Standpoints and Freedom

Pamela Hieronymi, hieronymi@ucla.edu

I am presenting a piece from the second chapter of a book manuscript, and so I will begin by briefly stating where we are in the book.

The book is about free will and moral responsibility, and its first two chapters are meant to isolate what I take to be the intuitive problem of free will.

In the first chapter I present difficulties that I think do not present vexing *philosophical* problems about freedom or agency: threats to our freedom posed by interfering agents, such as meddling neuroscientists, powerful gods, and oppressive political regimes, as well as hinderances to and defects of agency, such as diseases or drugs. All of these are real threats to freedom, ones we should do our best to avoid. But, although these raise important ethical questions, and sometimes difficult philosophical questions in ethics, I suggest that they do not pose any particularly vexed philosophical difficulty about freedom or agency, itself. They are, we might say, problems in life, not in theory.

I then contrast these threats to our freedom—interferences, hinderances, and defects—with the threat that seems to be posed by deterministic physics, or mechanism. Despite the metaphorical excesses philosophers sometimes indulge, determinism is not an interfering agent—it is not analogous to a powerful god or meddling neuroscientists. It is rather a scientific claim, a claim about how the world works, one which implies that the processes that underlie and explain our agency, the processes that underlie and explain the making and executing of our decisions, unfold strictly from earlier states. But notice that the processes that underlie and explain the usual operation of our agency could not be interferences with, hinderances to, or even defects of it.

And yet, it seems, when we think about the processes that underlie and explain our agency, and when we imagine that they unfold strictly from earlier states and events, we feel our freedom is threatened—in fact, we feel we are not *really* free at all. Moreover, as noticed by many, we feel the same intuitive threat if we imagine that our actions and decisions unfold entirely from earlier states and events in a merely lucky or probabilistic way.

This poses an especially sharp philosophical problem, in part because it is also the case that we cannot understand an event as an action, at all, unless we are able to explain it by appeal to psychological facts. To see an event as an action, we must see it as something that occurred because someone meant for something to occur. But the fact that someone meant for something to occur is a psychological fact. And, we—or, most of us—now believe that such psychological facts emerge, in their entirety, from the stuff of the earth: from nature or nurture, working in some contested combination, along with some luck. But, once we see our own actions as a part of the unfolding history of the natural world, a history that starts long before our decisions, long before even our birth, it seems to us that we are not free. And so we arrive at a vexing philosophical problem: we must see an event as explained by certain sources, to see it as an action, and we—or, most of us—think those sources are, in turn, entirely explained by prior worldly goings on. But, if our actions are entirely explained by those goings-on, we then feel we are not free—perhaps that we are not really acting, at all.

Many seek refuge in the insistence that the psychological emerges from the physical and cannot be reduced back to it.[1] But, I would argue, this fails to appreciate the strength of the intuitive problem. Shifting from neurons and chemicals to wants, desires, and beliefs, loves and commitments, fears and insecurities, self-esteem and jealousy, does not remove the worry. Loves and commitments, self-esteem and jealousy, are explained by prior states and events. Perhaps those explanations are not deterministic, but—again—probabilistic explanation is no less worrying. If some unfortunate soul, due to his or her formative circumstances, lacks the strength of ego or capacity for empathy needed to regulate his or her desires in more sociable ways, then, it seems, he or she cannot regulate his or her desires in more sociable ways. And whether she has the strength of ego or capacity for empathy is a matter of nature, nurture, luck, and his or her past choices. But his or her past choices are ultimately a result of nature, nurture, and luck. And, of course, just the same is true of each of us. If what you do is a function of what you are like, with or without some slippage, and what you are like is a product of what came before, with or without some luck, it seems, intuitively, that you are not free. And yet it seems undeniable that what you do is a function of what you are like, with or without some slippage, and that what you are like is a function of what comes before, with or without some luck. Thus, while it may be true that the human emerges from the physical in a way that defies reduction back to it, this will not, ultimately, assuage our concerns about our freedom. If we were bothered by Newton, Freud will do just as well.

---

[1] NTS: they might thereby deny the transitivity of explanation?

And thus I arrive at what I take to be the intuitive problem of free will—a philosophical problem about agency. It is this: when we explain free action, we seem to explain it away. The goal of this second chapter is to try to locate the source of this problem: why should focusing on the processes or forces that underlie and explain our activities make them seem unfree or unreal?

I am not alone in thinking that explanation poses a special problem for agency—I am typically joined, in this, by contemporary neo-Kantians. The standard contemporary response to the problem is what I will call "two-standpoints" compatibilism. My task for today is to explain this response and to explain why I find it unsatisfying, both as a diagnosis of the intuitive problem and as solution to it. I will end by saying briefly where I think the real source of the problem lies.

The two standpoints in question are typically distinguished by the activities undertaken from them: There is, on the one hand, a "practical," "deliberative," "first-person," or "subjective" point of view from which we decide and act, and, on the other, a "theoretical," "explanatory," "third-person," or "objective" point of view, from which we observe, describe, and explain. This distinction in standpoints captures the intuitive problem: when we occupy the first point of view, we take ourselves to be free. But when we occupy the second, when we reflect upon our agency and start to describe or explain it, we appear to ourselves, not as agents but as objects, and our actions appear as mere events—from this point of view, the stingy provisions of a step-motherly nature seem to curtail our possibilities. In fact, they seem to make our decisions for us. We seem to ourselves mere machines, pushed along by external determinants.

