers compatibilist responses to the Consequence Argument and surveys a variety of contemporary compatibilist perspectives on freedom and responsibility (essays by Berofsky, Haji, Russell and Taylor and Dennett) . Part V considers issues about moral responsibility and alternative possibilities posed by so called "Frankfurt- style examples" and discusses doctrines of "semicompatibilism" (essays by Fischer, Ekstrom, and Widerker) .

The Intelligibility Question. The essays of Part V I consider various incompatibilist or libertarian perspectives on free will and agency and address the question of whether traditional doctrines of free will that require indeterminism can be made intelligible. They also consider issues about the nature and explanation of

THE CONTOURS OF CONTEMPORARY FREE WILL DEBATES 37
action, the relation of reasons to causes, control, rationality, and metaphysical issues about mind and body, agency and personhood (essays by O ' Connor, Clarke, Ginet and Kane) .

Part VII considers recent nonstandard views on free will, including Successor Views to hard determinism. The essays of this part also discuss a variety of further topics related to free will, including metaethical issues about the objectivity and subjectivity of value, morality and ultimate desert, criminal punishment, autonomy, self-control and weakness of will, illusion, and metaphilosophy (essays by Strawson, Honderich, Pereboom, Smilansky, Double, and Mele) .

Finally, Part VIII takes a look at all three questions-determinism, compatibility, and intelligibility-from the perspective of the neurosciences, which have begun to influence debates about free will in the past decade and are likely to

have further influence in the immediate future (essays by Libet and Walter) .

N O T E S

1. Matson 1987, vol. 1: 158.

2. I contemplated including essays on the history of free will debates in nonWestern as well as in Western cultures but found this task excessive for one volume.

Another volume with a historical or comparative focus would be a valuable project in its own right, but it also goes beyond the scope of this work. Some recent works on historical figures in the Western tradition do have relevance to contemporary debates about free will, and many are cited in various essays of this volume. Some prominent examples (by no means a complete list) include, on Aristotle, Hardie (1968) , Kenny (1969b) , Sorabji (1980) , Irwin (1980), (1988) , Fine (1981) , Broadie (1991) , Meyer (1993) ; on other ancient thinkers, Mi. White (1985) ; on Augustine, Babcock (1988) , MacDonald (1999) , Hunt (1996a, 1999) ; on Thomas Aquinas, Stump (1990) , Kenny (1993) , Loughran (1994) , Gallagher (1994) , Pink (1997) , MacDonald (1998): on medieval thinkers in the period from Aquinas to Scotus, Kent, (1995) ; on Ockham, M. Adams (1987) ; on Molina and Molinism, Freddoso (1983, 1988): on Locke, Yaffe (forthcoming) ; on Leibniz, R. Adams (1982, 1994) , Blumenfeld (1988a) , Sleight (1990, 1994) , Paull (1992) , Murray (1995) ; on Thomas Reid, Rowe (1991) , O'Connor (1994) ; on Hume, Russell (1995) ; and on Kant, Wood (1984) ; Allison (1990) .

3. Introductions to the problem of free will which appeared in the past twenty years include C. Williams (1980) , Trusted (1984) , Flew and Vesey (1987) , Thornton (1990) , Honderich (1993) , Felt (1994) , Dilman (1999) , Ekstrom (2000) . Anthologies of readings on free will in the same twenty-year period include Watson (1982) , Fischer (1986) , Fischer and Ravizza (1993), O'Connor (1995) , Mongkin and Kellner (1998) , Ekstrom (2001a) , Kane (2001) , and for theological issues about free will and divine foreknowledge, Fischer (1989) .

4. There is an enormous literature on the "philosophy of action" (or "theory of action") in the period covered by this volume, much of it relevant to issues about free

38 I N T R O D U C T I O N

will and frequently cited in this volume. Influential book-length treatments and anthologies include Goldman (1970) , von Wright (1971), Castaneda (1975) , Brand and Walton

(1976) , Tuomela (1977) , Aune (1977), Thomson (1977), Thalberg (1977), Davis (1979) , Peacocke (1979), Davidson (1980), Hornsby (1980), O' Shaughnessy (1980) , Brand (1984), Bratman (1987), Donegan (1987), Frankfurt (1988), Dretske (1988), J. Bishop (1989) , Velleman (1989), Wilson (1989), Ginet (1990), Lennon (1990) , Schick (1991), Mele (1992), (1997b), Audi (1993) , Bennett (1995), McCann (1998).

5. To be sure, this personal or practical standpoint involves other presuppositions as well, for example, that we as persons are enduring objects with a continuing identity in time. If one reduces persons to successions of causally related physical and mental events-denying they are enduring substances-as does Derek Parfit (1984) or much of the Buddhist tradition, then it may be argued that free agency and free will go by the board as well, in the normal way we understand them from a practical standpoint. (See Timothy O'Connor's essay in this volume for a discussion of some of these issues.) In an insightful article, Mark Siderits (1987) argues that the Buddhist reductionist conception of persons yields a novel compatibilist position on free will.

6. Woody (1998) provides an in-depth study of the historical and personal sources of free will issues.

7. For further discussion of definitions of determinism, see the essay by Robert C. Bishop in this volume. The sense of determinism I have in mind here that is relevant to free will is what Jordan Howard Sobel (1998) calls determinism by "ancient causes." Sobel identifies ninety varieties of determinism, but he indicates that the ones that pose problems for free will are those that imply that all events have ancient causes (i.e., they are events that, for any past time, have sufficient antecedent causes earlier than that time). Other influential works on determinism of the past thirty years include Berofsky (1971), Montague (1974), Earman (1986), and Honderich (1988) .

8. Those persons, like Dennett, who believe that freedom in every sense worth having is compatible with determinism, usually note that notions of freedom in ordinary language, such as freedom from coercion, from addiction, physical restraint, or political oppression are all consistent with determinism. An interesting variation of this compatibilist strategy is exhaustively pursued by Christine Swanton (1992), who lists scores of ordinary statements in which we talk about agents' being free and tries to devise

a "coherence theory" of freedom, as she calls it, satisfying all of them.

9. E.g., those of Hodgson, Bishop, Kapitan, Russell, Taylor and Dennett, Honderich, Pereboom, Strawson, Smilansky, Libet, and Walter.

10. Honderich (1988, 1993).

11. I have been speaking here of various kinds of scientific or causal forms of determinism (by antecedent causes and laws of nature) , which have been the focus of most of the attention in modern debates about free will. There continue to be worries about other sources of determinism of theological and logical, rather than causal, varieties in contemporary philosophy. These are discussed in a later section of this introduction and in essays in this volume by Linda Zagzebski (on theological determinism) and Mark Bernstein (on logical fatalism), chs. 2 and 3, respectively.

12. While the vast majority of contemporary arguments for incompatibilism have taken this route, there are exceptions. A few contemporary philosophers have chosen to argue for incompatibilism by way of condition (b) -the requirement of ultimate origins or sources-rather than from condition (a) for alternative possibilities. For discussion of this alternative approach to incompatibilism, see section 8 of this introduction and essays by Strawson, Kane, and Russell, chs. 19, 18, and 10, respectively. In addition, there has been a tradition of arguing that a notion of indeterminacy is logically built into the notion of a free choice (see Mackay 1967; Popper 1972), or that Godel's incompleteness theorem has implications concerning the compatibility of human reasoning and determinism (Lucas 1970; Penrose 1989). Finally, an unusual self-referential argument for incompatibilism is presented in Boyle, Grisez, and Tollefson (1976).

13. Pike's theological form of the argument is discussed in Linda Zagzebski's essay (ch. 2) in this volume which discusses the theological aspects of the free will problem. I return to Zagzebski's essay and the theological issues later in this introduction.

14. E.g., Widerker (1987), Talbott (1988), Hasker (1989) , Ginet (1990) , O'Connor (1993, 2000) , Fischer (1994), Warfield (1996) and Finch and Warfield (1998). Differences of various formulations and their implications are discussed in Kapitan's essay, ch. 6.

15. For further discussion of classical compatibilism, see the essays of Russell and

Berofsky.

16. The role and interpretation of such counterfactual conditionals have therefore been important topics of discussion in free will debates. See the essays in this volume by Kapitan, Berofsky and Taylor and Dennett, chs. 6, 8, and 11, respectively.

17. Shatz (1997) is an insightful discussion of character examples and one of the most thorough discussions of this topic in the recent literature. See also van Inwagen (1989) and Kane (1996: ch. 3) .

18. Cf. Kane (1996: ch. 3) and Shatz (1997). Also see M . Zimmerman (1989, 199 6) and Haji (1998) for this sort o f argument applied to morally blameworthy actions.

19. Robert Audi is a compatibilist who, unlike most other recent compatibilists, has emphasized the importance of responsibility for character as well as for individual actions (see, e.g., Audi 1991b for a perceptive account of responsibility for character) .

Audi's view allows him t o concede the points made i n this paragraph about a need for the power to do otherwise, but he is also one of those compatibilists who offers a compatibilist account of the power to do otherwise (1974, 1993) rather than denying its importance for responsibility.

20. One can see why Frankfurt-style examples have also had an impact on debates about theological determinism. For discussion of their theological implications, see the essays in this volume of Zagzebski and Widerker, chs. 2 and 14, respectively.

21. Semi-compatibilists must of course show what moral responsibility does require if it does not require the power to do otherwise. Fischer (1994) and Fischer and Ravizza (1998) attempt to do this in terms of notions of (what they call) guidance control and reasons-responsiveness. There is a discussion of their use of these notions in the essay by Russell in this volume.

