

## FREE WILL AND RATIONAL COHERENCY

Patricia Greenspan  
University of Maryland

Philosophers often picture undetermined action on the model of Epicurus's random swerves of atoms. For an agent acting rationally to do otherwise than she actually does would mean swerving away from the course prescribed by her own preferences and values. As Hume famously argued, undetermined action would lack the kind of connection to an agent's character or motives that we need for ascriptions of moral responsibility. In contemporary terms, whether or not one did the right thing would be a matter of chance or luck.

Current authors mainly accept Hume's point, if not as an argument against indeterminist free will, then as a constraint on the form it can take. Besides contemporary versions of agent-causation, which attempt to make out the agent as a special kind of cause of free action, a strategy exemplified by Robert Kane's event-causal version of libertarianism takes free action as based on resolving a practical conflict in a way that *shapes* the agent's character or motives.<sup>1</sup> Doing something else in light of the same reasons would make sense on Kane's account only where prior deliberation was insufficient to decide the issue. In making a choice the agent would then be deciding to stress certain of her reasons, to weight them more heavily than competing reasons, presumably with implications for future choice as well. For it would seem to be rationally incoherent, an incomprehensible practical swerve, to assign a reason greater weight only on a single occasion, possibly *against* the results of prior deliberation.

I think a practical swerve can make sense rationally, however, as long as it rests in the right sort of way on assigning weights to competing reasons. On a conception of practical reasons that gives the agent a role in establishing their weights, we can allow for rational reconsideration of a prior deliberative conclusion, not on the basis of new information, but just at the agent's option. According to my own preferred account, the normative

role of practical reasons is to offer or answer criticism of action, and an agent is in a position to discount criticism of her own, possibly just on a single occasion. If her decision to discount is to some extent arbitrary, that need not raise problems of rational control, assuming that it occurs within an overall structure exhibiting rationality.

That, in highly compact form, is the view I want to propose here. It turns on recognizing an *active* component in deliberation. The role of an agent in assigning weights to her reasons—*weighting* them, rather than merely weighing them—was defended at one point by Robert Nozick, though Nozick held that the weight assigned to a reason had to be taken as setting a precedent for future choice.<sup>2</sup> A different sort of contrast is to Alfred Mele's recent work on event-causal libertarianism, which allows for indeterminism at two stages in the production of action. One is passive: the stage at which reasons occur to us during deliberation, where full control is not to be expected. The other is active: the stage of action on considered judgment, where doing otherwise would be irrational, an instance of *akrasia*.<sup>3</sup> On Mele's account an agent has sufficient control as long as she learns to minimize *akrasia* over time, shaping the probabilities of her future options. But I think we can allow for doing otherwise, on a libertarian account, if we instead make room *within* deliberation for an undetermined act of the sort that Nozick had in mind: a decision about how to weight one's reasons.

In short: I agree with Kane, Nozick, and Mele in taking some version of self-shaping as the source of a free agent's rational control, but I also want to apply it to the way an agent shapes her own deliberation at a given time. At the stage of deciding to accord greater weight to certain reasons an arbitrary decision may be unproblematic, with no requirement that a rational agent go on to decide similar cases in the same way, or even that she learn to minimize practical inconsistency in her future choices.

I do recognize other grounds for consistency constraints on rational re-deliberation. Most obviously, it cannot be done too often, if an agent is to pursue effectively any of her practical aims. Also, on the view of reasons I favor, it cannot extend to reasons of the sort that yield moral requirements, understood as offering criticism from the standpoints of other agents. But the focus of my argument here will be on nonmoral cases. I hold that an agent can sometimes decide to discount a reason without irrationality and that part of what some of us want from free will is just that ability to shape our deliberation, deciding how to decide. Since many current authors apparently think free will has a point only as a requisite of moral responsibility, I begin, in Section 1, with a general discussion intended to convey my somewhat different take on the issue. My argument in defense of the rationality of practical swerves based on active deliberation will be postponed until Section 2.

It should be evident by now that I speak of freedom, free will, and free action more or less interchangeably. Similarly for "choice" and "decision."