This same appeal to standpoints is also thought to ground a (to my mind peculiar) form of compatibilist response to the problem: When we occupy the second standpoint, our freedom does not appear. But, it is said, we are not entitled to conclude, from the fact that our freedom does not appear when we theorize ourselves as empirical subjects, that our freedom is only an illusion of the practical perspective. This illicit conclusion could only be reached by improperly privileging the theoretical point of view over the practical, when neither could be given priority. Even though the two points of view paint what seem to be contrasting pictures, we need not—in fact, cannot—choose between them. This is not worrying, because they concern different subject matters or conceptual schemes. The two points of view are, so to speak, *so* incompatible, that they cannot even be brought into genuine conflict. And thus we arrive at a peculiar kind of compatibilism.

I have just sketched, in bare outline, the two-standpoints approach. But notice, the outline requires filling in. *Simply* appealing to distinct "standpoints" is a compelling way to *describe* the intuitive problem. But, if we are going to do more than provide a gripping metaphor in which to state our problem, we need to know something about the two points of view—what constitutes them, why we occupy them, etc.—that might allow us to understand why they cannot be combined and so cannot genuinely conflict. And that further story might then help us understand why, when we explain our own agency, we seem to explain it away. Kant himself provided such a story, with his appeal to in-principle unknowable aspects of reality. But that is not a story that many, today, would embrace.

Notice, too, that a *mere* appeal to distinct "conceptual frameworks" or "levels of description" will not do justice to the intuitive difficulty. The intuitive problem is not the simple one that arises when we shift vocabularies or change aspects: Learning that music is explained by sound waves does not make us think music has disappeared, or that there is no such thing as "real" music, or that the music is no longer genuine. Learning that pain can be explained as neural and brain happenings has no tendency to make us think we do not *really* feel it (likewise with consciousness). In contrast, learning that (what we thought of as) agency is explicable by prior conditions makes us think there is no such thing. Our philosophical question is, Why should this be?

In a section I am cutting for time, I consider the very different labels often used, in contemporary discussion, to mark the two standpoints: "practical," "deliberative," "first-person," or "subjective," on the one hand, and "theoretical," "explanatory," "third-person," or "objective" on the other. I distinguish between (what I believe are) several distinct distinctions, and I try to show that, in each case, either the distinction does not track the apparent disappearance of agency, or, if it does, that is because we are have applied the labels by taking for granted an understanding of the intuitive problem—and so we will not illuminate the problem by appeal to a distinction between such "standpoints."

Here is one quick example from this cut section: Consider the distinction between the "theoretical" and the "deliberative" point of view. It seems ill-drawn. When I theorize about some subject matter—Newtonian mechanics, perhaps—I may well deliberate. Do I then leave the "theoretical" point of view? Surely not. So perhaps the intended distinction is really between the "theoretical" and the "practical"—the point of view of describing, explaining and understanding, on the one hand, and of decision making and acting, on the other. But, of course, in making my decisions—in deciding whether to take my umbrella, e.g., or whether to flip my omelette—I may also do some thinking about how the world works. When I do so, must I then leave the "practical" point of view, temporarily, and adopt instead the theoretical one, before returning to my practical deliberations? If so, what, exactly, is this point of view of decision-making? Do I occupy it only at the moment of decision? When is that moment? Or perhaps I enter the practical point of view whenever I consider what people call the "normative": what is good or right or required. But surely I can theorize about such things, without making any practical decisions—and I may do so while viewing actions as entirely explained by past circumstances.

By pressing such points, I argue that, to the extent that we can draw a distinction that tracks our intuitive problem, we do by relying on our understanding of the intuitive problem. And so the distinction cannot then be put forward as a diagnosis of it, and certainly not as a solution to it.

In the end, I do not think the intuitive problem relies on a distinction between standpoints or points of view—even though it arises naturally when we reflect upon ourselves. Rather, in the end, I think the source of the intuitive problem lies in the thought (or feeling) that our own wills are not in our control, and that this thought (or feeling) arises naturally when we reflect upon ourselves, due to our confusion about what controlling our own will would require. The goal of this chapter is to arrive at that diagnosis.

However, before moving there, I would like to spend more time thinking about the "two standpoints." I think we can do a better job identifying the "two standpoints," one of which has to do with decision-making and one of which has to do with explanation, and I do think that those two "standpoints" can sometimes come into conflict. By laying out them out more precisely, and by examining more carefully how they do and do not conflict, I hope to support my claim that we will find neither the source of nor the solution to the intuitive problem here.

THE TRUTH IN THE STANDPOINTS TALK: QUESTIONS, NOT POINTS OF VIEW

To begin, recall what I call[ed in the Introduction] *the ordinary notion of control.* When we think about what it is to exercise control, we naturally think of the control we exercise over ordinary objects, such as cars, coffee cups, and chairs, or the control we enjoy with respect to our own intentional actions, such as doing a back-flip or writing our name. These cases invite a certain model, according to which to control a thing is to be able to conform it to your will, or, less grandly, to be able to bring the thing to be as you would have it to be. Thus it comes to seem that, in order to control a thing (your handwriting or your future), you need to have in mind how you would have it to be, and you need to be able to bring it to be as you would have it. Crudely put, exercising control of the ordinary sort is a matter of representing some change and causing the change you represent.

