

Introduction

This book is about the psychology of our freedom. By *freedom* I mean the freedom of alternative possibilities: the freedom to do things or not do them, or – as I shall also put it – *control* over whether we do those things or not. It is just this freedom that we think we possess in relation to much of our action. We ordinarily think that we are free, say, to go out in the evening or stay in – that it is within our control, or up to us, which of these actions we perform.

The Psychology of Freedom examines what sort of mind this freedom of action requires. It seeks to determine what mental capacities and states we need if we are to be free to act otherwise than as we actually do. In particular, the book examines what sort of *will* a free agent must possess. By the will I mean our capacity to determine in advance how we shall act through decision-making, thereby forming intentions about which actions we shall perform.

Much of the book discusses what goes on when we take a decision to act, and how the nature and rationality of our decisions relate to the nature and rationality of the actions which those decisions explain. This is because our ordinary conception of our freedom of action involves a conception of the will which is very demanding – and perhaps not even coherent. It is a controversial and unsolved problem what goes on in the decision-making by which we determine our actions. And this problem about the will is part of a wider problem about the nature of freedom itself.

Our ordinary conception of our freedom of action – I argue – requires that decisions combine two characteristics which, as we shall see, are not obviously consistent. Our conception of our freedom of action demands, first, that decisions to act serve an *executive* function – that they serve to apply practical reason as it concerns the actions decided upon. The point of taking decisions about which actions to perform, after all, is to ensure that we end

up performing the right actions. And secondly, our conception of our freedom demands that decisions count as deliberate or intentional actions themselves.

The book explains why our conception of our freedom requires that decisions combine these two characteristics, why these characteristics are in tension, and how the apparent inconsistency between them can in fact be resolved by a new theory of practical reason and agency.

The book's argument is complex. But then the issues which it addresses are not simple either. So this introduction will present, in outline, the book's general argument. I shall begin by saying something about what I take to be our ordinary conception of our freedom.

THE PSYCHOLOGISING CONCEPTION OF FREEDOM

In the seventeenth century, Thomas Hobbes asserted a quite novel and very simple theory of freedom – a theory which certainly contains no special obscurity in its psychology. According to Hobbes, an agent is free – has control over which actions he performs – to the extent that he has a capacity to act however he desires. To be free, then, one needs nothing more by way of a will or action-determining capacity, than a capacity to hold desires to act, and to form beliefs about means to performing the actions desired.

But Hobbes' theory was deeply revisionary, both of our ordinary conception of our freedom, and of our ordinary conception of the will. There is more to our freedom of action, as we ordinarily conceive it, than Hobbes allowed. And there is correspondingly more to the kind of will which that freedom requires.

It is within our control which actions we perform only because we are capable of taking decisions about how we shall act, and it is within our control or up to us which such decisions we take. Is not that a claim about our freedom to which we naturally assent?

If so, our ordinary conception of our freedom is *Psychologising*. By which I mean that our freedom of action is conceived to depend on a further, psychological analogue of that freedom – on a freedom of some at least of the mental events and states which motivate and give rise to our action. Freedom of action is conceived to depend on a prior freedom of will.

Introduction

3

And that means that our freedom of action depends on our possessing a rather special psychology. For some at least of the mental occurrences which motivate and give rise to free actions must constitute yet further actions themselves. Whether I decide to stay in or decide to go out is supposed to be within my control – up to me – just as is whether I stay in or go out. But that means that, just like staying in or going out, taking a decision to stay in, or taking a decision to go out, must be something which I can deliberately do. Decision-making must be a further form of deliberate agency or doing, additional to the agency of the action which it explains. A Psychologising conception of our freedom ties our freedom of action to our possession of a prior capacity for active decision-making – to our possession of an action-generating agency of the will.

It is important to distinguish a Psychologising conception of our freedom from any volitional or conational theory of action. Some philosophers think that our bodily actions are explained by or involve volitions or conations whereby we *try* to move our bodies. But, as I argue in this book, decisions are importantly different from anything which constitutes trying. To take a decision to act is not, like trying, to initiate bodily movement, but rather to form a persisting psychological state – a state of intention in which we are left motivated to act as decided. Decisions are forms of agency which explain intentions; whereas tryings, like bodily actions, are quite distinct forms of agency which intentions explain.