I interpret decisions as mental acts. It might also be useful at this stage to outline some of the basic assumptions I make about reasons—meaning practical reasons, reasons for or against doing something. I understand reasons themselves as facts, not mental states. I am working from an “objective” notion of reasons, that is, according to which they are independent of an agent’s beliefs. The fact that smoking causes cancer, for instance, is a reason for a certain agent not to smoke, whether or not she knows it. Indeed, I take it that she “has” that reason whether or not she has any tendency to be motivated by it (whether or not she cares about her health, say). That classifies me as an “externalist” about reasons, in current terminology.<sup>4</sup> Since my concern here is with the influence of reasons on choice, though, I generally take it for granted that the agent to whom I ascribe a reason recognizes both the relevant fact and its normative bearing on her intentions and behavior.

However, according to what I want to argue, accepting a reason in this sense does not require according it any weight in deciding what to do. This means that I would deny that there are external reasons as usually conceived: reasons that necessarily motivate any rational agent. It is enough to say that ignoring certain reasons (serious reasons of health, say) would be wrong-headed, contrary to one’s interests and hence a violation of ideal rationality, whether or not irrational in the narrow sense, of “rationally incoherent,” which my discussion here presupposes. Note, though, that on my account an agent’s influence on the weights of her reasons is not merely a reflection of her beliefs or desires, but instead depends on her decision to discount certain reasons. It operates in the way that promising does, via an act of practical commitment. If we take the pre-given weights of an agent’s reasons in the nonmoral cases I focus on here to be a function of her interests, discounting comes out as a decision to *waive* certain interests. Within broad limits, I maintain, this is something an agent has rational authority to do: what we commonly call “setting priorities.”

Some readers will find my talk of “deciding how to decide” familiar from current work on “bounded rationality.”<sup>5</sup> Despite overlaps, however, and a general affinity for that approach, what I have in mind here is somewhat different. Bounded rationality involves adopting decision-making heuristics appropriate to a choice situation of limited time, information, or cognitive resources. But on my account an agent determining the weight of her reasons is not attempting to *estimate* their weight according to some independent measure that may be hard or impossible to apply. She is actively *endowing* them with weight—or in the first instance, depriving opposing reasons of weight, by discounting them.

I use the term “discounting,” moreover, without reference to work in economics on a “discount rate” in our estimation of value across time.<sup>6</sup> A fuller treatment of discounting reasons might recognize lesser degrees than complete cancellation, but what is in question here is a decision to discount that need not be based on distance in time. In short, my defense

of discounting is not put forth as a way of accommodating our limitations as agents, cognitive or motivational—or even informational limitations or other features of the surrounding situation—but instead as a basis for active deliberation, a manifestation of our freedom to choose.

### 1. Valuing Freedom in its Own Right

Let me at this point postpone further discussion of reasons until after a somewhat broader discussion of free will. Why should I care if my acts all turned out to be determined and on that account (let us suppose) unfree? The usual thought is that I would then get no credit for my achievements, or for doing the right thing. Freedom is valued as a requisite of moral responsibility, understood as what makes us liable to praise or blame. However, I think we can detach free will—in the libertarian sense of indeterminist free will—from responsibility and still find reason to value it. Some recent authors have tried to spell out further reasons why we might want to be free in the libertarian sense, and at least some of these are reasons why we might want freedom in its own right, even if we were satisfied with a compatibilist account of responsibility.

There is a term in the literature that would fit what I have in mind here: “semi-compatibilism,” which John Fischer and Mark Ravizza apply to their own view.<sup>7</sup> Semi-compatibilism allows for two senses of freedom, one amounting to the freedom to do otherwise, the other required for responsibility. But Fischer and Ravizza immediately dismiss the first sense as irrelevant to practical reasoning. The view I mean to suggest instead might be thought of as *robust* semi-compatibilism, since it does assign value to freedom in the first sense, the sense that entails having options.

However, since my emphasis is on the incompatibilist component of semi-compatibilism, it might be better to represent the view as a form of libertarianism: libertarianism “proper,” one might say, since libertarianism in the first instance is a position on *liberty*—free will—despite the tendency of many philosophers to speak of free will more or less interchangeably with moral responsibility or to conjoin the two notions as if they necessarily went hand in hand. I think we could conceivably work out an adequate basis for moral responsibility that is compatible with determinism but would not supply all that I want from free will.