It is clear enough why this ordinary notion of control, with its two-part structure of controller and object controlled, leads us to think of ourselves, insofar as we are agents, as a power to effect changes in the world. Notice, though, that it also allows us, in a certain way, to ignore ourselves us as we make a decision: When you control some object (your pencil or your pan), you must have in mind the object of your control, the change you mean to effect, and (somehow) the fact that you will effect it, but you need not have in mind the psychological operations by which you exercise control. The particular features of your will that will explain your decision can remain, so to speak, behind the lens, or out of view, as you decide. And thus we introduce the visual metaphor. You are occupying what it is natural to call a "first-person," "practical," "deliberative" perspective, looking out at the world, so to speak, *from* your will, from your own point of view, rather than considering your will as though from the point of view of another. When looking out from your will, you need not have in mind any of its features.[2]

But, of course, you are not barred from considering the features of your own mind, even from your own point of view. We are sophisticated and reflective creatures, and we can think carefully about our own wills. We can sometimes understand our motives. We can often explain why we did what we did—not only by appeal to those considerations we took to count in favor of acting, but also by appeal to those features of our minds that explain why we took those considerations to so count.

---

[2] So, when thinking of ourselves as agents of ordinary control, we must think of ourselves as a power to effect change, but we may not think of ourselves as more than that. NTS: careful. you argued in chp 1 that we must consider the features of a mind to see something as an action. here you are saying we need not think of agents as having minds with features. fit these together explicitly? When you, yourself, act, you need not think of the features. Must you, to recognize your own past or future action? It seems so.

But notice a relatively simple point: even if you fully understand the operation of your mind, even if you can explain your every thought and move, you cannot exercise control over your future *simply* by understanding, observing, describing, or explaining the operation of your own mind or will. To exercise control over your future, you have to make something like a decision. And, if you are going to make anything like a decision, you need to make it. No amount of observing, describing, or thinking about how the decision-making process is going to unfold will unfurl it.

TWO ROUTES TO THE FUTURE

We need to examine this last fact more closely. Notice, first, that determining what you *shall* do, in the sense of making a decision about your future, can be distinguished from determining what you *will* do, in the sense of making a prediction about your future. You might predict that you will lose the match. This is different from deciding to throw the match.[3] Both will leave you with what is, in some sense, the same view of your future: you will lose the match. But, in the first case, you come to this view by considering ordinary evidence—considerations that show it likely that your opponent will better you. In the second case, you do so by considering, instead, features of your situation that you take to count in favor of bringing about your own loss.[4]

Likewise, you might predict—in fact, you might know—that the neuroscientists of the last chapter (who have implanted a device in your head and are able to control your thoughts remotely) are going to send you out for a walk. This, alone, will not get you walking. If you are to go for a walk *intentionally*—if the neuroscientists are to get you to walk by controlling your mind, rather than just your body—then you will have to go for a walk because you meant to; you will have to decide to go for a walk. So, if they are going to make you to walk intentionally, then they need to make you decide. But predicting, believing, or even knowing, that you are going to make a decision is not the same as making it.

---

[3] See G. E. M. Anscombe, *Intention* (Oxford: Blackwell Publishing Co., 1957); Stuart Hampshire, *Freedom of the Individual* (New York: Harper and Row, 1965). Cf. also CITE Hampshire, Strawson

[4] [IF you do so for reasons, you do so for such reasons]

Thus there are, it seems, two routes, so to speak, to the conclusion that you will lose or that you will go for a walk: one route is occupied by predictions (in prospect, and explanations, in retrospect), while the other is occupied by decisions (in prospect, and (something like) justifications, in retrospect). You travel the first route by answering the ("theoretical") question of whether you *will* go for a walk—where that is a question you could ask about anyone (whether I will go for a walk, whether Luce will go for a walk, whether Rodney will go for a walk...). In settling this first question, you arrive at an ordinary belief, one which happens to be about yourself. The considerations you use to settle the first question (if you use any) will be those you take to show it *likely* that you (or Luce, or Rodney) will go for a walk. You travel the second route by answering a different question —not whether you will go for a walk, but, rather, whether to go for a walk. This second question is not, so to speak, about anyone,[5] and so cannot be asked about anyone else. It is, in some sense, essentially "first-personal." In settling this second question, you arrive at an intention to go for a walk. And whatever considerations you use to settle the second question will be—*in virtue* of your so using them—considerations you take, in some way, to count in favor of (or against) walking.[6]

Importantly, though, predictions and decisions routinely interact. Good decision-making often requires making predictions about yourself (whether you are likely to choke in the clutch or to forget your password). You might decide to throw the match because you predict you will lose it, anyway, and you would like to save your strength. You might decide to go for a walk because you believe the neuroscientists will make you walk and you would rather not wait around any longer.

