This book is a selection of my essays on moral responsibility. I seek to present what I call an overall “framework” for moral responsibility. In some cases the essays introduce or refine ideas that were originally presented elsewhere. Also, in some of the essays I attempt to defend my views against critiques, with the hope that, in the end, the basic structure of the views—and the arguments for them—emerge with greater clarity and force. In this introductory essay, I shall sketch some of the leading ideas in my framework, and I shall take the liberty (on a few occasions) of referring to some additional work (not included in My Way) that develops the material in further detail.

I. A Framework for Moral Responsibility

I.1. Motivation and the Concept of Responsibility

The framework I present for moral responsibility involves a portfolio of different ideas in a certain arrangement. I start by presenting some basic “motivating ideas”—some considerations that render my overall approach attractive. Perhaps the key idea here stems from the appeal of a certain sort of “resiliency”. I believe that our fundamental status as agents—our being deeply different from mere nonhuman animals insofar as we engage in practical reasoning and are morally responsible for our behavior—should not depend on the subtle ruminations of theoretical physicists. That is, I do not think that our status as genuine

---

agents should hang on a thread—that it should depend on whether or not natural laws have associated with them say 99% probabilities or 100% probabilities. In my view, that sort of empirical difference should not make a difference as to our moral responsibility. So, for example, if in the future I am convinced that the fundamental laws of nature are—or can be regimented as—(among other things) universally generalized conditionals with 100% probabilities rather than similar conditionals with 99% probabilities, this would not issue in any inclination to give up my view of myself and others as genuine agents and legitimate participants in the practices constitutive of moral responsibility. Of course, this is simply one consideration, and it specifies a desideratum of an adequate theory of moral responsibility: it does not in itself count as evidence of the truth of the theory. For example, it does not in itself address any of the difficult skeptical worries about the relationship between causal determinism, free will, and moral responsibility.

Note that, insofar as I take it that it would be desirable to have an account of moral responsibility according to which our fundamental status as morally responsible agents does not “hang on a thread”, it would follow also that we should not give up our views of ourselves as deeply different from nonhuman animals (in the relevant ways) if we are convinced (in the future) that the fundamental laws of nature have irreducible indeterminacies associated with them; suppose, for example, that we discover that these laws are indeed (among other things) universally quantified conditionals with 99% probabilities. In my view, this in itself should not issue in any inclination to discard or revise our views of ourselves and others as genuine agents and subject to moral responsibility. Again, I am here simply articulating what I take to be a desideratum of an adequate theory of moral responsibility; it does not in itself provide evidence of the truth of the theory. For example, it does not in itself provide any sort of answer to the difficult skeptical worries about the relationship between causal indeterminism and control (and moral responsibility).

A second element in the overall framework for moral responsibility consists in an articulation of the “concept” of moral responsibility. I accept some sort of distinction between the concept and its conditions of application; I of course recognize that the legitimacy of this sort of distinction has been called into question. And yet I continue to think that there is some reasonable way (or ways) of making the relevant kind of distinction, even if it is not straightforwardly a matter of distinguishing “analytic” from “synthetic” truths, or matters of meaning from empirical matters. I simply presuppose that there is some tolerably clear way of distinguishing (roughly speaking) the concept of moral responsibility from the conditions in which moral responsibility actually obtains.
As I said above, my overall framework for moral responsibility is a portfolio of ideas in a certain arrangement. Part of the portfolio is a suite of options in regard to the concept of moral responsibility; but I do not take a stand on these options. That is, I chart out different views that seek to articulate our inchoate concept of moral responsibility: but I do not argue that one (rather than the others) is the correct specification. I am not even sure that there is one unique specification. Rather, I focus most of my attention on specifying the conditions of application of the concept of moral responsibility, and I contend that accepting these conditions is completely compatible with accepting any of the specific options with respect to the concept of moral responsibility.