I do not plan to argue for an account of moral responsibility in this paper, but let me explain why I approach freedom in different terms. In a nutshell: I take moral responsibility to be primarily a backward-looking notion, whereas freedom in the sense I intend is forward-looking. In holding someone responsible we look back in time from an act to its intentional cause, as an appropriate object of praise or blame, whereas in assessing someone’s freedom we focus on her power to affect the future<sup>8</sup>. Both notions, moral

responsibility and freedom, can be understood as having degrees, but we might conceivably want more freedom, or a further form of freedom, than we need just in order to take credit for what we do.

Detaching freedom from responsibility might seem harder for cases involving blame. Blame is typically unwelcome to its object, so it seems unfair to subject an agent to blame she cannot avoid. But even if so, what it takes to ground a minimal degree of blame might be less than we need for a reasonable expectation of self-control. I once argued a version of this point for psychopaths, who apparently lack the emotional resources normal agents can rely on to inhibit anti-social urges.<sup>9</sup> Psychopaths are a complicated case—most philosophers focus on the question whether they satisfy knowledge conditions on responsibility, rather than conditions involving freedom—but they can be seen as extreme examples of the kinds of agents we refer to as “incorrigible,” using a term that conveys a degree of blame, but also the practical inefficacy of the corresponding demand. Though the urges psychopaths act on do not seem to be overpowering, and they may have access to other means of control, learning to control their urges reliably would seem to be harder for them than for normal agents. Assuming that it is hard enough to be unreasonable to expect of them, at any rate in the short term, but that they do what they do intentionally—and for reasons of their own, not as a result of external manipulation or internal constraint, and so on (fill in your favorite compatibilist conditions)—it makes sense to count them as unfree but morally responsible. They are liable at least to a minimal degree of blame for intentionally inflicting harm on others, even if their motivational deficit puts a limit on *how much* we should blame them.

There is more to this story, but the treatment of psychopaths and other abnormal cases is not essential to my purposes here. If determinism is in question, the question is not whether doing otherwise would be too hard to expect of someone, but whether it is possible at all, as it would not seem to be if our acts all resulted from deterministic causal chains extending back beyond our agency. I take it that an intentional agent—someone whose acts exhibit good or ill will, in P. F. Strawson’s terms—can still count as appropriate object of praise or blame for what she does, even if we also manage to identify earlier causes, including some that make other agents responsible for it.<sup>10</sup> This is to say that moral responsibility, in the sense involving liability to praise or blame, may stop short of “ultimate” responsibility, the notion that Kane uses to sum up what we want from libertarian free will. But just because ultimate responsibility is backward-looking, it does not capture all I want from free will. As Kane also puts it, in forward-looking terms, I want my acts to be “up to me”—which I take to mean that I have the power to decide whether to perform them.

Mele speaks of “initiatory” power, understood as a form of independence of the past.<sup>11</sup> He defends this just as an indeterministic element of freedom that some of us reasonably value—value in ourselves, that is, since it affects

the importance of our individual lives and choices by giving us a particular kind of causal bearing on our actions—though Mele automatically applies what he says here to moral responsibility.<sup>12</sup> It is not so obvious, however, that agents who want to have this power must want to limit praise and blame to those who have it. Even supposing that they value it in others, it is unclear why that should have implications for moral responsibility.

Mele's point in distinguishing initiatory power is to identify a reason for allowing an element of indeterminism in action that can be combined with acceptance of "Frankfurt-type" cases, where a mechanism is in place that would block alternative possibilities.<sup>13</sup> Frankfurt-type cases are designed to show that an agent who acts for reasons of her own is morally responsible for what she does—and does it "freely," or "of her own free will"—whether or not she can do something else instead. But these adverbial attributions of freedom concern the origins of action or the manner in which it is done, not something forward-looking. What I want to say it also makes sense to value—*apart from* any bearing it may have on moral responsibility—is just what we would lack in a Frankfurt-type case: the freedom to do otherwise.