---

[5] I am tempted to say, to answer this question is not to ascribe a predicate to a subject, and so affirm a proposition, but rather to commit yourself to effect some represented change in the world. ADD Thompson?

[6] This is not to say that you *believe* they count in favor of or against x-ing. It is rather that, in *using* them to settle, positively, the question of whether to x, by employing them in this way, you treat them as counting in favor of or against x-ing. AND comment about other forms of practical reasons (undermining, e.g.), and why they eventually come to bear on the practical question. THANK Wiland

Notice, again, that this entirely routine interaction of prediction and decision shows that the "standpoint" from which we make a decision can—and, in fact, ought to—avail itself of the "conceptual framework" of explanation. What is distinctive about the standpoint of decision-making is not concepts employed, but the question addressed.

So, while we might want to continue to use the metaphor of "standpoint" or "point of view," I suggest we understand the "two standpoints" by appeal to these two questions: the predictive question, of whether you will do something, and the practical question, of whether to do it. We can then consider whether addressing one kind of question, or answering it in a certain way, allows or precludes addressing the other, or answering it in a certain way.[7] Sometimes it will.

INTERACTING QUESTIONS

Sometimes, when we make predictions, we thereby change *which* practical question we ought to address. If I realize that I simply cannot beat my opponent—if I realize that, no matter what I do or how hard I try, I will not win—then I cannot sensibly address the question of whether to win.[8] I cannot sensibly address this question because I have realized that whether I win is not up to me in the following specific sense: whether I win does not depend on my decisions, planning, skills, or effort. And thus I cannot sensibly represent winning as a change I shall bring about. And so I cannot consider whether to bring it about. I should instead adopt what I will call the *fatalistic attitude* towards winning: I should set aside the question of whether to win, and instead address some other question, such as the question of whether to do my best anyway, or to give it my all, or, maybe, to decline to compete this round.[9]

---

[7] CITE literature in formal decision theory?

[8] I *could* address it, it is *possible* for me to do so, but I would be guilty of some error. Also, CITE Bratman on the simple view

[9] [NTS: Prof P takes the fatalistic attitude toward a decision that is his to make. That's his "bad faith." The inevitability point it is confused with it. Also: another illustration of the importance of the "theoretical" to decision-making.]

It is tempting to put this point this way: the fact that my loss is inevitable makes it unreasonable for me to address the question of whether to win.[10] But this is not right. It is not the *inevitability* of my loss that renders the practical question unreasonable. It is rather the fact that whether I win does not depend on my decisions, planning, skills, or effort.

This claim needs support. To start, notice that the fact that an outcome does not depend on my decisions, etc., is, by itself, sufficient to render addressing the question of whether to bring it about unreasonable. It will do so even if the outcome is *not* inevitable. Suppose I suffer from an illness from which I might, but might not, recover. And suppose that whether I recover does not depend, in any way, on my decision-making, etc. Even though my recovery is not inevitable, I still cannot sensibly address the question of whether to recover—because my recovery does not depend on my decision-making.[11]

Even so, one might think that, if an outcome *is* inevitable, that fact, alone, makes it unreasonable to make a decision about it. But notice how odd this position turns out to be: it claims that the fact that an outcome is inevitable makes it unreasonable to make a decision about it, *even if* the inevitable outcome depends on your (admittedly inevitable) decision. This position thus declares it unreasonable to do what you will inevitably do, simply because it is inevitable that you will do it. That hardly seems reasonable—after all, the thing you will inevitably do may well be, otherwise, the most reasonable option available. (Perhaps it is inevitable that you will take the more attractive offer.

---

[10] To be precise, it is not my *loss* that is inevitable: I could simply refuse to play, and so avoid the loss. What is inevitable is, rather, that I will not beat this opponent; I will not win. For ease of exposition, I will overlook this wrinkle.

[11] As outcomes become more likely the issue becomes more difficult. There is some discussion about whether I can decide to make my free throw. CITE. I suspect this example prompts disagreement in part because learning and even accomplishing a skilled action, such as a free throw, typically involves visualizing success (repeatedly). Visualizing success and then succeeding seems (philosophically, at least) similar to representing a change and bringing it about—and so similar to deciding. But we might want to distinguish visualizing the ball going through the net from deciding to throw the ball through the net. (The two will certainly have different Bratman-style conditions and will leave one open to different questions and criticisms.) In any case, for present purposes we need not determine the point at which the unlikeliness of success requires one to change the question one addresses, if one is to remain sensible.

According to this view, the fact that it is inevitable renders the decision to take the more attractive offer unreasonable. But that seems unreasonable.)

I suspect that what underlies the temptation to think that inevitability, alone, renders decision-making unreasonable is the thought that decisions, themselves, cannot actually be inevitable—and so, if an outcome in fact depends on my decision, then it is not really inevitable, after all.[12]

But, why think decisions are never inevitable? (Or, as inevitable as anything else we take into account, when making our way through the world.) We do not generally think so, when considering other people: you may think it is inevitable that your friend will decline the offer, or investigate the misbehavior, or insult the chair. It may be said, though, that you cannot have the same view of yourself: that you cannot think your own future actions are inevitable. But, again, I think this is simply not so. It may well be inevitable that I will accept a certain job when it is offered, or tell the truth in court, or attend to the needs of my child. (As noted in the last chapter, opening a decision to contingency does not render it more free or more my own.) And, if it is inevitable that I make a certain decision, I see no bar to my knowing that.