To be a bit more specific about the concept of moral responsibility, perhaps the most salient view might be called the “Strawsonian” view, following the classic presentation by Peter Strawson in “Freedom and Resentment”.2 On this view, being morally responsible is a matter of being an appropriate target of a set of distinctive attitudes Strawson dubbed the “reactive attitudes”, such as gratitude, love, respect, hatred, and resentment, and appropriate participants in activities, such as moral praise and blame and punishment, which presuppose the application of the relevant attitudes. It was important to Strawson that the “appropriateness” of the attitudes does not depend on the target agent’s meeting some “theoretical condition”, such as possessing free will; additionally, it does not depend (for Strawson) on the world’s meeting certain conditions, such as that causal determinism is false (or, for that matter, true).

Another account of the concept of moral responsibility is associated with the metaphor of a “moral ledger”. On the moral ledger view, we are morally responsible insofar as we are apt targets of specifically moral judgments. On this view, we are deeply different from nonhuman animals in that we can have moral properties—we can act rightly or wrongly, we can be good or bad, courageous or cowardly, and so forth. On yet another view, we are morally responsible insofar as we can legitimately be asked to provide explanations or accounts of our behavior. As I said above, I do not know whether there is a single correct specification of our concept of moral responsibility; perhaps “moral responsibility” is what Wittgenstein called a “family-resemblance” term. In any case, I contend that my account of the conditions in which moral responsibility obtains is compatible with any of the plausible attempts to specify the concept.

I accept the traditional view that moral responsibility involves a *freedom* or *control* component and an *epistemic* component. But whereas I agree that moral responsibility requires control, I distinguish two kinds of control: guidance and regulative control. The two kinds of control can be prized apart analytically through the use of certain thought-experiments (the Frankfurt-style examples). One kind of control involves genuine metaphysical access to alternative possibilities (freedom to choose and do otherwise); I call this “regulative control”. The other kind of control does not require such access. It is a distinctive kind of control that does not involve freedom to choose or do otherwise; I call this “guidance control”. My claim is that guidance control is the freedom-relevant or control component of moral responsibility; thus, an agent can legitimately be held morally responsible for his behavior, even though he lacks regulative control (or freedom to choose and do otherwise).

To develop these notions of control (and their relationship), consider the following cases.\(^3\) Let us suppose that I am driving my car. It is functioning well, and I wish to make a right turn. As a result of my intention to turn right, I signal, turn the steering wheel, and carefully guide the car to the right. Further, I here assume that I was able to form the intention not to turn the car to the right but to turn the car to the left instead. In this ordinary case, I guide the car to the right, but I could have guided it to the left. I control the car, and also I have a certain sort of control over the car’s movements. Insofar as I actually guide the car in a certain way, I shall say that I have “guidance control”. Further, insofar as I have the power to guide the car in a different way, I shall say that I have “regulative control”. (Of course, here we are not making any “special assumptions”, such as that causal determinism obtains or God exists.)

Consider, now, a second case. Here I again guide my car in the normal way to the right. The car’s steering apparatus *works properly* when I steer the car to the right. But unknown to me, the car’s steering apparatus is broken in such a way that, if I were to try to turn it in some other direction, the car would veer off to the right in precisely the way it actually goes to the right. Since I actually do not try to do anything but turn to the right, the apparatus functions normally, and the car’s movements are precisely as they would have been if there had been no problem with the steering apparatus. Indeed, my guidance of the car to the right is precisely the same in this case and the first car case.

\(^3\) I take these cases from Fischer, *My Way*, p. 39, from which I borrow the following three paragraphs.
Here, as in the first car case, it appears that I control the movement of the car in the sense of guiding it in a certain way to the right. I do not simply cause it to go to the right (say, as a result of sneezing or having an epileptic seizure or involuntary spasm). Thus, I exhibit guidance control of the car. (I control the car, but I do not have control over the car’s movements). Generally, we assume that guidance control and regulative control go together. But this case (which has some of the salient structural features of a “Frankfurt-type case”) helps to show that they can at least in principle pull apart: one can have guidance control without regulative control.