Fischer and Ravizza call this "regulative control," but their technical language makes it seem all too easy to forgo. We might speak instead of "determinative" power, but to avoid the overlap with "determinism," I prefer "decision-power." I take this to include, not just independence of the past, but also some real influence on the future. I want to be *in charge of* what I do—and also of what I decide to do, and even what reasons my decision should be based on. My core argument below defends that last possibility, the application of decision-power to reasons. Meanwhile, I think we can take a leaf from Mele's defense of initiatory power and be content to say that *some* of us value this fuller kind of power, and can do so without irrationality, even if it goes beyond what we need for moral responsibility.<sup>14</sup> Indeed, it will be enough to say that we value it in some endeavors or areas of life, whether or not we could dispense with it in others. In negative terms: surely some of us would be put off to find that another agent was set to interfere with our decision-making on significant matters, even if we were about to make the decision he happened to favor on our own.

I also want to add that, besides simply being valued, belief in our own decision-power may be important to our motivation in certain areas—though, again, not everyone need feel the same way. Randolph Clarke reminds us that part of what we want from free will is to "make a difference."<sup>15</sup> My interest in pursuing certain of my goals depends on thinking that the pursuit of them depends crucially on me. Fischer uses an example of artistic creation to debunk the relevance of this notion to free will in favor of self-expression: a statue would still have value for the sculptor as a statement of his creativity, even if he should find out that someone else would have sculpted it if he had not.<sup>16</sup> But we can grant that the statue would still have *some* value for a sculptor in this situation and still think its value

for him would be lessened, along with his motivation to create it, if he knew his creative efforts would make no real difference.

Of course, the degree of effort we exert does make a difference to what we manage to achieve. But if we saw our efforts as mere links in a deterministic causal chain extending back beyond us—or thought that a Frankfurt-type mechanism was set up to ensure that we chose to exert them—a somewhat disengaged attitude toward the need to exert them would seem to be appropriate. Instrumentally, our efforts would have the same point: in a case where a certain degree of effort was necessary to achieve some end we have, it would not follow (per the “lazy argument”) that we no longer would need to try so hard. But non-instrumentally, if we knew our efforts were determined in advance of our decision, we would seem to be justified in substituting detached curiosity, about just how much effort we would muster, for practical concern *to* muster the requisite effort. We would see ourselves as essentially just conducting motion from another source, driven to make whatever effort we do, rather than the drivers; or in Frankfurt-type cases, we would see ourselves as driving along a fixed track. We might need to suppress this view for instrumental purposes at the time of action and work up the illusion of setting our own course, but for those of us who could see through the illusion in more removed contexts, the result would surely affect our long-range plans, along with our sense of individual importance.

Fischer’s artistic analogy may not be quite apt here, since on some views artistic creation is driven by nonrational motives that the artist cannot control. In that case, the value of his work would not be diminished by ultimate attribution to external sources, as long as these gave rise to a unique sensibility that shaped his creation. So instead substitute a form of work that we see as reflecting our rational faculties. In deciding what to write in a paper, for instance, it matters to me that I have options. I might conceivably agree to write something whose topic and structure were narrowly prescribed, but I would consider that a less desirable task, even granting that another sort of person might be just as happy with it, or even relieved to avoid difficult choices. Similarly for constructing the “narrative” of one’s life. I want to be the one who decides what particular projects I pursue, rather than simply expressing a nature I was given.

My discussion of “making a difference” and related suggestions for explaining the value of free will has mixed together forward- and backward-looking considerations. Just as blame on the common view should be limited to acts whose agents are free, at some time before they act, to do otherwise, part of my reason for wanting to be free myself may indeed be the credit I get for having chosen well. My essential claim here is just that it makes sense for me to want a fuller kind of freedom in prospect, as I consider what to do, than I feel any need to attribute either to myself or to others in retrospect, when assessing moral responsibility for past choices.