The two-standpoints theorist can make a ready retreat to more secure ground: whether or not a decision or outcome is in fact inevitable, and whether or not I can know that about myself (in a reflective moment), I cannot sensibly *see* it as inevitable, *as* I make a decision about it. I cannot see it as inevitable, she might say, from the "standpoint" or "point of view" of decision-making.

The visual metaphor again makes the point difficult to assess, but I think there is something to this thought. After all, when you address the question of whether to walk, for example, you are, it seems, addressing the question of whether *or not* to walk. The question you decide admits of two answers: yes and no. And so it might seem that, in addressing this question, you must, in some sense, take there to be two possibilities: you could settle it positively or negatively. And, further, if you settle, positively, the question of whether to walk, you should, and you usually will, work into the rest of your thinking and planning the fact that you will walk*,* while, if you settle it negatively, you should, and usually will, work into the rest of your thinking and planning the fact that you will not walk.[13] And so, when you address the question of whether to walk, it seems you are, in some sense, contemplating two contrasting futures, each of which depends on the outcome of your decision. Thus, it might seem, to address this practical question you must treat the future as open. And thus, it might seem, so long as you continue to accept the inevitability of a given outcome, you cannot sensibly address the question of whether to bring it about.

---

[12] I am grateful to Daniela Dover for commentary on this point.

As compelling as this last thought seems, it is not correct. First, it is not obvious that, in order to settle the question of whether to do something, you must, in any robust sense, contemplate or entertain the possibility of not doing it. Nonetheless, for the sake of argument, let us grant that, when you address the question of whether to walk, you are contemplating two contrasting futures. Let us also remind ourselves that you are also, in some way, acknowledging that which future is realized depends on your decision. In contemplating the two scenarios, you are considering whether or not to bring about some change. You have not *yet* decided the question. However, in the case we are considering, you also you believe, of yourself, that you will certainly decide the question one way rather than the other. There is no bar, it seems to me, to contemplating a future that *would* occur if you *were* to make the decision you believe you will certainly not make, nor any unreasonableness in doing so.[14] There is certainly nothing contradictory about doing so. You would, of course, be guilty of the unreasonableness already considered if you thought that the outcome will come about *regardless* of your decision, and yet proceed to make a decision about it. But you do not so regard it.

---

[13] CITE Bratman

[14] One might say it is unreasonable simply because it is a waste of time: you already know what you are going to do, so why contemplate this other possible world? Whether it is a waste of time depends, I think, on whether you know also *why* you are going to do it—whether you have your reasons for action at hand. If your secure prediction leaves your future reasons opaque to you, then you may well contemplate the two futures as you make your decision. That may be the way in which you find your reasons. But if you already have your reasons at hand, then it may be a waste of time to contemplate the future that you already believe clearly inferior. That is why I said, at the start of the paragraph, that you may not need to contemplate the alternative future.

And so, I think, contemplating the two contrasting futures while addressing the practical question can sensibly be done even while continuing to believe that one of the outcomes is inevitable.[15] Some illustrative cases have already been mentioned: I may know, in advance, that I will decide to accept a certain long-desired opportunity, tell the truth when asked, or care for my children. This does not prevent me from making my decision.

One might resist by replying that, if I *know* that I will certainly accept the opportunity or tell the truth, that is because I have *already* made the decision to do so. In any such case, my knowledge of my future action is *practical* knowledge, built on my decision. And, the opponent might continue, in advance of such practical knowledge, I cannot know what I will decide.[16]

While it may sometimes be true that I decide, far in advance, to accept the opportunity if offered or to care for my child, I doubt we must or should understand all such cases in this way. It seems possible, after all, to make predictions about your own decisions. Suppose that, while speaking to my therapist, I consider what I will do on the witness stand, and thereby come to see (what is plain to you and to him) that it is inevitable that I will tell the truth. Coming to this conclusion (or even doing so sensibly) does not require that I have *already* made the decision.[17]

If I then, later, turn to address the question of whether to tell the truth, must I, to be reasonable, expunge from my own mind what I have learned about myself? Must I suspend or revise my prediction, in order to make my decision?

---

[15] The view of the potential futures thus provides a different, and contrasting, picture, or point of view, than the view of the predictions. Here we might again want to talk about "two standpoints." But these two standpoints do not employ different concepts, and they are not in principle incompatible—in fact, it is an important part of planning to be able to reconcile them, to be able to plan to do what you see is possible. Nonetheless, in some situations they can be brought into problematic tension, as we will see below.

[16] THANK Greg for highlighting practical knowledge

[17] Perhaps I can takes steps, between now and then, to change things so that I will not: perhaps I can set up other, overwhelming incentives that I predict will motivate me to lie. My ability to do this does not undermine the relevant claim that my decision is inevitable EXPLAIN.

At just this point, the opponent may, again, appeal to standpoints: She may say, I need not suspend or revise the *prediction*, but I must, rather, enter a different point of view. From the predictive, third-person, or theoretical, point of view, I might believe that I will certainly tell the truth, but when I turn to make my decision, I adopt the practical, or first-person point of view, and I cannot continue to believe that.