22. The possibility of separating moral responsibility and free will in this regard is a new wrinkle in free will debates. Once one conceives the possibility of disentangling them, one might consider going the other way-regarding moral responsibility as incompatible with determinism, while freedom is compatible with determinism. Such a view has in fact a modern defender in Bruce Waller (1990), whose view is discussed in Richard Double's essay for this volume ch. 23. The separation might also be considered in theological contexts. Some incompatibilists, such as Eleonore Stump (1990, 199 6a)

and Linda Zagzebski (1991, 2000) , influenced by Frankfurt-style examples, have also en40

I N T R O D U C T I O N

tertain the view that, while free will is incompatible with causal determinism, moral responsibility is not. See Zagzebski's essay in this volume, ch. 2.

23. Two forthcoming collections of essays involving critical discussion of Frankfurt's views are Buss and Overton, eds. (n.d.) and Betzler and Guckes, eds. (2000a, and b).

24. A number of other distinguished philosophers, such as Bernard Williams (1986) and Thomas Scanlon (1988), defend normative approaches to the free will problem of compatibilist kinds that differ in certain respects from those mentioned in these paragraphs. Williams is a compatibilist, but he does not think determinism can be reconciled with our "ethical conceptual scheme . . . as it stands" (1986: 12) , which is strongly influenced by Judeo-Christian and Kantian ideas of moral duty and conscience. Incompatibilism does seem to fit this scheme, Williams thinks. But he also thinks that modern "morality" conceived in this Kantian or deontological manner is deficient and thinks we have to "recast our ethical conceptions" (p. 13) by returning to ancient Aristotelian models of the ethical life which he argues are compatible with determinism. One's views about free will therefore depend on how one conceives the ethical life. (See Richard Double's essay in this volume on this point, ch. 23) . Likewise, Scanlon's compatibilist view (1988) is related to his influential contractualist ethical theory. What gives free choice and action their special value and moral significance, Scanlon holds, is the desire to regulate one's behavior by standards that other persons could not reasonably reject in an informed and unforced contractual agreement and the desire to justify one's behavior to others in accord with such standards. What the satisfaction of these desires requires by way of freedom and responsibility is the capacity for critically reflective, rational selfgovernance (which Scanlon spells out in a manner similar to hierarchical theorists like Frankfurt). Yet another compatibilist who believes that the view one takes on the freedom/determinism issue depends upon one's conception of ethics and the good life is Richard Warner (1987).

25. Fischer and Ravizza are of course only (semi-) compatibilists about moral responsibility, not freedom. But their account of responsibility in terms of "reasonsresponsiveness"

employs strategies, examined by Russell, that throw light on compatibilist views in general.

26. O'Connor says that while Swinburne appears to be an AC theorist, some of his statements suggest otherwise.

27. Rowe (1987, 1991) develops his view as an interpretation and defense of the theory of Thomas Reid, the seventeenth-century philosopher who is regarded by many AC theorists as the modern originator and inspiration for their view.

28. Other noteworthy recent defenses of a dualist solution to the free will problem include Foster (1991) and Moreland and Rae (2000).

29. The expression "Valerian" refers to the French poet Paul Valery, who argued that freedom and creativity involved the "intelligent selection" from a number of undetermined alternatives. Dennett (1978) quotes Valery to this effect; and subsequently Mark Bernstein dubbed views of this sort "Valerian libertarianisms." See also Double (1988a) and Kane (1988).

30. Fischer (1995) contains another more recent suggestion for a view of this kind.

31. See the essays of Mele and Ekstrom in this volume, chs. 24 and 13, respectively, for defenses of their respective views against such charges.

32. Even where "causal" has the ordinary sense of "event-causal."

THE CONTROVERSY OF TEMPORARY FREE WILL DEBATES 41

33. Honderich comes closest, but even the determinism he affirms is not the universal determinism of classical hard determinism, as I explain later.

34. Examples of philosophers who have focused more attention on this so-called "ultimacy" condition include Paul Gombert (1975), Richard Sorabji (1980), Robert Nozick (1981), Robert Kane (1985, 1996a), Thomas Nagel (1986), Ted Honderich (1988), W. S. Anglin (1990), Martha Klein (1990), Derk Pereboom 2001, and Strawson himself (1986). Not all of these figures think this condition can be realized (indeed many of them argue against the possibility of its realization); but they all think it is a significant feature of free will in the traditional sense.

35. Paul Russell's essay (ch. 10) also considers issues of ultimacy, but from a compatibilist perspective.

36. The assertion in the title of his essay that "Compatibilism and Incompatibilism

are False" might suggest that Honderich is denying thesis (1) of the kernel, but this is not so. What he is claiming is that compatibilists go wrong when they insist that compatibilist freedom is the only kind worth having and incompatibilists go wrong when they claim incompatibilist freedom is the only kind worth having. Both kinds support significant "life-hopes."

37. Another philosopher who has construed the free will issue in terms of conflicts between certain of our beliefs and our deeply held desires is Nicholas Nathan. In his work *Will and World* (1992), Nathan develops a general view about the nature of philosophical issues as involving conflicts between beliefs and desires, applying such a metaphilosophical theory not only to the question of free will, but to other philosophical problems as well.

38. This volume, p. 479.

39. There is a rich recent literature on the problem of weakness of will or akrasia, much of which Mele refers to in his chapter. Significant recent works include Audi (1979), Bigelow, Dodds, and Pargetter (1990), Charlton (1988), Dunn (1987), Davidson (1970), Gosling (1990), Hurley (1993), Jackson (1984), Kenny (1975), Mortimore (1971), Pears (1985), Pugmire (1994), K. Robinson (1991), A. Rorty (1980a), Walker (1989), and Watson (1977).

40. These theological questions about free will also played a pivotal role in the Arab retrieval of ancient philosophical texts, so important in the development of late medieval Western culture. When Muslim scholars (about a century after the death of Mohammed) asked the caliphs if they could look at the ancient scrolls of the Greek philosophers hidden away in the libraries of the Middle East, one of their primary motives was to see if the pagan philosophers would provide insight into the vexing question of predestination and free will, which the Koran did not resolve.

41. Wilbur (1979: 163).

Introduction: Conceptions of Freedom

THESIS: There are several forms of liberty. Whether they are conflicting or complementary is a matter of historical circumstance.

“History, it has been said, is the field of study in which one cannot begin at the beginning.”¹ Telling a story requires decisions that could have been made differently – in particular, where to start the story. For philosophers, the story often begins with the task of clarifying the topic. For many of them this is where the story ends, too, but this is not that kind of book. This is a history of liberty, not a history of *theorizing* about liberty. Still, the topic calls for a clarifying philosophical introduction.

Histories of Liberties

What, then, does it mean to be free? Like many core philosophical concepts, the concept of liberty is not easy to pin down. Ludwig Wittgenstein observed that we talk about games with ease, even though it is not easy to say what a game is. Solitaire, football, Dungeons and Dragons, chess, and hopscotch are games. But is there anything important that they all have in common? Do the things we call games share a common essence in virtue of which the term ‘game’ properly applies? Wittgenstein thought not. We could say that all games involve forms of play, but that is only to say that

we use the word 'play' as we use the word 'game,' to refer to a range of activities whose differences are obvious but whose similarities are obscure. Part of Wittgenstein's point is that we often know *how* to use words like 'game' or 'liberty' well enough to communicate with no apparent difficulty, even when we lack a precise recipe for how to use these words. Languages evolve over centuries as tools we use to convey information and ideas about issues that actually arise in our living together. Moreover, we are constantly running into cases that are in some way novel or ambiguous, and our linguistic practices do not resolve them in advance. The historical fact about language in general is that we revise our categories as we go, as needed. The edges (if not the cores) of our categories are fluid, which is part of what makes our categories as adaptable, and thus as useful, as they are.

Part of our job as philosophers is to make our language, concepts, and questions more precise. This job is never easy. As Nietzsche once noted, only that which has no history is definable.² Liberty, however we define it, has a history. Partly because of that, defining it is indeed a serious problem. In ordinary discourse, we use the terms 'freedom' or 'liberty' to refer to various ideas; these are related in important ways, but there may not be any essence that the ideas all share. Or, if there is a shared essence, we may not be able to say exactly what it is. Perhaps the things we call freedom bear a 'family resemblance' to each other. That is, in a large family we may observe that two siblings have the same nose, while two others have the same chin or hair color. Even if no characteristic is shared by every sibling, overlapping patterns of family resemblance still mark the siblings as members of the same family.

Perhaps free speech and free trade are usefully viewed as members of the same family.³ They may turn out to have a history of going hand in hand, even though they are logically separable. Here we categorize forms of liberty as much as our present purpose requires. We don't assume there is any essence awaiting our discovery; neither do we assume otherwise.