Taking a decision to act, then, is not just an action-generating action. Unlike tryings or conations, I argue, a decision is an action which generates action by way of constituting the formation of an intervening action explanatory psychological state. That makes decision-making a highly distinctive kind of action-generating action – what I term a *second-order* action, by contrast to the *first-order* actions it generates.

It is the argument of this book, that our concept of agency is deployed at two levels. First we have our everyday actions and attempts at action – our first-order agency. Then we have, among the psychological states which explain and rationalise our first-order agency, some such as intentions whose formation constitutes agency too. That agency, of which the most intuitive case is decision-making, is our second-order agency. By tying our action control to a prior decision control, a Psychologising conception of

freedom ties the freedom of our first-order agency to a capacity for second-order agency.

The purpose of *The Psychology of Freedom* is to set out and defend precisely this Psychologising conception of our freedom of action. The book does this by providing a theory of what second-order agency is, and why free agents need a capacity to perform it. And this theory of second-order agency is proposed within a theory of agency in general – a theory of what unites our decisions and the actions which they explain to make them all actions.

The book addresses two crucial questions which the Psychologising conception of our freedom raises, and which any successful defence of it must answer. First, why should freedom of action depend on a prior freedom of will? And secondly, do we actually possess the second-order, action-generating agency of the will which freedom of will involves? I shall now say something about each of these questions in turn.

WHY SHOULD FREEDOM OF ACTION DEPEND ON FREEDOM OF WILL?

It is not simply our ordinary conception of our freedom which is Psychologising. There have been plenty of philosophical accounts of freedom which are Psychologising too. We find Psychologising accounts of freedom within Stoicism, throughout mediaeval Scholasticism, and in Kant. These accounts typically appeal, as I shall also be appealing, to the assumption that freedom of action involves a capacity for rational self-determination – a capacity, which we possess as agents who are practically rational, to reason or deliberate about how to act, and then to apply those deliberations in the exercise of control over our action. Freedom of action then depends on freedom of will, because this capacity for rational self-determination requires the possession of a free will.

My own Psychologising theory is of this type. But it is going to differ from many past Psychologising theories of freedom at one important point. In many past theories, freedom of will was often conceived as the freedom of a deliberative capacity – as a freedom of our practical reason. This is true, in particular, of Kant's theory of freedom. Freedom of will, according to Kant, is a freedom

Introduction

5

which we possess as rational deliberators – for, in his view, as rational deliberators we are the sovereign legislators of reason's requirements on action.

My theory is rather different. On my theory, the will on whose freedom our freedom of action depends, is not in any sense a deliberative capacity. Rather, the will is an *executive* capacity – a capacity for *applying* our prior deliberations about how to act. A decision is a second-order executive action – an action by which we ensure that we subsequently perform the first-order actions which, as deliberators, we have judged it desirable to perform. Freedom of action depends on freedom of will, I shall be arguing, because to exercise action control in a way which applies our deliberations about how to act, our action control must extend into the future. We must be able to exercise control in advance over what future actions we perform – a control that we exercise through a prior action-generating agency of decision-making.

Our ordinary conception of the will is, I argue, precisely a conception of it as a faculty for second-order executive agency – an agency the function of which is to apply reason as it governs the first-order agency of the actions decided upon. For, as we ordinarily conceive it, decision-making provides a highly distinctive, *reason-applying* way of controlling one's future actions.

We can, of course, use ordinary, first-order actions to control our future actions. Jon Elster's *Ulysses and the Sirens* contains a well-known discussion of various ways of doing just this. Ulysses is a case in point, who had himself bound to a mast to prevent himself from subsequently trying to leave his ship to join the Sirens. But there are other, more everyday and less drastic examples. Concerned for our health, we might perform some first-order action now in order to get ourselves to eat less in future. For example, we might now tell our friends that we are going on a diet. That way we bring the fear of losing face in as a further motive to induce ourselves to eat less in future.

But such use of present first-order action to control one's future action is essentially manipulative. One is twisting the arm of one's future self to get it to perform actions which now, prior to the arm twisting, one favours performing on quite other grounds. Fear of losing face is in the future going to be the motive which induces one to eat less. It is a concern for one's health, however, which is presently motivating one to make use of that fear to ensure that

one eats less. This arm-twisting use of present first-order action to control future action can even threaten or wholly remove one's future freedom of action. Ulysses' use of present action to control his future action placed a brute physical constraint on what he might do in the future.