But, again, I do not see what we gain by appealing to standpoints (other than some unclarity). The fact that the two *questions* are distinct, and that answering the predictive question will not, itself, amount to making a decision, is enough to do the work we need done, without restricting us further than seems real. In the case we are considering, I have settled the predictive question, but I have not yet settled the practical question: I believe (from my own point of view) both that I will certainly tell the truth and that I will tell the truth because I will decide to do so. This may be true, even if I have not *yet* decided to do so. And, as noted, no amount of predicting, nor any degree of confidence in a prediction, will simply amount to decision. And so I still have work to do. I have to get to the business of deciding. But I do not see why, in order to do *that*, I must enter anything like another "point of view" or "standpoint." I must, instead, address the practical question. And, again, I see no conflict in addressing the practical question while maintaining, "in view," my firm conviction in my prediction.[18]

To close out this point, let me consider an especially extreme case, by returning to the neuroscientists. In thinking about this case, though, it is important to remember two things. First, I am not, at the moment, wanting to make the stronger claim that inevitability is no threat to *freedom*. I am rather wanting to make a narrower point: that believed inevitability—a confident prediction—does not, by itself, render addressing and answering the practical question impossible or even unreasonable. Second, I have already granted that the meddling interferences of the neuroscientists

---

[18] CITE Martin Luther case. It is typically cited as a case in which I am faced with an inability on my part—so called "volitional necessity." But weaker cases will make the point I am after: inevitability is no bar to sensible decision-making.

are a genuine threat to freedom. But, again, I am employing them, here, to address a much narrower question: once more, whether a confident prediction will render addressing and answering the practical question impossible or unreasonable.

With those caveats in mind, recall that the neuroscientists are going to send me for an *intentional* walk—that is, they are going to control my thoughts, not just my body. Now suppose that I know the scientists will make me walk, and, further, that I have no objection to walking. Suppose I even think, all things considered, I ought to walk each day. Perhaps I have asked the scientists to make sure that I get out for a good walk today, and I know they both can and will. Perhaps they have told me they will send me walking at 9:23, and I believe them. I look at the clock. It is 9:23. I think, "Shall I go for a walk?" and answer, "Sure." And out I go.[19] In addressing the question of whether to walk—even the question of whether or not to to walk—I need not ever doubt that I will walk. And yet, I claim, I proceeded sensibly.

Here ends my attempt to support the claim that believing an outcome inevitable need not render a making a decision to bring about that outcome unreasonable—so long as the outcome depends on the decision.

It will be noticed that I have thus far focused on (what I will call) the happy cases, cases in which what I regard as inevitable and what I would have myself choose align. One might reasonably wonder: What of the unhappy cases? What if I am in the dark story in which the neuroscientists will make me decide to do something I despise? Or perhaps the Oracle tells me that the Fates have determined I will kill my father. Or maybe I simply know, of myself, that I will not follow through on my decision—as well intentioned as I am now, I will procrastinate or give into temptation. I think I should let my child cry, but I know I will not. I know I should complete the review in a timely way, but I know I will not. Surely, it will be said, in cases like these, my firm prediction will in some way interfere with my decision-making.

---

[19] I believe that I have just described the prayer life and subsequent decision-making processes of many. Other people sometimes employ life coaches. Others have elaborate strategies for ensuring that they will make certain decisions. And, when the time comes, they make the decision [and it is not true that they made the decision in advance, cf. note X].

I will next consider the unhappy cases in some detail. But my final position will be this: if believed inevitability is not a problem in the happy cases, then believed inevitability does not, *in itself,* make it impossible or unreasonable either to address or to settle the practical question. Rather, I will now suggest, in the unhappy cases, certain *sources* of inevitability present hinderances to or interferences with agency. But these hinderances and interferences are, once again, problems in life not in theory. They can pose serious ethical problems, and serious problems for ethical philosophy, but they are not themselves vexed philosophical problems for agency itself. (And so, if we want to understand the intuitive problem of free will, we will again have to look elsewhere.)

In examining the unhappy cases, in which what I confidently predict I will choose is not what I would have myself choose, let us start with a very simple case, one that poses no threat to freedom at all. Suppose you are again subject to the neuroscientists, and suppose you confidently predict they will send you walking at 9:23. 9:23 comes, but you do not want to walk. You face the decision, and you think, "Nah, not right now." In that case, you will not walk (or, at least, you will not do so intentionally). The neuroscientists will have failed. In this case, the fact that you did not, at 9:23, want to walk, and so decided against walking, shows that your confident prediction was inaccurate. Such a case poses no problem for your freedom, but it also poses no threat to my claims. My claim, again, is that confident prediction about a particular outcome does not render decision-making about that outcome unreasonable, so long as the outcome depends on the decision. I have also been assuming that decisions, like any other event in the world, can be inevitable. But I need not assume either that the neuroscientists are omnipotent or that you are infallible in your confident predictions. So, it may be that they, or you, get it wrong sometimes.