Freedom from and freedom to

Isaiah Berlin describes two kinds of ‘freedom’ or ‘liberty.’ (Berlin uses the terms interchangeably, and so do we.) We often equate being free with an absence of constraints, impediments, or interference. For instance, the American Constitution protects freedom of speech by prohibiting Congress from passing laws that constrain speech. Berlin called this a *negative liberty*. Negative liberty connotes freedom *from* – that is, from constraints or interference. The ‘great contrast’ between it and positive liberty is that the latter has to do with self-government. The positive sense of liberty, Berlin says, is in play when the question is not “How far does government interfere with me?” but rather “Who governs me?”⁴

Berlin is often interpreted as trying to draw the following contrast. Someone is free – free *to* as opposed to free *from* – when she has a relevant capacity. So, for a bird to be free *to* fly, it must have wings and energy to take off. It is not enough that no one stops the bird. For me to be in this sense free to fly implies that I have a working aircraft at my disposal, and not merely that flight control has cleared me for takeoff. Positive freedom in this sense – freedom *to* – connotes possession of a relevant resource or capability. But, however illuminating this contrast may be (and we will come back to it), Berlin’s original aim seems to have been to draw a related but different contrast between being free from constraints, especially constraints imposed by others, and positive freedom, conceived of as exercising whatever capabilities one has in an autonomous way.⁵ In different words, the distinction between positive and negative freedom is a distinction between being free to choose goals of one’s own and being unimpeded in pursuing those goals.

Berlin sees negative (political) liberty as an absence of obstacles imposed by others.⁶ Thus he says:

If I say that I am unable to jump more than ten feet in the air, or cannot read because I am blind, or cannot understand the darker

pages of Hegel, it would be eccentric to say that I am to that degree enslaved or coerced. . . . You lack political liberty or freedom only if you are prevented from attaining a goal by human beings. Mere incapacity to attain a goal is not lack of political freedom.⁷

Berlin's negative/positive metaphor naturally suggests that the two categories are supposed, jointly, to exhaust the possibilities. Not so. Berlin says that historians have documented two hundred ways of using the term, and he is writing only about two central ones.⁸

According to human rights activist Natan Sharansky, the simple and ultimate test of whether you live in a free society boils down to the following question: can you speak your mind without fear?⁹ The locutions 'free from' and 'free to' are merely handy figures of speech, and here is a case where they can mislead.¹⁰ We would naturally speak of being free *to* speak one's mind; but what Sharansky means is being free from laws or tyrants who suppress opinions, rather than having the technological or rhetorical capabilities necessary for effectively expressing one's opinions to any given audience. Nothing stops us from being concerned about the latter, but as a matter of fact Sharansky's concern, and the concern of the framers of the US Constitution, was about freedom of speech as a negative freedom.

Benjamin Constant, writing in the wake of the French Revolution, distinguished the 'liberty of the ancients' from the 'liberty of the moderns.' Constant's idea is that the liberty of the ancients involves active participation in government, whereas the liberty of the moderns is more a matter of having control over one's own life within the rule of law.

According to Constant, a citizen of modern England, France, or America conceives of liberty as a

right to be subjected only to the laws, and to be neither arrested, detained, put to death or maltreated in any way by the arbitrary will of one or more individuals. It is the right of everyone to express their opinion, choose a profession and practice it, to dispose of property,

and even to abuse it; to come and go without permission, and without having to account for their motives or undertakings. It is everyone's right to associate with other individuals, either to discuss their interests, or to profess the religion which they and their associates prefer, or even simply to occupy their days or hours in a way which is most compatible with their inclinations or whims.¹¹

Constant continues:

Now compare this liberty with that of the ancients. The latter consisted in exercising collectively, but directly, several parts of the complete sovereignty; in deliberating, in the public square, over war and peace . . . But if this was what the ancients called liberty, they admitted as compatible with this collective freedom the complete subjection of the individual to the authority of the community.¹²

As we interpret Berlin and Constant, what Constant calls 'liberty of the ancients' is one example of what Berlin calls 'positive freedom.' Specifically, the liberty of the ancients is a collective form of freedom: people being free to deliberate and to choose their own goals. What Constant calls 'liberty of the moderns' is, by contrast, an example of what Berlin calls 'negative freedom'; it is, specifically, an individual form of freedom from external impediments.

A brief history of liberty cannot cover everything. We concentrate on liberty in its individual forms. However, we do not neglect the topic of collective freedom altogether. Our "Prehistory" chapter discusses a collective form of negative freedom, namely being free from subjugation by neighboring nations, while our "Civil Rights" chapter discusses a collective positive freedom – the empowering of subjugated groups.

Working toward an analysis of the concept of freedom is a theoretical task, but many claim that the consequences of the exercise are not merely theoretical. Constant wrote that confusing the two (that is, the ancient and the modern) conceptions of liberty was "in the all too famous days of our revolution, the cause of many an evil. France

was exhausted by useless experiments, the authors of which, irritated by their poor success, sought to force her to enjoy the good she did not want, and denied her the good which she did want.”¹³ Likewise, after distinguishing between negative and positive liberty, Isaiah Berlin went on to say that the two are not merely different conceptual categories, but rival political ideals, with conflicting implications about the proper role and scope of government.¹⁴ Right or wrong, Constant and Berlin make the debate more interesting, for their assumption that different conceptions of liberty entail different political regimes recasts the semantic issue as a political one, where the debate is not merely about how to use the language but about how to use the police.

The remainder of this chapter identifies some of the many forms of liberty. Later chapters discuss the histories of some (but not all) of these forms.¹⁵

Negative liberty

(a) Hobbes describes liberty as an “absence of external impediments.”¹⁶

By external impediments, Hobbes meant obstacles that “may oft take away part of a man’s power to do what he would; but cannot hinder him from using the power left him, according as his judgment, and reason shall dictate to him.”¹⁷ On Hobbes’s view, any obstacle whatever is an impediment to liberty.

(b) More specifically, we can define ‘liberty’ as an absence of impediments imposed *by other people*.

Suppose some obstacle leaves me unable to move my car. Perhaps a tree fell on it. Or perhaps you parked in a way that boxed me in. I am impeded either way, yet the latter is a different kind of impediment; because, if you imposed it, then we can ask whether the law should forbid your imposing such obstacles. This is what Berlin seems to have had in mind when he discussed political freedom.

(c) Even more specifically, we can define 'liberty' as an absence of obstacles *deliberately* imposed by other people.

Your unknowingly parking in my favorite parking spot is not the same as your deliberately parking there, in the knowledge that I always park there. Either act renders me unfree to park in my customary spot, yet they leave me in different situations. The accidental parking is a mere inconvenience. If I take this inconvenience personally, I am overreacting. To take my spot deliberately, though, is to send me some sort of signal – perhaps that I don't command as much respect as I thought. The accident may leave me feeling irritated in a way, but it does not leave me wondering what you are trying to tell me.

Consider another example. Your accidentally running over my bicycle is, morally, not the same as your deliberately running over it. Either act leaves me unable to ride my bicycle; but the *accident* requires you to apologize, me to accept your apology, and both of us to do the kinds of things neighbors do to make sure there are no hard feelings. (You should offer to fix the bike, at which point I should consider whether I was at fault to leave the bike in harm's way.) The deliberate assault, though, requires me to defend myself rather than to be a good neighbor. This example marks the difference between an accidental and a deliberate imposition; and now the moral overtones of the difference are unmistakable.

(d) Accordingly, we can define 'liberty' as an absence of obstacles *wrongfully* imposed by other people.

Suppose you tow my car away because I was illegally and dangerously parked, and you are a duly appointed official hired to do such things. Compare this to a situation where you tow my car away because it is a lawless town and towing my car is your way of extorting money from me for the car's return. In the second case, I am furious and perhaps terrified. In the first case, by contrast, I am irritated and disappointed, but I cannot tell myself that the obstacle to my driving away was wrongfully imposed. I decided to park in a certain way, but I cannot tell myself that my decision to park in that dangerous and

illegal way ought to have been respected. When you interfere with my deciding to park there, you are in the right, not me. So, the issue highlighted by this definition concerns obstacles that create *grounds for complaint*.

Although Locke and Hobbes had negative conceptions, each of them seeing liberty as an absence of obstacles, Locke's characterization of it is slightly moralized:

the end of law is not to abolish or restrain, but to *preserve and enlarge freedom* . . . where there is no law, there is no freedom: for liberty is, to be free from restraint and violence from others . . . freedom is not . . . a liberty for everyman to do what he lists: (for who could be free, when every other man's humour might domineer over him?) . . .¹⁸

Two centuries later, in 1881, T. H. Green would agree that freedom, rightly understood, is not a mere absence of impediments. In particular, "We do not mean a freedom that can be enjoyed by one man at a cost of a loss of freedom to others."¹⁹ Moreover,

When we measure the progress of society by its growth in freedom, we measure it by the increasing development and exercise on the whole of those powers of contributing to social good with which we believe the members of the society to be endowed; in short by the greater power on the part of the citizens as a body to make the most and best of themselves. Thus, though of course there can be no freedom among men who act not willingly but under compulsion, the mere enabling a man to do as he likes, is in itself no contribution to true freedom.²⁰

One way to understand Green is to see him as holding that real freedom has two parts: our having opportunities to perfect ourselves in cooperation with others, and our taking responsibility for pursuing such opportunities in a way that does not compromise the opportunities of others. On this reading, real freedom on Green's view is not freedom *from* responsibility but freedom *to* be responsible: responsible, namely, for pursuing our own perfection and for

making sure we do no harm in the process. Note that Green's conception of freedom is not essentially individualistic. We can freely take responsibility for ourselves as individuals, to be sure, but we can also take responsibility for ourselves as a group (as members of a family, community, church, mutual aid society, or business). So long as we are not, as Green says, under compulsion, the form of responsibility we take will be a form of freedom.²¹

On any of these conceptions, we might want to say that potential as well as actual impediments could compromise our liberty. Suppose I am a slave, but my master never tells me what to do. If as a matter of fact I live as I choose, it makes sense to say I have more freedom than other slaves have. But it also makes sense to say I am not as free as people who similarly live as they choose but have no master, because mine could at any moment start ordering me around.