Indeed, even in assessing the value of freedom, it seems to make a difference whether the situation we are evaluating is viewed in prospect or in retrospect. Consider a Frankfurt-type case in which someone's parents have made arrangements to have him married off to the very spouse he is about to choose on his own.<sup>17</sup> Perhaps he would not mind finding out about the arrangements sometime later. The choice was still *his* choice, in the sense that he actually made it, rendering their arrangements unnecessary. But if he found out about their arrangements before he made the choice, he might very reasonably resent their interference, even assuming an otherwise untroubled relationship with his parents.<sup>18</sup> They have deprived him of something many of us want, something whose loss might even affect his attitude toward the prospective union: namely, a choice.

Similarly for career choice and any number of everyday choices. Though it is a commonplace of popular psychology that *too much* choice is bad for us, that applies to choices among options hard to rank, whereas here "choice" simply means having *any* options, even if the relevant reasons strongly favor one of them. I go on in my next section to show how rationality can allow for choice under those circumstances. Some may have no use for it—much as some think political freedom has no value apart from its tendency to promote the value of outcomes—but it amounts to the authority of an individual agent to shape her deliberation.

In the recent free will literature Fischer asks us to consider how we would react if determinism were someday established.<sup>19</sup> Let me sum up this section by answering on my own behalf. I would not expect our institutions or our relationships founded on notions of responsibility to fall apart. However, I *would* feel a sense of personal loss, of a sort that seems likely to affect my motivation to pursue certain life-projects. Lacking knowledge, I of course would be curious about which projects I would pursue and how they would turn out. Presumably I would work as hard at making them turn out well as past causes allow, but past causes would include my knowledge that determinism is true, which (if I know myself) would tend to engender a degree of detachment. I would be watching my life unfold, rather than trying to push it in certain directions. Or that is how I would look at things at moments when I could afford to take the implications of determinism to heart. Some might find this a healthier outlook; in popular talk it is the *opposite* of being "driven" (meaning self-driven). But it would come at a motivational price. To the extent that I could manage to drop the illusion that I had options, I would feel less committed to the course of life I happened to pursue.

## 2. In Defense of Active Deliberation

Now let me ask how I can retain the view of myself as a basically rational agent with decision-power, including even the power to assign weight to my



reasons. What I have termed active deliberation raises a version of a common question about libertarian accounts: whether and how it can make rational sense for a single past to give rise to different futures, each containing a different alternative act. “Same past/different futures” is represented as a problem for libertarian free will in terms that include “luck,” “chance,” “randomness,” and “arbitrariness,” but in application to reasons the last term seems most apt to convey what amounts to a swerve within deliberation. How, we might ask, except by making an arbitrary choice, ungrounded in any reasons, can an agent move from recognizing a single set of reasons applicable to a given case, one of which seems to outweigh the others, to either of two different practical conclusions, at her option? This makes no sense on a common picture of reasons as yielding practical requirements when they are strong enough to outweigh competing reasons.

There is another way of understanding reasons, however, that can make sense of an option to re-weight them, by appeal to higher-order reasons that are themselves optional: reasons it would not be irrational to ignore. The choice as to whether to act on such higher-order reasons would indeed involve an element of arbitrariness, but no departure from rationality, if I am right. The details of the view are somewhat complicated, but it can be seen as spelling out an idea that is familiar enough from everyday practical reasoning: setting new priorities.

Let me illustrate this with a decision I had to make a few summers ago about where to spend a week’s vacation after a conference in Italy. I had reasons both for traveling on to Rome, since I had never seen the coliseum and other ancient historical sites, and for visiting the Italian Riviera, since I like to relax by the sea. Trying to get to both places in the time allotted would have made the trip impossibly hectic, so I had to choose. We might suppose that initially I recognized a clear ranking of my two competing reasons, assigning more importance to visiting essential cultural sites during my short time in the area than to spending that particular week on a beach. But intuitively, it would still seem to have been open to me just to decide to set relaxation rather than tourism as my priority for that vacation and go to the Riviera. Whether or not that was the best decision I could have made, it would have made rational sense. I had the option of discounting touristic considerations as a distraction from an aim I chose to stress.