Let us turn, next, to the familiar example of Professor Procrastinate. Procrastinate is asked to review an article within a given time frame. To agree to complete a review on time is to commit to a plan of action that will require a number of other decisions along the way.[20] And we can assume that, if Procrastinate made the right subsequent decisions—if he put away his book, avoided making another cup of coffee, spent less time surfing the internet, etc.—he would certainly succeed. So, unlike our ill-fated competitor, whether Procrastinate succeeds depends (and, we can suppose, entirely depends) on his own decision-making. His difficulty is that he can predict (and, we are granting, can accurately predict) that he will not make the required decisions, when the time comes. How should Procrastinate proceed? What decisions can he sensibly make?

First, and importantly, notice that, unlike in the case of the ill-fated competitor, it seems sensible for Procrastinate to *address* the question of whether to complete the review—in fact it seems he is *required* to address that question and return an answer to the editor. Whereas the ill-fated competitor could not sensibly address the question of whether to win, since winning did not depend on her decisions, etc., Procrastinate *must* address it—because the outcome does depend on his decision. But Procrastinate knows as well as we do that he will not write the review, even if he accepts the invitation. So how is Procrastinate to answer the question he must address?

Like any of us, Procrastinate cannot sensibly agree to write the review unless he can be reasonably confident that will do it. Most of us are entitled to that confidence, without first drawing up elaborate plans: we can reasonably count on ourselves to sort it out as we go. But, given Procrastinate's poor record on such things, he cannot sensibly proceed in this way. If he did so, then, in light of his track record, he would be guilty of bad planning. Nor can he sensibly decide to rely on strategies that have failed in the past. And so Procrastinate needs to have in mind some reasonably detailed plan, in which he can have confidence. Notice, though: if he has such a plan, and if his plan is tolerably reasonable, then, it seems, it is no longer inevitable that he will fail, and he can sensibly accept the request.[21]

---

[20] CITE Michael Bratman's pioneering work.

[21] [CITE Beri's work about the "reasonable"]

But what if Procrastinate finds himself unable to come up with any reasonable plan? What if he continues to regard it as inevitable that he will fail? He cannot, then, sensibly agree to complete the review. Can he sensibly decide to decline it?

It seems problematic, in some way, for Procrastinate to decline the review because he regretfully predicts that he will not complete it, if, as we have stipulated, whether he completes the review depends entirely on his own decision-making, planning, efforts, etc. It seems to be in some way in bad faith. In fact, I think there are at least two different problems, what might be thought of as two different kinds of bad faith, to be distinguished.

First, Procrastinate would be in bad faith if he treated the prediction, *itself*, so to speak, as settling the practical question. As we have noted several times now, the predictive question and the practical question are distinct, and answering one does not simply amount to answering the other. So, your confident prediction that you *will* walk is not yet a decision *to* walk, and Procrastinate's confident prediction that he will not complete the review is not yet a decision to decline it. One form of bad faith—what, it seems to me, Sartre had in mind—would try to ignore this distinction and allow the prediction just to stand in, so to speak, for the decision. But this will not do. If you walk intentionally, you will walk because you mean to walk—and so you will need to settle for yourself, positively, the question of whether to walk. Likewise, if Procrastinate declines the review, he will have to settle, positively, the question of whether to decline it. He cannot avoid that decision.

So a prediction cannot simply stand in for a decision. Nor does a prediction, simply by itself, bear on a decision. But you might take it so to bear. You might take your confident prediction that you will walk to bear on the question of whether to walk.[22] Perhaps you are impatient and would rather avoid waiting around: might as well get this over with. Or maybe you think walking, now, will somehow help you to retain some sense of control over your own future. Or maybe, since resistance is futile, you would simply like to save your strength. Alternatively, maybe you think it is a reason not to walk—since it shows the walking unfree—but, in the end, other needs won out. So while a prediction cannot simply stand in for a decision, it can become one consideration (typically one among many) in light of which you decide.

---

[22] You will then take it, in some or another (perhaps indirect) way, either to count in favor of or to count against walking Why will I take it to so bear, at least in some indirect way? Because any consideration that I employ, in answering the question of whether to act, must ultimately, in some or another way, come to bear on that question, and so ultimately, in some or another way, be taken to either count in favor of or count against walking. Thanks to Eric Wiland for comments on this point.

Just the same is true for Procrastinate. His prediction, that he will not complete the review, cannot simply stand in for his decision. But it is relatively easy to see how he might take the prediction to count in favor of declining the review. He might reason, "I am sure not to complete the review, so it will be best for everyone if I decline now." In so deciding, Procrastinate has not simply treated his prediction as if it were a decision. He is making a decision, one for which he can be held, and hold himself, to account. So a charge of one form of bad faith [the form that, I think, goes with the charge of self-deception] will not stick.

But his decision can still seem problematic. He is still deciding not to complete the review because he foresees he will not, and he takes the fact that he will not do it to count in favor of not doing it. What is the remaining problem?