On a negative conception of liberty, it will be a matter of historical contingency whether a given liberty makes for happier or healthier or wealthier lives. Negative liberties are not guaranteed to make us better off, but neither is vitamin C, or exercise – so guarantees can be beside the point. The point of negative liberty has less to do with what liberty guarantees and more to do with what liberty gives people the chance to do for themselves.

There is a difference between guaranteeing in the sense of rendering inevitable (as when government price controls render shortages inevitable) and guaranteeing in the sense of expressing a firm intention (as when government declares no child will be left behind). Clearly, *guaranteeing* something in the latter sense is no guarantee in the former sense. A legal guarantee expresses the government's commitment to produce some result, but this doesn't mean that the government will in fact produce that result. Imagine a world where, every time a government legally guarantees that people will achieve a given level of welfare, an evil demon makes sure that people do not. In that world, if you wanted people to be well off, you wouldn't want

to be issuing legal guarantees. You'd *permit* people to be badly off, because that would be their only chance to prosper in that demon-plagued world.

Of course, we don't live in a world of evil demons, so perhaps the example is irrelevant. Yet plenty of factors in this world can and do disrupt, corrupt, or pervert our best-laid plans and legal guarantees. Therefore imagining a world devoid of corruption and of unintended consequences is no more relevant than imagining a world of evil demons. We have to *check* how legal guarantees actually work in our world.

Despite the lack of guarantees, history may well reveal that respecting negative liberties has a long, successful, non-accidental track record of making for better lives. In any case, we won't settle any debate about what negative liberty does for people by conceptual analysis alone.²² We need to investigate what happens to people when negative liberties are reasonably secure, and what happens when they are not.

Positive liberty

(e) In a more positive vein, we can treat freedom as an ability to do what we want rather than as an absence of impediments. Berlin would reject this notion in an analysis of political freedom (whether positive or negative). Berlin, as has been noted, would not label the inability to jump ten feet in the air a lack of *political* freedom, unless the inability in question were caused by other people.²³ Still, even if such inabilities have no bearing on political freedom, they remain a part of the conceptual landscape of positive freedom.

Many Greeks of Plato's time conceived of freedom as a capacity for living a certain lifestyle. Having to work for a living was close to being a slave. Wage workers work under duress, or so it was thought. But if this is a contentious idea (one that Berlin and quite possibly Constant would have rejected), its undeniable grain of truth is that there is a difference between being independently wealthy and not being so. In ancient times being independently wealthy meant having

time – being able to enjoy leisure. Nowadays even average workers are independently wealthy in this sense. They work eight hours a day, not fourteen. Typically they work five days a week, not seven.

Even on this positive (in particular, capacity-oriented) view of freedom, though, it will be a contingent matter whether increasing freedom makes for better lives. Parents want better lives for their children, but does this mean that they want their child to be free to drive the family car? Not necessarily. Even when we are adults, some of our wants are self-destructive, and having the power to satisfy them won't necessarily be good for us: it will depend on the nature of these wants, or on our level of maturity. Maturity is partly a matter of being free to satisfy self-destructive wants without actually giving in to them. Maturity is, likewise, a matter of acknowledging that actions have consequences, and that the consequences of one's actions are something for which one should take responsibility.

For these sorts of reasons, Plato rejected conceiving of positive freedom as an effective license to do what we want. He worried that people could be slaves to their desires. He viewed freedom more as a capacity for effective self-governance than as a capacity to satisfy one's appetites.²⁴ Plato would have been more sympathetic to something like the following:

(f) Moralizing the previous definition, we can think of freedom as a power to do what is right.

(g) Kant distinguished between the grounds of dignity and the grounds of full moral worth.²⁵ A person's dignity consists of being at liberty to choose to respect the moral law, as per (f). By contrast, a person's full moral worth, and the fullest realization of freedom, involve not only possessing liberty in the sense of (f) but going ahead and exercising it, out of reverence for moral law. Rousseau in France, like his contemporary Kant in Prussia, spoke of freedom as "obedience to a law one prescribes for oneself."²⁶ Chapter 6 discusses what it takes to achieve something like (g) when one already has achieved freedom in the sense of (f).

(h) We can define 'freedom' as a power to do what is right, free from all temptation to do otherwise.

Conception (h) leaves room for stressing that there are internal as well as external impediments to freedom. Moreover, it explicitly incorporates both positive (freedom *to*) and negative (freedom *from*) elements.²⁷ Where Hobbes's conception often is interpreted as being more like (a), Kant's conception of what it is like to be truly, fully free (to be a *holy* will) was more like (h). This Kantian conception (which has roots in Aristotle's discussion of weakness of will and in Plato's discussion of the tyrannical soul) is moralized; it is a power to do what is right, unimpeded by contrary desire.

These last two conceptions of freedom raise a question: Is living by morality a form of servitude or of freedom? Morality demands that I do some things and refrain from doing others. Does this make me unfree? We can answer this question in more than one way; but, here too, in order to answer the question clearly, we need to be clear about how we are using the terms. In this case, the question is not empirical. We settle the question by analyzing ordinary language together with some stipulation, not by gathering social scientific observations.

For example, we may choose to place weight on ideas like the following: A person of integrity (as we understand this notion) may be unwilling to act against her principles, yet the constraints under which she lives were not arbitrarily imposed by her parents or some other authority figure. Instead, they are self-imposed. She may not dictate the content of moral law. (She cannot simply *decide* whether telling the truth is moral law.) However, she does freely choose to respect it. In a way, she seems freest of all. You may have heard the legend of Martin Luther saying before a court, "Here I stand, I can do no other." If Luther really could not have brought himself to act against his principles, does this make him unfree, or free?

Consider a poetic remark of Viktor Frankl's. "It did not really matter what we expected from life, but rather what life expected from us."²⁸ Frankl's remark implicitly suggests that we are here on this earth

for a reason. We have a mission. A typical reader finds remarks like Frankl's to be inspiring rather than stifling. Why?

(i) We note the possibility of a whole family of related conceptions according to which liberty is a power to do what we want, without self-imposed baggage (in other words being free of commitments or, more generally, free of plans, promises, hang-ups, and self-conceptions that no longer fit the person one has become).

This conception of freedom (i), unlike (h), is not moralized. John Stuart Mill's idea of a free person is that of "a person whose desires and impulses are his own – are the expression of his own nature, as it has been developed and modified by his own culture . . ." ²⁹ This conception of fully rational self-direction comes closer to what Berlin seems to have meant by positive freedom. ³⁰ Persons who are free in this sense are autonomous: legally, politically, and psychologically in a position to decide for themselves what their lives are for.

This sort of psychological freedom, and the way it relates to other forms of freedom, is the subject of our final chapter. Here we leave the discussion with a question: Insofar as freedom involves being able to do what one wants, does this mean that we can be more free simply by not wanting very much? If we are not at liberty to emigrate, can we avoid this being a limitation on our freedom simply by talking ourselves into not wanting to emigrate? ³¹ The connection between being free and getting what we want is subtle, and only partly a matter of linguistic convention.

Republican freedom

Philip Pettit says: "The negative conception of freedom as non-interference and the positive conception of freedom as self-mastery are not the only available ideals of liberty; a third alternative is the conception of freedom as non-domination which requires that no one is able to interfere on an arbitrary basis – at their pleasure . . ." ³² Pettit adds that this republican ideal of freedom as non-domination

“regarded all those who are subject to another’s arbitrary will as unfree, even if the other does not actually interfere with them; there is no interference in such a case but there is a loss of liberty. The non-interfering master remains still a master and a source of domination.”³³

We have discussed some elements of this republican conception already. The negative conception of liberty as absence of impediments wrongfully imposed by others is a related notion. Another related notion is the positive conception of liberty as self-mastery – that is, as the power to do as we will. As Pettit draws the distinction, republican freedom shares with negative freedom the idea that freedom is an absence, and with positive freedom the idea that freedom is about mastery.³⁴ Republican freedom does not, however, entail self-mastery, but merely its most crucial political prerequisite: the absence of mastery by others.

We will continue to speak of positive and negative liberty in the pages to come, but we remain aware that, as Berlin and as his critics stressed, positive versus negative liberty is a false dichotomy. As noted, negative and positive liberty can themselves be viewed as clusters of related concepts. Moreover, there are other fruitful ways of carving up the conceptual landscape, and Pettit’s is one of them.

Responsibility

Any freedom worth defending has responsibility as a corollary. (There is an existentialist conception of freedom, associated with Jean-Paul Sartre, according to which a person is responsible for literally everything, including events that occurred before the person was born.)³⁵ Societies that allow their citizens latitude for self-governance also need to trust citizens with some level of responsibility for their own conduct.

Having a lot of liberty starts to sound like having a lot of responsibility. Liberal societies give people a chance to choose a conception

of the life worth living; but such opportunity to invent ourselves is at the same time a responsibility. What makes liberty good, then? Perhaps having a lot of liberty and a lot of responsibility simply goes with being fully human. Or perhaps it is the prerequisite of living a dignified, *adult* human life – accepting, and not cowering from, the fact that a lot can go wrong when we have a lot of freedom.