To spell out what such cases involve let me ignore free will for a while and outline what I call the critical conception of practical reasons, according to which the normative function of a reason is to offer or answer criticism. This is a version of a general approach to reasons that several recent authors have defended, according to which some reasons do not yield requirements, so that action on them is optional.<sup>20</sup> However, on my account even a reason of the right sort to yield a requirement—a reason that offers a criticism of competing options—may be discounted by appeal to an optional higher-order reason. In my Rome/Riviera case, for instance, the fact that the

Coliseum is in Rome was a reason I recognized in favor of going there, though according to the critical conception I would be rationally required to go there only in light of unanswered criticism for failing to go there.<sup>21</sup> But I would have had the option of re-setting my priorities by appeal to the value of relaxation as a reason for discounting that criticism.

We can think of reasons of the sort that offer criticism as *negative* or *critical* reasons—entries, in effect, on the “con” side of the ledger, counting *against* the option in question. Entries on the “pro” side, counting in favor of some option, amount to *positive* or *favoring* reasons. In contrast to standard approaches to reasons, the critical conception makes out the negative case as primary. Framing the distinction in positive/negative terms can be misleading, though. For one thing, we need to preserve the logical point that a reason in favor of doing something is a reason against not doing it—and hence against doing anything else that would interfere with doing it. However, that can be accomplished just by counting such negative reasons as trivial consequences of positive reasons, lodging only insignificant criticism, since they merely sum up in negative terms information from the “pro” side of the ledger, about the positive features of an alternative option. So my discussion will be limited to nontrivial reasons, whose underlying criticism is seen as significant insofar as it makes out some option as objectionable or defective.

Many nontrivial reasons, however, are naturally stated in positive form but should be understood as fundamentally negative. The reasons that give rise to practical requirements are often represented as favoring certain options, the acts we ought to do, but in the first instance they amount to critical reasons, ruling out alternatives. They may be more likely to motivate us when expressed positively, but their normative force really depends on offering criticism. The only favoring reason they need to add to this is one that supplies permission to take the remaining option, answering any reasons against it. Thus, my guidebook’s inclusion of the Coliseum on its “must-see” list for tourists in Italy means, if taken literally, not just that the Coliseum is a major attraction of the area, but also that there would be something defective about a trip there that failed to include it.

On the critical conception a reason counts as a favoring reason, which ascribes some valuable feature to an option, just insofar as it yields an answer to criticism—not necessarily criticism anyone has lodged or is expected to lodge, but rather, potential criticism. So the function of a favoring reason on this account is in the first instance defensive or justificatory: it serves to render an option eligible for choice, to borrow from Joseph Raz’s general gloss on reasons.<sup>22</sup> Where it exceeds what is needed to answer applicable criticism, it also has a commendatory function. By contrast, a negative or critical reason *disqualifies* an option, rules it out as ineligible for choice. Or more precisely, it tends to do so, since its underlying criticism might be adequately answered by favoring reasons.

Can reasons ever be *purely* positive, in the sense of implying no significant criticism of alternatives? I say yes and will henceforth take this for granted in my references to positive or favoring reasons. The point is important to my argument here, since it allows for reasons it can be rational to discount without appeal to further reasons. The fact that sea air is healthful, say, offered me a reason to go to the Riviera during my vacation, in the sense of citing a benefit of going there. But assuming that my health would be fine without sea air, this reason did not count against going somewhere else instead. Even if I had no serious reason to go elsewhere, I could simply have turned down the benefits in question, without any need to defend that choice.

Sometimes, too, one can rationally turn down a reason of the sort that otherwise would yield a requirement: a reason offering a criticism that one cannot answer adequately by citing competing reasons. For there are other ways of answering criticism besides countering it with reasons at the same level. In Rome/Riviera we might suppose that the benefits of relaxing on the Riviera were not important enough to answer criticism of missing the Coliseum. Instead, though, if I made a decision to prioritize relaxation on the trip, that would give rise to a higher-order critical reason, against failing to follow through on my decision, for discounting my reason against skipping Rome.

In effect, then, my account allows for two different forms of rational discounting, depending on what sort of reason is in question. For reasons of the sort that imply significant criticism of some act or other practical option—negative or critical reasons—discounting has to be justified by higher-order reasons.<sup>23</sup> By contrast, we can discount a favoring reason, which simply cites the benefits of some option, without appeal to higher-order reasons, and in that sense arbitrarily. In Rome/Riviera I could discount the health benefits of sea air, say, because I had no particular need for them that summer. But by the same token, if my higher-order reason for *discounting* my reason against skipping Rome was itself just a positive or favoring reason, I might have discounted it too.