The second car case should elicit the intuition that we do not need regulative control (genuine metaphysical access to alternative possibilities) in order to have the kind of control involved in moral responsibility. The second car case is rather like John Locke’s famous example of a man who is in a room which, unknown to him, is locked; the man thinks about whether to leave the room, but decides to stay in the room for his own reasons. The fact that the door is locked plays no role in the man’s practical reasoning. Locke says that the man stays in the room voluntarily, although he could not have left the room. Similarly, it seems that I exhibit guidance control of the car, although I could not have caused the car to go to the left.

But in Locke’s case the man did have various options available to him. After all, he could have decided to open the door, he could have tried to open it, and so forth; and similarly, in the second car example I could have decided to steer the car to the left, I could have tried to do so, and so forth. Some philosophers might then insist that it is in virtue of the existence of these alternative possibilities that the agent is morally responsible. And it must be conceded that we have not yet produced an example in which an agent is intuitively thought to be morally responsible and yet has no alternative possibilities (no regulative control). This is precisely where Harry Frankfurt adds an element to the cases.4 In Frankfurt’s examples, a “counterfactual intervener” stands by ready to intervene in the relevant agent’s brain processes, if he shows even an inclination to choose to do otherwise. Although Frankfurt was rather vague about exactly how the counterfactual intervener can succeed in expunging all access to alternative possibilities, Frankfurt’s followers have filled in the template in various ways.

Over the years, I have provided a sustained argument that the Frankfurt-examples show that moral responsibility does not require genuine metaphysical access to alternative possibilities (regulative

---

control). The leading idea in my argument is that any proponent of the regulative control requirement should say that the alternative possibilities in question must be robust, and not mere flickers of freedom. If the basis of moral responsibility is access to an alternative possibility, the alternative possibility cannot be any old possibility of something different happening; such a mere flicker of freedom would be too thin a reed to support the superstructure of moral responsibility. The situation here is precisely like the problem faced by proponents of indeterministic accounts of moral responsibility; how can the mere addition of a certain sort of alternative possibility—say an event the happening or not-happening of which is entirely arbitrary or accidental, from the agent’s point of view—render it true that the agent has the control associated with moral responsibility? I have asked the proponent of regulative control a similar question, “How can the mere addition of an alternative possibility in the Frankfurt cases—say an event the happening or not-happening of which is entirely arbitrary or accidental, from the agent’s point of view—render it true that the agent has the control associated with moral responsibility?” Note that the worry behind this question is exactly why the prominent libertarian philosopher, Robert Kane, has essentially agreed with me on this point, positing the “dual voluntariness” requirement for moral responsibility.

My claim, then, is that versions of the Frankfurt cases can be given in which it is very plausible to say that the agent in question is morally responsible for his behavior, and yet he has no access to the relevant sort of alternative possibility—a sufficiently robust alternative possibility. I made this point in my 1981 article, “Responsibility and Control,” which is the basis for “Responsibility and Alternative Possibilities,” Chapter Two of My Way.5 Recently, I have defended this fundamental point, as applied to both explicitly indeterministic and explicitly causally deterministic versions of the Frankfurt Cases.6

In addition to presenting a systematic defense of the contention that the Frankfurt cases show that moral responsibility does not require regulative control, I have pointed out that the rejection of the requirement of regulative control does not depend on the Frankfurt cases. There are various other routes to the same conclusion, including the Strawsonian contention that our ordinary responsibility practices do

---


not presuppose the requirement of regulative control.\footnote{For a version of this sort of strategy, see R. Jay Wallace, Responsibility and the Moral Sentiments (Cambridge, Mass.: Harvard University Press, 1994).} Also, Daniel Dennett has presented various arguments against the requirement of regulative control.\footnote{Daniel Dennett, Elbow Room (Cambridge, Mass.: MIT Press, 1984); and Freedom Evolves (New York: Viking, 2003).} I believe that the fact that there are various different routes to the same conclusion helps to establish the plausibility of the conclusion; if one finds thought-experiments such as the Frankfurt-cases unattractive, or if one finds the Frankfurt-cases unconvincing, there are still good reasons to accept that moral responsibility does not require regulative control.