In sum, making the best of one's prospects for living a good life – the kind of life one considers happy, or inspiring, or whatever – is inevitably a personal responsibility to a great extent. We operate in a framework of largely self-imposed constraints. We embrace some goals as realistic yet inspiring, and we reject others. We narrow down our options so that what we have left is a manageable set.³⁶

Picking a conception

Time-honored conceptions of liberty tend to be time-honored for a reason. They play different, often complementary roles in common-sense thinking. So we see a point in trying to narrow down the list. If the word 'liberty' is used in so many ways, this might reveal a confusion in common language. Alternatively, the differing uses might reveal something important, such as the fact that context matters. Particular historical contexts will make some aspects of freedom (social, political, economic, religious) more salient than others. Victorian-era social pressure is one context. Slavery is another. The Protestant Reformation is another. Freedom from the risk of polio is another. President Roosevelt's call for moving toward a society that achieves "freedom from want" is another. Seeking freedom from the stress of overcommitment is yet another. There is value in trying to identify the essence that these various freedoms all share, but there is also much to gain from acknowledging the differences. Each of these freedoms is something people have for good reason struggled to secure. One is concerned with liberty in all such contexts, but

the concerns one aims to mark by using the word are only related, not identical.

Although these various conceptions of liberty are sometimes treated as competitors, we often see them as being complementary. Some theorists see a minimal set of protected negative liberties as being all we need in order to launch a society that, over generations, produces explosive gains in positive liberty. Other theorists seek guarantees and do not find them in a system of mere negative liberty. I might be free from interference by government, free from oppression by a rigid caste system, and so on, yet I might remain unable to do much because of lack of wealth. Negative freedom, some would say, is the freedom to be poor, to sleep on a public sidewalk, and the like.

We would not want to let debate about negative freedom's real effects degenerate into a terminological dispute. Perhaps, as a matter of fact, negative freedom often leads to poverty. How would we know? Manipulating definitions would not tell us much. The point of defining terms is not to cut off debate about whether negative freedom leads to poverty – but to facilitate debate: not to *stipulate* that negative liberty leads by definition to prosperity, but to be precise enough to make a question answerable. For example, where there are fewer obstacles to seeking employment of one's choice (fewer migration restrictions, fewer licensing or union membership requirements), are there fewer unemployed people? If so, then we can infer (not in the way a logician deduces but rather in the way a scientist guardedly infers causal connections from empirical regularities) that negative freedom is positively liberating in that particular way. We can ask well-defined questions about the consequences of specific forms of negative freedom, such as freedom from trade restrictions or from state-mandated religion. If we can document trends, making the debate less about whether a trend is real and more about why the world sometimes departs from it, we have made progress in lowering the barriers of understanding – which is what we realistically hope for from philosophy.

Isaiah Berlin and many classical liberals are suspicious of ‘positive liberty,’ thinking that acknowledging its value can be misinterpreted as licensing socialism – or, more generally, as licensing bureaucrats to force us to be “free.” Nevertheless, simply acknowledging positive liberty as a valuable species of the genus liberty does not commit us to any particular view about what regime promotes it best. We share Berlin’s concern about giving governments a license to do whatever it takes in order to promote positive liberty. (In the real world, to give government officials the power to do *x* is to *hope* that officials will use it to do *x*, knowing that, no matter who actually ends up holding such office, the person in question will duly pay lip service to doing *x*, then will use the power for purposes of her own.) None of the conceptions of freedom discussed earlier entails that it should be the government’s job to secure that kind of freedom. Defining terms cannot settle a government’s proper role as protector or promoter of particular liberties. We must examine historical, sociological, and economic evidence to see what actually happens when people rely on any institution, including a government, to play a given role.

At the risk of oversimplification, each of the first four chapters starts with negative liberty, treated as freedom from external impediments deliberately imposed. We argue that, in various ways, securing this freedom has a history of enabling people to achieve positive freedom. That is to say, in (negatively) free countries, people generally have more *real choice*. What is real choice? Charles Taylor (1979) distinguishes between negative freedom as an opportunity concept (that is, as a state of having options) and positive freedom as an exercise concept (that is, as a state of having exercised one’s options in such a way as to achieve self-realization). In speaking of “real choice,” we intend to encompass both an opportunity and an exercise concept: to have real choice is to have options together with the capacities to exercise these options successfully.³⁷ Chapter 5 (on civil rights) might be seen mainly as starting from a republican conception of freedom as non-domination and working toward the conclusion that

non-domination, too, has a history of fostering positive freedom in the sense just defined. Chapter 6 (on psychological freedom) turns to an awkward, by no means merely theoretical, puzzle lying at the core of (one sense of) positive freedom. The puzzle: we could be, as they say, our own worst enemy. We might have an enviable set of options, yet we might want too much, or too little, or know too little about what we truly want, to be able to handle our world of options as it is. In sum, not all impediments to freedom are external.

Institutions

This book's overall purpose is to tell a story of liberty, and to tell it briefly, not to *argue* for liberty or for any particular way of defining the term 'liberty.' The claims we make here are not ideologically neutral, but neither are they ideologically loaded. They don't tell you how to vote. They don't settle debates between socialists and capitalists, communitarians and liberals, modern liberals and classical liberals.

When we ask, "What sorts of institutions promote liberty?", the only answer that does not admit of counterexamples and cannot be accused of oversimplification is: it depends. Naturally, we want better answers than that. So we will risk saying more, knowing that, in a world of general trends and exceptions to the rule, whatever we go on to say will be subject to counterexamples, will run the risk of oversimplification, and of needing correction in light of new information, as historical research always is.

We focus on western Europe and on the English-speaking world partly because this is where we and most of our readers are, but more importantly because this is where most of our information is. We wish we knew more about the history of liberty in non-western cultures, but we incorporate what we know as best we can. We find the history of liberty fascinating regardless of what culture it comes from. We hope that readers will feel the same way.

The role of government

Whatever Constant and Berlin meant to be doing by warning that different conceptions of liberty are also different political ideals, the takeaway value is not that different conceptions of liberty automatically translate into different government mandates, but rather that naïve conceptions of what is possible or what is at stake can leave people vastly overconfident about the likelihood of achieving a result by creating the political power to obtain that result by force.

Constant and Berlin would have agreed. Constant, for example, writing after the French Revolution, was aware of the potentially tyrannical nature of government. Even a democratic government needs to be sharply limited by a constitution.

The holders of authority are . . . ready to spare us all sort of troubles, except those of obeying and paying! They will say to us: "what, in the end, is the aim of your efforts, the motive of your labours, the object of all your hopes? Is it not happiness? Well, leave this happiness to us and we shall give it to you." No, Sirs, we must not leave it to them. No matter how touching such a tender commitment may be, let us ask the authorities to keep within their limits. Let them confine themselves to being just. We shall assume the responsibility of being happy for ourselves.³⁸

Thus, only so much of practical consequence hangs on how we define our terms. By contrast, a lot depends on what we want out of our lives and our communities – whether we want to be free to stand or fall by our own merit, or whether we want to be free from the risks and costs that go with personal responsibility. A lot depends on how big a hammer we want our government to be, which in part turns on how confident we are that the hammer will be used by, rather than against, our children.³⁹ In sum, it matters that we understand: a) what our options are; b) that we did choose among them; and c) that in many ways large and small, as individuals and as societies, we will face the choice again.

Beyond ideal theory

What makes a hammer good? One answer is: hammers are good insofar as they serve a purpose.⁴⁰

To say what makes a hammer good, we wouldn't normally think we'd need to know what an *ideal* hammer is. Neither would we need to learn what sort of hammer would be useful to angels or giants, as intermediate steps in evaluating our own hammers as pale imitations of tools fit for the hands of angels or giants. All that matters for our purposes is whether the hammer works for us just as we are.

Institutions, as we see them, are like hammers. Institutions are tools. Institutions that help real people live together in peace and prosperity are to that extent worth preserving. Institutions that do the opposite, whatever they might symbolize or however blameless or famous their creators may have been, leave us without a reason to preserve them.

John Rawls, the author of *A Theory of Justice*, has another approach to evaluating institutions. He thinks that political philosophers who study institutions ought to start with what he calls 'ideal theory,' which involves asking "what kind of regime . . . would be right and just, could it be effectively and workably maintained?"⁴¹ He assumes that people are motivated by a sense of justice, fully understand what they are doing, and will accordingly be competent and uncorrupted. Rawls then asks which institutions are the most just under those conditions. His question is probably worth answering at some point. We, likewise, might want to know at some point what sort of hammer would be best for a giant or an angel. We learn something about hammers when we answer that question. But a theory of institutions, for us, is about how people achieve peace, a rising standard of living, and mutual respect in our actual complex world. Thus Rawls's question has little (possibly nothing) to do with ours: "Which institutions have a history of enabling people to achieve peace, a rising standard of living, and mutual respect under real-world conditions, where the willingness and the ability to comply are variable, and very much depend on what citizens are being asked to do?"⁴²

One misuse of ideal theory would result from inferring that, if some institutions are best under 'ideal' conditions, then our real-world institutions ought to come as close as possible to those institutions. Not so. Different conditions call for different tools. Ideal conditions might call for a wrench where non-ideal conditions call for a hammer. In other words, ideal theory is like designing cars on the assumption that they'll never encounter slippery pavement, or will never be driven by bad drivers. If we had no such worries, we might not bother installing air bags. Here and now, though, we have compelling practical reasons not to build cars like that. Analogously, if power didn't corrupt, if people were omniscient and unfailingly altruistic, we might want to entrust government with a great deal of power. But, if people are corruptible, if power is above all what corrupts, if people's generosity depends very much on circumstances, and if those who hold power never know exactly what will happen when they implement a given policy, the kind of government we have reason to favor might be far more limited.