Here is one very tempting way of answering (a way that I used to endorse, but now think is mistaken): The question Procrastinate is addressing is the question of whether he *shall* complete the review, and he cannot, in addressing that question, treat as given the fact that he *will not* complete it —because whether he will complete it is precisely the matter under consideration. Likewise, the question he is addressing (whether to complete the review) takes into consideration the subsequent decisions required to do so—in deciding whether to complete the review, Procrastinate is also, therein, deciding whether to do what is necessary to complete it (whether to set aside the required time, whether to get up early, etc.). And so he cannot, when addressing this question, treat as given the fact that he will not do what is required. That is, again, precisely what is under consideration.

As tempting as this response is, I now think it is not right—again, because of the happy cases. In those cases, there is no difficulty with taking your own future decisions to be inevitable—in fact, there is no difficulty in taking the inevitability of your future decision to act to count in favor of deciding so to act. I may be confident that I will care for and attend to my children, and that confident prediction may be part of my reason for deciding to adopt children; or I may be confident that I will relentlessly pursue justice, and that confident prediction may be part of my reason for accepting a certain challenging job. So it seems to me that the difficulty with Procrastinate cannot be that he treats his future decisions as inevitable, nor even that he takes the inevitability of his future decisions into account when making them. Rather, the problem, I suggest, is simply that he, himself, regards the future decisions he plans to make as poor ones, even as he plans to make them. Or, rather, to put the point in a cleaner way [23]: the problem is that he takes the fatalistic attitude towards his own future decisions and starts to plan around them.

But why is that a problem? Return to the thought that, when you decide on some course of action, you are contemplating a future. You are also committing to a plan, a plan that might include a range of sub-decisions.[24] As you make the decision to complete the larger task or project, you are also, therein, committing to make those needed decisions along the way (that is why, in light of his past failures, Procrastinate cannot agree to complete the review without a reasonable plan). So, in committing to his plan, Procrastinate is committing to make the required decisions along the way. Those decisions are, then, in some sense, included in the decision he is making now. So if Procrastinate, in deciding to decline, is treating his future procrastinating decisions as facts to plan around, he is in some way treating those decisions as though they are not up to him. But, we have agreed, they are up to him. So he is in some way incoherent. It is as though he is counting on something or someone other than he—or, other than the he now making the decision to decline—

---

[23] NTS: "poor" might drag you into guise of the good. But that should be needless.

[24] CITE Bratman.

to shore up the decision to procrastinate, when the time comes. Something or someone else must, so to speak, hold those future decisions in place, or explain them, to make sense of his decision-making now. And that other something else, whatever it is, is a threat to his freedom, in the sense that it is an *interference with*, *hinderance to*, *constraint on*, or *defect in*, the operation of his agency.

What, then, is a procrastinator to do? In the end, I think that, if Procrastinate really cannot find a plan that would enable him to have any confidence that he will complete the review, then he ought to engage in this lesser kind of bad faith—he ought to plan around his own regrettable decisions, fatalistically, in the same way that the dominated opponent must plan around her loss. But, whereas the dominated opponent works around her loss because it does not depend on her own decision-making, Procrastinate is working around what is, everyone agrees, up to him. And so he is treating his future his self as though he were another (and unreasonable) person. (It is clear why it is tempting to call this taking up a "third-person" point of view on yourself. It is also clear why it might seem an evasion of responsibility—how are we now to hold you responsible for these future decisions you now disavow?) It is a bad position to be in. But the problem is not exactly one of self-deception or inauthenticity (as it would be, if he pretended that the prediction settles the matter). Procrastinate may be vividly and accurately aware of his predicament, and he may be doing everything he can to figure out to do, to take responsibility for it. It is rather a problem of disunity —and it is a *defect* of agency. This is not a problem posed simply by the inevitability of a future decision.

Much the same can be said, and, I think, in the case of the Fates and the Evil Neuroscientists, where our hero faces, not a defect of his own agency, but rather external manipulation and interference. [I will skip these, for time, luckily].

CONCLUSION

So, what have we accomplished? Our question was: why does explaining agency seem to explain it away? Why does agency seem unreal, once we learn that it is natural? Some have thought that we can explain this by pointing out that the "standpoint" of explanation and the "standpoint" of decision-making are not only distinct, but so different as to make it the case that they cannot be brought into conflict.

I have suggested that the distinction between standpoints is best understood as a distinction between questions: between a predictive question and a practical question. I have insisted that answering one question does not amount to answering the other, and I have suggested (though I have not directly argued) that this fact is all we need to understand the apparent difference in "standpoint" or "point of view." I have examined how and why answering the predictive question with certainty can make it unreasonable to address or to answer the practical question. Unreasonableness appeared only in what I called the unhappy cases. In these cases, the "point of view" of prediction and the "point of view" of decision do part ways. But these are also, I have argued, cases in which you see your agency as subject to some hinderance, interference, or defect.

Thus we have not, I think, yet found anything to help us understand why, when we explain our own free actions, we seem to explain them away. [We could think we were just confused, but this is not satisfying]. And so I think our original problem remains. I believe I can say what it is. I think it is not, in the end, a problem with standpoints or points of view, however rich and important these ideas are. It is, rather, a somewhat simple problem about our ordinary notion of control—our ordinary notion of control will not allow us to see our own will, our own decision-making (or concluding, believing, caring) as in our control. But if our decision-making, concluding, believing, and caring are not in our control, then it seems that nothing really is. That is the problem I propose to address in the coming chapters.