So my preliminary conclusion is that if causal determinism rules out moral responsibility, this is \textit{not} in virtue of its eliminating regulative control (if it indeed does eliminate regulative control). This is an important point; I believe it is the “moral of the Frankfurt stories”, no matter how they are told and re-told. Further, if this point is correct, it allows us to side-step the traditional debates about the relationship between such doctrines as God’s omniscience and causal determinism, on the one hand, and “freedom to do otherwise” or regulative control, on the other. That is, we can sidestep these debates if we are simply interested in moral responsibility. Insofar as these traditional debates have issued in what I have called Dialectical Stalemates—black holes in Dialectical Space-time—avoiding them may open the possibility of genuine philosophical progress.

This having been said, I have never suggested that the mere fact that regulative control is not required for moral responsibility would allow us to conclude \textit{straightaway} that causal determinism is compatible with moral responsibility. Indeed, in my 1982 paper, “Responsibility and Control,” I emphasized that causal determinism might rule out moral responsibility \textit{directly} (and not in virtue of ruling out alternative possibilities). I thus identified what has come to be called “Source Incompatibilism”, and I pointed out that it must be taken seriously. I concluded that theorists of moral responsibility should focus on the \textit{actual sequence} that issues in any particular choice or behavior.

In subsequent work, I have identified and evaluated a number of different factors that might be invoked to explain why causal determination rules out moral responsibility directly. I have concluded that none of these factors provides a good reason to suppose that causal determination \textit{in itself and apart from ruling out alternative possibilities} is incompatible with moral responsibility. I believe that when one shifts from consideration of the relationship between causal determination
and regulative control to a focus on actual-sequence features of causally deterministic processes, the philosophical terrain becomes significantly more hospitable to compatibilism.

Given that I do not think that causal determinism rules out moral responsibility by threatening regulative control, and I do not think that there are other good reasons to suppose that causal determinism rules out moral responsibility, I present an account of moral responsibility that is compatible with causal determinism. More precisely, I present an account of “guidance control”, the freedom-relevant condition necessary and sufficient for moral responsibility.

An insight from the Frankfurt cases helps to shape the account of guidance control: moral responsibility is a matter of the history of an action (or behavior)—of how the actual sequence unfolds—rather than the genuine metaphysical availability of alternative possibilities. On this view, alternative scenarios or non-actual possible worlds might be relevant to moral responsibility in virtue of helping to specify or analyze modal properties of the actual sequence, but not in virtue of indicating or providing an analysis of genuine access to alternative possibilities.

Note that, in a Frankfurt-type case, the actual sequence proceeds “in the normal way” or via the “normal” process of practical reasoning. In contrast, in the alternative scenario (which never actually gets triggered and thus never becomes part of the actual sequence of events in our world), there is (say) direct electronic stimulation of the brain—intuitively, a different way or a different kind of mechanism. (By “mechanism” I simply mean, roughly speaking, “way”—I do not mean to reify anything.) I assume that we have intuitions at least about clear cases of “same mechanism,” and “different mechanism.” The actually operating mechanism (in a Frankfurt-type cases)—ordinary human practical reasoning, unimpaired by direct stimulation by neurosurgeons, and so forth—is in a salient sense responsive to reasons. That is, holding fixed that mechanism, the agent would presumably choose and act differently in a range of scenarios in which he is presented with good reasons to do so.

The above discussion suggests the rudiments of an account of guidance control of action. On this account, we hold fixed the kind of mechanism that actually issues in the choice and action, and we see whether the agent responds suitably to reasons (some of which are moral reasons). My account presupposes that the agent can recognize reasons, and, in particular, recognize certain reasons as moral reasons. The account distinguishes between reasons-recognition (the ability to recognize the reasons that exist) and reasons-reactivity (choice in accordance

---

The following three paragraphs are taken from *My Way*, pp. 17-19.
with reasons that are recognized as good and sufficient), and it makes different demands on reasons-recognition and reasons-reactivity. The sort of reasons-responsiveness linked to moral responsibility, on my view, is "moderate reasons-responsiveness."