Someone trying to do non-ideal theory would ask: How much power do we want our leaders to have, given that they are going to be about as good as actual leaders have been, historically speaking?⁴³ Both of us have heard students say: "This goal of social justice is so important that, if we need something like a KGB to achieve it, so be it. We'll just have to make sure the right people run the KGB." But there is no such thing as making sure that the right people run the KGB. People who gravitate toward KGB jobs do so for reasons of their own. Philosophers don't get to stipulate that their reasons are noble.

Beyond non-ideal theory

This book is not about ideal theory. It is not really about non-ideal theory either. We are not trying to say how the principles of justice for an ideal world would have to be modified so as to be fit for the non-ideal world we actually live in. Our first aim is simply to tell a story about liberty. This is not easy. First, it is no easy matter for historians to get their facts right. Even celebrated experts sometimes

misinterpret archeological evidence that falls squarely within their area of greatest expertise. Second, there is no such thing as saying every true thing under the sun. We have to be intensely selective, and, as we judge what to say, our story inevitably will not be the same as the one that someone else would have told.

As we go through the history, we will be theorizing from the bottom up, building toward conjectures about why histories unfold as they do. When we imagine an ideal we at the same time try to imagine what it would be like to get to that ideal starting from where we are.

Finally, we offer piecemeal hypotheses about what worked and what didn't. As we discuss in our final chapter, it is hard to avoid being biased; indeed few succeed. In particular, people often decide what to believe, then look for evidence that seems to confirm that they decided correctly. This bias is hard to avoid; but some books are a joy to read because they simply ask a question, then follow the trail of evidence wherever it may lead, celebrating such surprises as come the author's way. Not many books are like that. But that is our model.

So what sorts of institutions tend to make people free? Which ones make people better off? Many works on liberty try to answer these questions from the armchair, by imagining how potential institutions might do. In this book we look at how real institutions have done.

We conclude that the rule of law is needed in order to provide a framework that encourages experimentation and entrepreneurship. Historically, societies that get their property rights more or less correct tend to achieve prosperity; societies that do not always fail. Cultures of tolerance and openness to change lead to greater prosperity than do closed, intolerant cultures. Societies that economize on moral motivation – working with, rather than against, self-interest – fare better than societies that do not.⁴⁴

The result of freedom of thought, of freedom of association, of the division of labor within firms and of the specialization of roles that evolves between firms is that society becomes an unimaginably

complex web of cooperation, moving ever further away from individual self-sufficiency. Although it may sound somewhat paradoxical, this is actually a contribution to positive freedom, because, as particular roles within society become redundant, a given individual grows less dependent on particular providers of a given service. Freedom in the positive sense can and sometimes does burgeon along with the increasing complexity of this web of interdependence. This book is a story of those preconditions of real choice slowly coming together.

Discussion

- 1 What promotes positive freedom? What is the historical record? Are there cases where people had to choose between respecting negative liberty and promoting positive liberty? If so, what happened? When people have to choose between these two forms, do they end up getting neither?
- 2 Are there circumstances where simply respecting negative liberty is a way (perhaps even the best way) of promoting positive freedom? If so, what would those circumstances be?
- 3 A constitutional framework is a set of relatively fixed rules formulated at a high level of generality. What else does a society need so as to promote positive freedom effectively?
- 4 Suppose we say the point of government is to promote liberty, and the point of liberty is to promote welfare. What would we need to show in order to turn these premises into a solid argument for some form of welfare state?
- 5 To Marx, your degree of freedom in exercising an option depends on how unacceptable your next best alternative is. Suppose Alf does not mind working, whereas Betty would (almost) rather die than have to work for a living. Does this mean that Betty, by virtue of having to work, is more *unfree* than Alf, or simply that Betty *resents* her unfreedom more than Alf does?

- 6 In a marketplace governed by voluntary exchange, there is a general truth that, if you want to induce others to provide you with goods and services, you have to reciprocate, which is to say that you have to work in return for the work others do for you. In some sense, we would be more free if we did not have to reciprocate – if we could consume products of other people's labor for free, without having to offer our own labor in return. Having to reciprocate – to pay for what we consume – makes us less free than if we could consume for free, where 'free' means at someone else's expense. Along this dimension, how free is it good to be?

Acknowledgments

For helpful comments, we thank Jerry Gaus, Benjamin Garcia Helgado, Jeff McMahon, Guido Pincione, and Horacio Spector.

Notes

1. Spiegel 1983, 2.
2. Nietzsche 1994, 57.
3. Gerald MacCallum has argued that all conceptions of freedom have something in common. Ian Carter summarizes:

MacCallum defines the basic concept of freedom – the concept on which everyone agrees – as follows: a subject, or agent, is free from certain constraints, or preventing conditions, to do or become certain things. Freedom is therefore a triadic relation – that is, a relation between *three things*: an agent, certain preventing conditions, and certain doings or becomings of the agent. Any statement about freedom or unfreedom can be translated into a statement of the above form by specifying *what* is free or unfree, *from* what it is free or unfree, and *what* it is free or unfree *to do or become*. Any claim about the presence or absence of freedom in a given situation will therefore make certain assumptions about what counts as an agent, what counts as a constraint or limitation on freedom, and what counts as a purpose that the agent can be described as either free or unfree to carry out. (Carter 2003, 9)

4. Berlin 1997, 177.
5. See Horacio Spector 2009, "Two Dimensions of Freedom," manuscript, 2ff. To the contrast between negative and positive freedom, Horacio Spector adds an orthogonal contrast between natural and civil liberty. The two distinctions, conjoined, define four conceptions of liberty: (1) negative natural liberty is freedom as non-interference (mere liberty); (2) negative civil liberty is civil liberty (which adds rights and immunities to mere liberty so as to secure non-domination, that is, freedom from arbitrary power); (3) positive natural liberty is the capacity for individual self-government (which adds powers); and (4) positive civil liberty is the capacity for civil (collective) self-government.
6. Carter's entry in the *Stanford Encyclopedia of Philosophy* (Carter 2007) distinguishes between the location and the source of obstacles.
7. Berlin 1997, 169.
8. Ibid., 168. Berlin's reference to historians may be to Lord Acton, who said much the same thing.
9. As related by legal scholar Isaac Lifshitz in conversation, Jerusalem, 2008.
10. Again, the technical distinctions we use, such as positive versus negative, are getting at something, but our subject here is not technical. The edges of these concepts are notoriously hard to illuminate. We are seeking to minimize the vagueness and the potential for misunderstanding, even as we accept that eliminating all possibility of vagueness and misunderstanding is not a realistic goal. Success consists in being able to use various distinctions so as to triangulate and to develop a *fairly* accurate sense of what we are talking about.
11. Constant 1988, 311.
12. Ibid.
13. Ibid., 309.
14. For a superb discussion of Berlin, see Carter 2007.
15. Miller (1991, 10) gives a shorter, differently organized list. We do not pretend that ours is the best list for all purposes, but merely that it serves present purposes well enough.
16. Hobbes 1994, 79. The context is that Hobbes is describing the impediments to the Right of Nature as the liberty of people to do whatever they judge to be most conducive to self-preservation.
17. Hobbes 1994, 79.

18. Locke 1980, 32. We say that Locke's characterization is 'slightly' moralized insofar as the idea is not to define the moral so much as the practical limits of liberty. It takes a certain mutual forbearance for us to be truly free to achieve what we want to achieve.
19. Green 1986, 199.
20. Ibid.
21. Green is often seen as a transitional thinker, situated between the classical (individualistic) and the modern (welfare statist) forms of liberalism. On our interpretation, though, Green's conception of freedom does not beg the question against classical liberalism; rather, it is neutral between these broad categories of liberalism.
22. We acknowledge the possibility that conceptual analysis alone might establish a given liberty's intrinsic value.
23. Berlin says this in a discussion of negative liberty, but he never suggests he would categorize such inability as a lack of liberty of any kind.
24. This seems to be the sort of conception that Locke was rejecting when he asked readers to consider the circumstance of a man finding himself in the same room with a person he longs to be with. Without question, he is in the room willingly; but, unknown to him, the door to the room has been locked from the outside, so that he could not leave if he wanted to. Locke is arguing in this passage that liberty is not a matter of doing what one wills, but a matter of having the power to do or forbear to do what one has a mind to do (Locke 1996, 96).
25. Kant 1996, 61–2 and passim.
26. Rousseau 1968, 65.
27. MacCallum (1967) argues, plausibly, that the best way of analyzing freedom is as a three-place relation specifying what agent is free from what impediments to achieve what end.
28. Frankl 1997, 122.
29. Mill 1978, 57.
30. Berlin took a dim view of rationalism in politics, though, so he would not have wanted to be seen as endorsing rational self-direction as an ideal of *collective* self-determination.
31. Berlin claims that "every form of autonomy has in it some element of this attitude. I eliminate the obstacles in my path by abandoning the path" (1997, 182). He goes on to say that "the definition of negative liberty as the ability to do what one wishes . . . will not do. . . . Ascetic

self-denial may be a source of integrity or serenity and spiritual strength, but it is difficult to see how it can be called an enlargement of liberty" (186).