The use of the second-order agency of our decision-making to control future action is, by contrast, essentially freedom-preserving. Having taken a decision to act – a decision which will cause us to act as we have decided – we none the less always retain a continuing freedom not to act as we have decided. And that is because decision-making is also a method of future action control which is essentially non-manipulative. Deciding to perform an action is not a form of arm-twisting. When one takes a decision to act, one is not putting brute constraints on one's future self's capacity for action. Nor is one creating inducements or penalties to get one's future self to act as one now thinks it should. And that is because as second-order agency, decision-making is, quite distinctively and unlike first-order action, a *motivation-perpetuating* cause of future action.

A decision to act affects future action, I shall be arguing, simply by perpetuating the motivating force of the considerations that have already motivated it. The decision ensures that the reasons which motivated one to take it also motivate one subsequently to perform the action decided upon. If I decide to eat less out of a concern for my health, then my decision ensures that that same concern for my health will thereafter motivate me actually to eat less. Taking a decision perpetuates our present motivation into the future, so that our motivation in the future will remain one with which we can, as we are at present constituted, immediately identify. When we take a decision to act, our future self is left a person who entirely shares our present rationality – and who, precisely because of that shared rationality, is motivated and willing to act as we have decided.

Decision-making, then, is a very special kind of future action-controlling action. By taking a decision to act, we are not only applying practical reason now; we also ensure that we go on applying that same practical reason into the future, in our very performance of the actions which our present deliberations recommend. Decision-making, then, is a form of action by which we exercise future action control in a way that ensures our continuing

Introduction

7

rationality as agents. And it is that form of future action control which, I shall be arguing, we need to be free.

Thus the book contains an original explanation of why freedom of action depends on freedom of will. And this explanation is itself based on a new theory of decision-making as a reason-applying, second-order executive action – an action by which we ensure that we act rationally thereafter.

IS THERE AN AGENCY OF THE WILL?

But is taking a particular decision to act a deliberate action itself? Whatever ordinary intuition might say on the matter, from Hobbes onwards plenty of philosophers have been pretty sure it is not. And many more philosophers have simply avoided the issue. This book is indebted to Michael Bratman for his important work on the action-coordinatory and planning function of intention. But in his *Intention, Plans and Practical Reason*, Bratman notably refrains from claiming that forming a particular intention to act – what we do when we take a decision to act – is itself an intentional action. And such restraint is entirely understandable. It is a serious question whether there really is an agency of the will.

It is not hard to see where the scepticism comes from. The actions which our decisions explain are purposive. We perform these actions for the sake of some end – either for their own sake, or for the sake of wider ends extending beyond what we are doing. We perform these actions, then, in the belief that they will attain desirable ends. And that is connected with the fact that the practical deliberation which explains our actions is precisely deliberation about how to act – it is precisely deliberation in which we consider what ends our actions might further.

But the forming of intentions to act does not seem to be purposive. It does not seem to be motivated by beliefs about what ends it will further. We form intentions to do A because we believe *doing A* – the action decided upon – will attain desirable ends, and not because we believe that *now intending to do A* will attain desirable ends. The deliberation that explains our intentions to act, then, is simply deliberation about which actions to perform. Our intentions are not generally based on practical deliberation about which intentions to form.

Again actions, at least in their uncontroversial first-order form, are subject to the will. We can perform actions A on the basis of having decided to do so. But the decisions which determine our actions are not themselves subject to the will in the same way. We cannot, it seems, take decisions to do A on the basis of having decided to take them.

Decisions to act, then, look importantly different from the actions which they explain. Actions are purposive, based on practical deliberation about whether to do them and about what ends they would further, and are subject to the will. Typically, decisions to act share none of these properties. Why should we suppose that these decisions to act are deliberate actions too? These dissimilarities between decisions and the actions they explain, have been the staple of scepticism about second-order agency since Hobbes. *The Psychology of Freedom* concedes that these dissimilarities do indeed exist, and uses its new theory of decision-making to explain them. At the same time, the book also provides a new theory of agency – a theory which shows why, these differences from ordinary first-order action notwithstanding, decisions to act are indeed deliberate actions themselves.

What might unite decisions to act and the actions which they explain to make them uniformly actions? Perhaps the uniting factor is practical reason itself. It may be that there is a distinctively *practical* reason – a distinctively practical way in which the justification, or lack of it, for all our *practice* or agency is determined. In which case decision-making can count as action if, just like the first-order action which it explains, it is governed by reason in this distinctively practical way.