But one could exhibit the right sort of reasons-responsiveness as a result (say) of clandestine, unconsented-to electronic stimulation of the brain (or hypnosis, brainwashing, and so forth). So moderate reasons-responsiveness of the actual-sequence mechanism is necessary but not sufficient for moral responsibility. I contend that there are two elements of guidance control: reasons-sensitivity of the appropriate sort and mechanism ownership. That is, the mechanism that issues in the behavior must (in an appropriate sense) be the agent’s own mechanism. (When one is secretly manipulated through clandestine mind control as in The Manchurian Candidate, one’s practical reasoning is not one’s own.)

I argue for a subjective approach to mechanism ownership. On this approach, one’s mechanism becomes one’s own in virtue of one’s having certain beliefs about one’s own agency and its effects in the world, that is, in virtue of seeing oneself in a certain way. (Of course, it is not simply a matter of saying certain things—one actually has to have the relevant constellation of beliefs.) On my view, an individual becomes morally responsible in part at least by taking responsibility; he makes his mechanism his own by taking responsibility for acting from that kind of mechanism. In a sense, then, one acquires control by taking control.

I ended my 1981 paper, "Responsibility and Control," by saying that we must "decode the information in the actual sequence" leading to behavior for which the agent can legitimately be held morally responsible and ascertain whether it is compatible with causal determination. The account of guidance control—with the two chief ingredients, moderate reasons-responsiveness and mechanism-ownership—are the secrets revealed by close scrutiny of the actual sequence, and I have argued that they are entirely compatible with causal determination. (Note that they are also entirely compatible with causal indeterminism; thus, on my approach, moral responsibility does not hang on a thread.)

Further, I have shown how we can build a comprehensive account of guidance control from an account of guidance control of actions. That is, we can develop an account of guidance control of omissions, consequence-particulars, consequence-universals, and perhaps even emotions and character traits by invoking certain basic ingredients contained in the account of guidance control of actions. I argue that it is a point in favor of my account of moral responsibility that it can give a comprehensive account that builds on simple, basic ingredients. Additionally,
I contend that this comprehensive account systematizes our intuitive judgments about a wide range of examples involving moral responsibility. It thus helps us to achieve a philosophical homeostasis, or, in John Rawls’s famous term, a reflective equilibrium.  

I. 3. My Way and Our Stories

In addition to presenting the motivation for the account of guidance control and the specifics of the account, I have sought to articulate the nature of the value we place on exhibiting guidance control (and thus so acting that we can legitimately be held morally responsible). I have thus pointed to an important connection between exhibiting the relevant kind of free will and the meaning of our lives. We do value acting freely (exhibiting guidance control), rather than (say) being a robot or even a conscious being controlled entirely by others. But what precisely is this value? If I am correct that we can exhibit guidance control without having regulative control, the value of exhibiting guidance control cannot be the value of making a difference (of a certain sort). Rather, I contend that the value of exhibiting guidance control—the value of acting freely—is the value of making a statement (of a certain sort). More specifically, my claim is that the value of acting freely is the value—whatever kind and amount that is—of a certain kind of artistic or aesthetic activity. On this view, when we act freely, we transform ourselves into creatures who have an irreducibly “narrative” dimension of value. I sketch the rudiments of this view in “Responsibility and Self-Expression,” and I seek to fill in the details in


There are certain features of my account of guidance control that a disconcerting cohort of (otherwise!) thoughtful philosophers have found rather less than irresistible, especially the subject element and the contention that “reactivity is all of a piece”. In the trio of articles above, I argue (among other things) that (if need be) I could adjust my account so as to do without these contentious features while still maintaining all of my major claims: that moral responsibility does not require regulative control, that causal determination is compatible with moral responsibility, that moral responsibility is an essentially historical notion, and so forth.
more recent work, especially the sequel to *My Way, Our Stories: Essays on Life, Death, and Free Will*.11