32. Pettit 1997, 271.
33. Ibid.
34. Ibid., 22.
35. The idea is, on its face, absurd, yet there is something interesting here. When someone says, "I am *not* responsible for the Holocaust. I descend from Germans, but the Holocaust happened before I was born," the person is freely *choosing* to regard her responsibilities as limited in this way. There are of course reasons for so choosing. On the other hand, it would be equally intelligible for someone to say, "I claim responsibility for the Holocaust in the sense that, even though it happened before I was born, I choose to take responsibility for making amends, for making sure it never happens again, and simply for remembering what happened." What existentialism should insist on is that people are *not* at liberty to see themselves as not responsible for *their choices*. Seeing oneself as responsible for one's choices is a core constituent of being of good faith.
36. See our final chapter.
37. See also Kant's distinction between *Willkur* (autonomy as a capacity to choose to be moral) and *Wille* (autonomy as the actual choice to be moral). As Horacio Spector notes in his unpublished paper of 2009 (see above, n. 5), Locke distinguishes a collective form of autonomy when he defines civil liberty. "The Natural Liberty of Man is to be free from any Superior Power on Earth, and not to be under the Will or Legislative Authority of Man, but to have only the Law of Nature for his Rule." Civil liberty is a liberty "to be under no other Legislative Power, but that established, by consent, in the Common-wealth, nor under the Dominion of any Will, or Restraint of any Law, but what the Legislative shall enact, according to the Trust put in it" (Locke 1980, opening paragraph of chapter IV). Spector also notes that Rousseau likewise adopts a collective conception of positive political liberty as democratic self-government. According to Rousseau, in the civil state we acquire *moral freedom*, "which alone makes man the master of himself; for to be governed by appetite alone is slavery, while obedience to a law one prescribes to oneself is freedom" (Rousseau 1968, 65).

38. Constant 1988, 326.
39. When someone rejects a cherished government program (e.g. rent-control), it is tempting to assume that the person rejects our values (such as affordable housing for the poor). However, even if you conclude that the government ought to promote some end, this doesn't tell you how it ought to do it. In particular, your conclusion doesn't prescribe direct or indirect means for the government to promote the end in question. For example, a government might attempt to promote commerce directly, by creating new corporations, offering subsidies and grants to businesses, providing tariff protections, and buying products; or indirectly, by providing a basic institutional framework (such as the rule of law, constitutional representative democracy, courts, and a well-functioning property rights regime). Once you settle on an end for the government, which way of achieving it – direct or indirect – works best is always an empirical question. Aiming at something directly is no guarantee of getting it.
40. What would make this a good answer? Perhaps the answer itself is good or bad insofar as it gets the job done or not – and whether it gets the job done depends on the nature of the job. It is one thing if our purpose is to start a conversation. But, if our purpose were instead to provide a definition which even the cleverest philosopher could not devise a counterexample for, we would be a long way from getting *that* job done. A hammer is a tool; in a less obvious way, so are the words we use to describe it.
41. Rawls 2001, 137. Rawls imagines away corruption and moral hazard. So, while he is asking what is possible, he is not asking what is likely. Elsewhere Rawls asks: "what would a just democratic society be like under reasonably favorable but still possible historical conditions allowed by the laws and tendencies of the social world? What ideals and principles would such a society try to realize given the circumstances of justice in a democratic culture as we know them?" (Rawls 2007, 11).
42. Rawls disagrees. He thinks the answer to his question tells us that certain types of social regimes (which Rawls calls property-owning democracy and liberal socialism) are intrinsically more just than others (state socialism, welfare state capitalism, and laissez-faire capitalism). His theory leaves it open that, in the real world, we might have to choose

Murray N. Rothbard, *The Ethics of Liberty*, New York University Press, New York, 1998.

12. Self-Defense

If every man has the absolute right to his justly-held property it then follows that he has the right to keep that property-to defend it by violence against violent invasion. Absolute pacifists who also assert their belief in property rights-such as Mr. Robert LeFevre-are caught in an inescapable inner contradiction: for if a man owns property and yet is denied the right to defend it against attack, then it is clear that a very important aspect of that ownership is being denied to him. To say that someone has the absolute right to a certain property but lacks the right to defend it against attack or invasion is also to say that he does not have total right to that property.

Furthermore, if every man has the right to defend his person and property against attack, then he must also have the right to hire or accept the aid of other people to do such defending: he may employ or accept defenders just as he may employ or accept the volunteer services of gardeners on his lawn.

How extensive is a man's right of self-defense of person and property? The basic answer must be: up to the point at which he begins to infringe on the property rights of someone else. For, in that case, his "defense" would in itself constitute a criminal invasion of the just property of some other man, which the latter could properly defend himself against.

It follows that defensive violence may only be used against an actual or directly threatened invasion of a person's property-and may not be used against any nonviolent "harm" that may befall a person's income or property value. Thus, suppose that A, B, C, D . . . etc. decide, for whatever reason, to boycott the sales of goods from Smith's factory or store. They picket, distribute leaflets, and make speeches-all in a non-invasive manner--calling on everyone to boycott Smith. Smith may lose considerable

income, and they may well be doing this for trivial or even immoral reasons; but the fact remains that organizing such a boycott is perfectly within their rights, and if Smith tried to use violence to break up such boycott activities he would be a criminal invader of their property. Defensive violence, therefore, must be confined to resisting invasive acts against person or property. But such invasion may include two corollaries to actual physical aggression: intimidation, or a direct threat of physical violence; and fraud, which involves the appropriation of someone else's property without his consent, and is therefore "implicit theft."

THE ETHICS OF LIBERTY

Thus, suppose someone approaches you on the street, whips out a gun, and demands your wallet. He might not have molested you physically during this encounter, but he has extracted money from you on the basis of a direct, overt threat that he would shoot you if you disobeyed his commands. He has used the threat of invasion to obtain your obedience to his commands, and this is equivalent to the invasion itself.

It is important to insist, however, that the threat of aggression be palpable, immediate, and direct; in short, that it be embodied in the initiation of an overt act. Any remote or indirect criterion—any "risk" or "threat"—is simply an excuse for invasive action by the supposed "defender" against the alleged "threat." One of the major arguments, for example, for the prohibition of alcohol in the 1920s was that the imbibing of alcohol increased the likelihood of (unspecified) people committing various crimes; therefore, prohibition was held to be a "defensive" act in defense of person and property. In fact, of course, it was brutally invasive of the rights of person and property, of the right to buy, sell, and use alcoholic beverages. In the same way, it could be held that (a) the failure to ingest vitamins makes people more irritable, that (b) the failure is therefore likely to increase crime, and that therefore (c) everyone should be forced to take the proper amount of vitamins daily. Once we bring in "threats" to person and property that are vague and future—i.e., are not overt and immediate—then all manner of tyranny becomes excusable. The only way to guard against such despotism is to keep the criterion of perceived invasion clear and immediate and overt. For, in the inevitable case of fuzzy or unclear actions, we must bend over backwards to require the threat of invasion to be direct and immediate, and therefore to allow people to do whatever they may be doing. In short, the burden of proof that the aggression has really begun must be on the person who employs the defensive violence.

Fraud as implicit theft stems from the right of free contract, derived in turn from the rights of private property. Thus, suppose that Smith and Jones agree on a contractual exchange of property titles: Smith will pay \$1000 in return for Jones's car. If Smith appropriates the car and then refuses to turn over \$1000 to Jones, then Smith has in

effect stolen the \$1000; Smith is an aggressor against \$1000 now properly belonging to Jones. Thus, failure to keep a contract of this type is tantamount to theft, and therefore to a physical appropriation of another's property fully as "violent" as trespass or simple burglary without armed assault.

SELF-DEFENSE

Fraudulent adulteration is equally implicit theft. If Smith pays \$1000 and receives from Jones not a specified make of car but an older and poorer car, this too is implicit theft: once again, someone's property has been appropriated in a contract, without the other person's property being turned over to him as agreed.¹

But we must not be led into the trap of holding that all contracts, whatever their nature, must be enforceable (i.e., that violence may properly be used in their enforcement). The only reason the above contracts are enforceable is that breaking such contracts involves an implicit theft of property. Those contracts which do not involve implicit theft should not be enforceable in a libertarian society. Suppose, for example, that A and B make an agreement, a "contract," to get married in six months; or that A promises that, in six months' time, A will give B a certain sum of money. If A breaks these agreements, he may perhaps be morally reprehensible, but he has not implicitly stolen the other person's property, and therefore such a contract cannot be enforced. To use violence in order to force A to carry out such contracts would be just as much a criminal invasion of A's rights as it would be if Smith decided to use violence against the men who boycotted his store. Simple promises, therefore, are not properly enforceable contracts, because breaking them does not involve invasion of property or implicit theft.

Debt contracts are properly enforceable, not because a promise is involved, but because the creditor's property is appropriated without his consent-i.e., stolen-if the debt is not paid. Thus, if Brown lends Green \$1000 this year in return for the delivery of \$1100 next year, and Green fails to pay the \$1100, the proper conclusion is that Green has appropriated \$1100 of Smith's property, which Green refuses to turn over-in effect, has stolen. This legal way of treating a debt--of holding that the creditor has a property in the debt-should be applied to all debt contracts.

Thus, it is not the business of law-properly the rules and instrumentalities by which person and property are violently defended to make people moral by use of legal violence. It is not the proper business of law to make people be truthful or to keep their promises. It is the business of legal violence to defend persons and their property from violent attack, from molestation or appropriation of their property

1. For a development of libertarian principles of the law of adulteration, see Wordsworth Donisthorpe, *Law In A Free State* (London: Macmillan, 1895), pp. 132-58.
2. For a further development of this thesis, see the section "Property Rights and the Theory of Contracts," pp. 133-48 below.