But what is practical reason? One possibility is that reason in its distinctively practical form comes to *means–end justifiability*. First-order actions are means–end justifiable – that is, they are justified in terms of the likelihood that performing those particular actions would further desirable ends. What if decisions to do A are means–end justifiable too? One way of defending the agency of decision-making, then, is by appeal to what I call the *Action model* of decision rationality. On this model, decisions to act are means–end justifiable – justifiable in terms of the desirable ends which they themselves are likely to further – just as are the actions which those decisions explain. And that makes decisions to act actions too.

DECISION RATIONALITY: THE ACTION MODEL VERSUS
 THE PRO ATTITUDE MODEL

However, it is not obvious that the Action model of decision rationality is true. The Action model of decision rationality seems to clash with the belief, which I have argued our conception of freedom also demands, that decisions to act have an executive function – that they serve to apply reason as it governs our subsequent action.

For if the function of decisions to act is executive – if the function of decisions to act is to apply reason as it governs first-order action – then should not the rationality of deciding to do A be explained simply in terms of the rationality of doing A? And so we have the *Pro Attitude model* of decision rationality – a view of decision rationality assumed in the work of philosophers such as Anscombe, Davidson, and Bratman. On the Pro Attitude model, justifications for deciding to do A are simply justifications for eventually doing A. Decisions to do A are justified simply in terms of the desirability of doing A – in terms of the likelihood of doing A furthering desirable ends, and not in terms of the likelihood of those decisions themselves furthering desirable ends. On the Pro Attitude model, whilst the actions which our decisions explain are means–end justifiable, our decisions are not.

We have then two competing models of decision rationality. One, the Pro Attitude model, says that decisions to do A are justified in terms of desirable ends which doing A would further. The other, the Action model, says that decisions to do A are justified by desirable ends which deciding to do A would further. And the potential difference between these models emerges clearly if we imagine, as Gregory Kavka has asked us to imagine, a huge cash prize offered simply for taking a decision to perform a particular action A – a prize which we win irrespective of whether we ever carry out the decision thereafter. It is just the taking of a decision to do A which the prize rewards, and not doing A.¹

Now on the Pro Attitude model, decisions to act cannot rationally be motivated by prizes awarded just for taking them. For such prizes do not justify the actions decided upon – winning the prize is not a desirable end which the action decided upon would further.

¹ See Kavka's 'The toxin puzzle'; *Analysis* (1983).

And so, the Pro Attitude model implies, such prizes cannot justify the deciding to do A. On the other hand, it might appear, the Action model implies that the prize does justify deciding to do A. For is not winning the prize a desirable end which deciding to do A would further?

But then, if decisions to do A *can* be made rational by prizes offered just for taking them, the rationality of taking a decision to do A need have no connexion at all with – and certainly will not guarantee – the rationality of subsequently doing A. The Action model seems inconsistent with the view that the essence of deciding to perform an action is executive – that decision-making serves to apply practical reason as it concerns the actions decided upon.

At the heart of the book, then, is a deep tension in our conception of the will – the very conception of the will as an executive agency that our belief in our freedom demands. We think of decisions as executing reason as it governs first-order action. The Pro Attitude model of decision rationality looks true. At the same time, we think of decisions as being actions themselves – and so as justified by the desirability of taking them, just as the actions which our decisions explain are justified by the desirability of their performance. The Action model of decision rationality looks true.

The Psychology of Freedom resolves this tension. It argues that the Pro Attitude model of decision rationality is false, and the Action model is true. And that makes decisions actions. On the other hand, it defends a restriction – which I term REASON-APPLY – on the ends which can rationally justify one's taking a decision. The Action model says that justifications for deciding to perform an action are desirable ends which taking that decision would further. But it is consistent with the Action model that not *every* desirable end which a decision would further can justify taking that decision. According to REASON-APPLY, desirable ends can *only* justify one's taking a decision to act if, once that decision is taken, they would thereafter be furthered by, and so justify, one's acting as decided.

That means that decisions are not justified by prizes offered just for taking them. For such prizes do not thereafter justify acting as decided. The Action model, when supplemented with REASON-APPLY, places restrictions on what can justify decisions that are very like – though, as we shall see, importantly not identical with – those imposed by the rival Pro Attitude model. And so we arrive at a version of the Action model that, I shall be arguing, is fully