II. Some Notes

As I stated above, I believe that our status as morally responsible agents should not “hang on a thread”; more specifically, neither the discovery that the laws of nature have associated with them 100% probabilities nor the discovery that they have associated with them 99.9% probabilities should incline us to give up our views of ourselves as deeply different from other creatures insofar as can engage in practical reasoning and be morally responsible. Much of my work has focused on the deterministic side of the equation, so to speak. But note that my account of guidance control is entirely compatible with the falsity (as well as the truth) of causal determinism. Indeed, Carl Ginet has kindly suggested that an indeterminist should accept the core of my account of the responsibility-conferring kind of control (in my view, guidance control), and simply add a condition specifying that causal determinism must be false.12 My account then is compatible with causal indeterminism.

It is, as I have emphasized, a considerable advantage of my approach that it renders agency and moral responsibility resilient to certain (although not all) empirical discoveries. Indeed, I consider it an important desideratum of an adequate account of moral responsibility that it does not depend on any contentious doctrine. So, for example, if one’s account of moral responsibility depended on (say) the existence of irreducible agent-causation or the falsity of “reductionism” (of a certain sort) in metaphysics or the falsity materialism about the mind, this would be a strike against the account. Similarly, if one’s theory of moral responsibility depended on a particular view about reasons—their ontological status or their “logic” or even their specific content—this would count against the theory.

Although I cannot argue for these claims here, I contend that my accounts of guidance control (and moral responsibility) are compatible with a wide range of plausible views about these contentious empirical and philosophical matters. For example, my account of guidance control certainly does not presuppose that there is irreducible, indeterministic agent-causation; it thus does not depend, for its acceptance, on

---


some sort of defense of this highly contentious doctrine. On the other hand, I believe that the core of the account is compatible with the existence of irreducible, indeterministic agent-causation. As with Ginet’s suggestion above, there would perhaps need to be certain adjustments or clarifications; but there is nothing in the core ideas of the account that requires either the truth or falsity of claims about agent-causation. I would contend that the situation is similar with respect to the other contentious doctrines mentioned above. It is then a big advantage of my account that it is significantly resilient; it can be “nested” within total packages with a wide range of particular views about the nature of the mind, normativity, and the world. In contrast, a libertarian’s philosophical views—or his life—would be turned upside down, if in the future the scientists discover that causal determinism is true. At least I can sleep well at night. (Or, perhaps better: my recurrent insomnia is at least not caused by the fact that crucial and central features of my philosophical views hang by a thread.)

I have always sought to take incompatibilism seriously. That is, I believe that the strongest argument for a compatibilist conclusion stems from granting the incompatibilist as much as possible. Approaches that simply reject out of hand the most plausible parts of incompatibilism—such as the Consequence Argument or worries about manipulation—are not appealing to me. After all, the Consequence Argument employs ingredients—such as the fixity of the past and natural laws—that are deeply ingrained in our commonsense ways of understanding the world and our agency, and it has been around—in one form or another, including versions that pertain to the prior truth values of propositions or God’s omniscience—for millennia.

My approach is to take seriously—and, indeed, to seek to accommodate—the extremely plausible kernel of incompatibilism: that causal determinism rules out regulative control (freedom to do otherwise). Additionally, I take seriously the possibility that our mental states can be manipulatively induced; thus, I seek to provide an explanation of the difference between such manipulation and mere causal determination. Whereas many compatibilists either ignore or simply don’t take seriously the Consequence Argument or the possibility of manipulative induction of mental states, I seek to capture what is true and important about these worries.

My doctrine of Semicompatibilism thus seeks to have its cake and eat it too. Semicompatibilism is the view that causal determinism is compatible with moral responsibility, quite apart from whether causal

---

determinism rules out regulative control. (Semicompatibilism is officially agnostic about whether causal determinism does indeed rule out regulative control.) Thus, a Semicompatibilist might accept the conclusion of the Conequence Argument, but still hold that causal determinism is compatible with moral responsibility. Thus, a Semicompatibilist can sleep at night; he can accommodate the kernel insights of the incompatibilist but also embrace the attractive features of compatibilism, most notably, the resiliency of our fundamental views of ourselves.