THE ETHICS OF LIBERTY

without their consent. To say more-to say, for example, that mere promises are properly enforceable-is to make an unwarranted fetish of "contracts" while forgetting why some of them are enforceable: in defense of the just rights of property.

Violent defense then must be confined to violent invasion-either actually, implicitly, or by direct and overt threat. But given this principle, how far does the right of violent defense go? For one thing, it would clearly be grotesque and criminally invasive to shoot a man across the street because his angry look seemed to you to portend an invasion. The danger must be immediate and overt, we might say, "clear and present-a criterion that properly applies not to restrictions on freedom of speech (never permissible, if we regard such freedom as a subset of the rights of person and property) but to the right to take coercive action against a supposedly imminent invader."

Secondly, we may ask: must we go along with those libertarians who claim that a storekeeper has the right to kill a lad as punishment for snatching a piece of his bubble gum? What we might call the "maximalist" position goes as follows: by stealing the bubble gum, the urchin puts himself outside the law. He demonstrates by his action that he does not hold or respect the correct theory of property rights. Therefore, he loses all of his rights, and the storekeeper is within his rights to kill the lad in retaliation.

I propose that this position suffers from a grotesque lack of proportion.

By concentrating on the storekeeper's right to his bubble gum, it totally ignores another highly precious property-right: every man's-including the urchin's- right of self-ownership. On what basis must we hold that a minuscule invasion of another's property lays one forfeit to the total loss of one's own? I propose another fundamental rule regarding crime: the criminal, or invader, loses his own right to the extent that he has deprived another man of his. If a man deprives another man of some of his self-ownership or its extension in physical property, to that extent does he lose his own rights. From this principle immediately

3. This requirement recalls the scholastic doctrine of the double effect. See G.E.M. Anscombe, "The Two Kinds of Error in Action," *Journal of Philosophy* 60 (1963): 393-401; Philippa R. Foot, *Virtues and Vices* (Berkeley: University of California Press, 1978), pp. 19-25.

4. On the maximalist view, furthermore, socialists, interventionists and utilitarians would, by virtue of their views, be liable to execution. I am indebted to Dr. David Gordon for this point.

5. The great libertarian Auberon Herbert, in Auberon Herbert and J.H. Levy, *Taxation and Anarchism* (London: Personal Rights Association, 1912), p. 38, put it this way:

SELF-DEFENSE

derives the proportionality theory of punishment-best summed up in the old adage: "let the punishment fit the crime."

We conclude that the shop keeper's shooting of the erring lad went

beyond this proportionate loss of rights, to wounding or killing the criminal; this going beyond is in itself an invasion of the property right in his own person of the bubble gum thief. In fact, the storekeeper has become a far greater criminal than the thief, for he has killed or wounded his victim—a far graver invasion of another's rights than the original shoplifting.

Should it be illegal, we may next inquire, to "incite to riot"? Suppose that Green exhorts a crowd: "Go! Burn! Loot! Kill!" and the mob proceeds to do just that, with Green having nothing further to do with these criminal activities. Since every man is free to adopt or not adopt any course of action he wishes, we cannot say that in some way Green determined the members of the mob to their criminal activities; we cannot make him, because of his exhortation, at all responsible for their crimes. "Inciting to riot," therefore, is a pure exercise of a man's right to speak without being thereby implicated in crime. On the other hand, it is obvious that if Green happened to be involved in a plan or conspiracy with others to commit various crimes, and that then Green told them to proceed, he would then be just as implicated in the crimes as are the others—more so, if he were the mastermind who headed the criminal gang. This is a seemingly subtle distinction which in practice is clear-cut: there is a world of difference between the head of a criminal gang and a soap-box orator during a riot; the former is not, properly to be charged simply with "incitement."

It should further be clear from our discussion of defense that every man has the absolute right to bear arms—whether for self-defense or any other licit purpose. The crime comes not from bearing arms, but from using them for purposes of threatened or actual invasion. It is curious, by the way that the laws have especially banned concealed weapons, when Am I right in saying that a man has forfeited his own rights (to the extent of the aggression he has committed) in attacking the rights of others? . . . It may be very difficult to translate into concrete terms the amount of aggression, and of resulting restraint; but all just law seems to be the effort to do this. We punish a man in a certain way if he has inflicted an injury which lays me up for a day; in another way if he takes my life. . . . There is generally underlying it [the law] the view (which is, I think, true) that the punishment or redress—both in civil and criminal matters—should be measured by the amount of aggression; in other words that the aggressor after a rough fashion—loses as much liberty as that of which he has deprived others.

6. For a development of this theory of punishment, see the section "Punishment and Proportionality," pp. 85-96 below.

82 THE ETHICS OF LIBERTY

it is precisely the open and unconcealed weapons which might be used for intimidation.

In every crime, in every invasion of rights, from the most negligible breach of contract up to murder, there are always two parties (or sets of parties) involved: the victim (the plaintiff) and the alleged criminal (the defendant). The purpose of every judicial proceeding is to find, as best

we can, who the criminal is or is not in any given case. Generally, these judicial rules make for the most widely acceptable means of finding out who the criminals may be. But the libertarian has one overriding caveat on these procedures: no force may be used against non-criminals. For any physical force used against a non-criminal is an invasion of that innocent person's rights, and is therefore itself criminal and impermissible. Take, for example, the police practice of beating and torturing suspects-or, at least, of tapping their wires. People who object to these practices are invariably accused by conservatives of "coddling criminals." But the whole point is that we don't know if these are criminals or not, and until convicted, they must be presumed not to be criminals and to enjoy all the rights of the innocent: in the words of the famous phrase, "they are innocent until proven guilty." (The only exception would be a victim exerting self-defense on the spot against an aggressor, for he knows that the criminal is invading his home.) "Coddling criminals" then becomes, in actuality, making sure that police do not criminally invade the rights of self-ownership of presumptive innocents whom they suspect of crime. In that case, the "coddler," and the restrainer of the police, proves to be far more of a genuine defender of property rights than is the conservative.

We may qualify this discussion in one important sense: police may use such coercive methods provided that the suspect turns out to be guilty, and provided that the police are treated as themselves criminal if the suspect is not proven guilty. For, in that case, the rule of no force against non-criminals would still apply. Suppose, for example, that police beat and torture a suspected murderer to find information (not to wring a confession, since obviously a coerced confession could never be considered valid). If the suspect turns out to be guilty, then the police should be exonerated, for then they have only ladled out to the murderer a parcel of what he deserves in return; his rights had already been forfeited by more than that extent. But if the suspect is not convicted, then that means that the police have beaten and tortured an innocent man, and that they in turn must be put into the dock for criminal assault. In short, in all cases, police must be treated in precisely the same way as anyone else; in a libertarian world, every man has equal liberty, equal

SELF-DEFENSE 83

rights under the libertarian law. There can be no special immunities, special licenses to commit crime. That means that police, in a libertarian society, must take their chances like anyone else; if they commit an act of invasion against someone, that someone had better turn out to deserve it, otherwise they are the criminals.

As a corollary, police can never be allowed to commit an invasion that is worse than, or that is more than proportionate to, the crime under investigation. Thus, the police can never be allowed to beat and torture someone charged with petty theft, since the beating is far more proportionate a violation of a man's rights than the theft, even if

the man is indeed the thief.

It should be clear that no man, in an attempt to exercise his right of self-defense, may coerce anyone else into defending him. For that would mean that the defender himself would be a criminal invader of the rights of others. Thus, if A is aggressing against B, B may not use force to compel C to join in defending him, for then B would be just as much a criminal aggressor against C. This immediately rules out conscription for defense, for conscription enslaves a man and forces him to fight on someone else's behalf. It also rules out such a deeply embedded part of our legal system as compulsory witnesses. No man should have the right to force anyone else to speak on any subject.

The familiar prohibition against coerced self-incrimination is all very well, but it should be extended to preserving the right not to incriminate anyone else, or indeed to say nothing at all. The freedom to speak is meaningless without the corollary freedom to keep silent.

If no force may be used against a noncriminal, then the current system of compulsory jury duty must also be abolished. Just as conscription is a form of slavery, so too is compulsory jury duty. Precisely because being a juror is so important a service, the service must not be filled by resentful serfs. And how can any society call itself "libertarian" that rests on a foundation of jury slavery? In the current system, the courts enslave jurors because they pay a daily wage so far below the market price that the inevitable shortage of jury labor has to be supplied by coercion. The problem is very much the same as the military draft, where the army pays far below the market wage for privates, cannot obtain the number of men they want at that wage, and then turns to conscription to supply the gap. Let the courts pay the market wage for jurors, and sufficient supply will be forthcoming. If there can be no compulsion against jurors or witnesses, then a libertarian legal order will have to eliminate the entire concept of the subpoena power. Witnesses, of course, may be requested to appear. But

THE ETHICS OF LIBERTY

this voluntarism must also apply to the defendants, since they have not yet been convicted of crime. In a libertarian society, the plaintiff would notify the defendant that the latter is being charged with a crime, and that a trial of the defendant will be underway. The defendant would be simply invited to appear. There would be no compulsion on him to appear. If he chose not to defend himself, then the trial would proceed in absentia, which of course would mean that the defendant's chances would be by that much diminished. Compulsion could only be used against the defendant after his final conviction. In the same way, a defendant could not be kept in jail before his conviction, unless, as in the case of police coercion, the jailer is prepared to face a kidnapping conviction if the defendant turns out to be innocent.