Let me now apply this to free will concerns. One of several cases of choice that Kane discusses as raising the problem of “same past/different futures” involves Molly, a law school graduate who has to decide whether to join a larger law firm in Dallas or a smaller firm in Austin.<sup>24</sup> Molly has deliberated at length on the merits of both firms and has concluded that joining the Dallas firm would be better—better for her career is what Kane says, but presumably he means to imply that it would be better for her overall, since he uses the case to raise questions about the rationality of an alternative decision. Taking the choice Molly makes to join the Dallas firm as undetermined would seem to mean that she could have opted for Austin instead after the very same process of deliberation. But as Kane puts it (speaking on behalf of a critic of libertarianism), for Molly to choose the

Austin firm, given exactly the same motives and prior process of reasoning that in fact led her to believe the Dallas firm was better, would be senseless, irrational, incoherent. In short, deliberative rationality seems to rule out “same past/different futures.”

Now, in Rome/Riviera the suggestion was that I could reverse the perceived weights of my reasons—and thus, in effect, do otherwise than prior deliberation on the merits of my options required—on the basis of a decision to set new priorities, favoring relaxation over cultural tourism. But my priorities might have been limited to that particular vacation. They need not have established any enduring motives or a precedent for future choice. Setting them committed me to enough behavioral consistency to follow through on my choice on that occasion, in order to get the intended benefit, but it did not commit me to later choices consistent with that one, even choices at around the same time, except where a different choice would have undermined the point of that one. There would have been no rational problem with my assigning *lesser* weight to relaxation than to cultural interest in choosing my lodgings on the Riviera, say, at least assuming I did not choose lodgings likely to make my stay just as hectic as touring Rome.

To be said to count as itself an instance of “same past/different futures” my choice of the Riviera over Rome would have had to be based on a priority-setting decision right at the time of action, rather than before. But to make the case more intuitively plausible, we might just shift the problem to an earlier time, within deliberation, at which I might have decided to discount my reason for touring Rome. We would then have a case of further deliberation, modifying my prior conclusion, but the point is that it would have been neither irrational nor rationally required.

Note that, besides simply being grounded in some reasons, my decision here would not have been akratic—against my better judgment—since it was made on the basis of a *change* in my better judgment, whether simultaneously with action or within deliberation. Nor need this instance of “judgment-shift” come out as weak-willed on other criteria, unless we add the further assumption that it was driven by a tempting desire that I really should have disregarded.<sup>25</sup> There might be versions of Rome/Riviera where my priority setting would just have been a cover for laziness or the like, but let us assume otherwise. I simply have the right, rationally speaking, to set priorities for my vacation.

Understood along the lines I have suggested, Rome/Riviera would resemble a variant of Molly’s case in which she had initially favored Austin but then decided to prioritize career goals and go to Dallas instead. Suppose that Molly initially had concluded that the Austin firm, because of its more supportive social environment, would be better for her overall, but then revised her priorities and chose Dallas. (A case of re-weighting in the opposite direction would work too, but in this direction it is less likely to seem weak-willed.) What makes Rome/Riviera different from this version of Molly’s

case is that my choice of the Riviera would not have had any very far-reaching implications for what I should do in the future, as Molly's decision to prioritize career goals would. However, I think we can accommodate both kinds of case just by accepting the element of arbitrariness in higher-order choice as compatible with rationality. Unless one fiddles with the notion of overall value to make it coincide with whatever an agent decides to prioritize, it is not obviously the sole criterion of rational choice, even as applied to one's life as a whole.

For a smaller-scale decision, consider Raz's case of deciding in light of just some of one's reasons: choosing a movie on the basis of its photographic excellence, while acknowledging that there are better movies to see, since other factors such as plot and character development are more important than photographic merit to the overall worth of a movie.<sup>26</sup> Though Raz interprets the relevant reasons as incommensurable, my proposed understanding of reasons would allow for such cases even granting that the reasons as well as the values in them admit of a clear ranking in terms of strength. Nor need we say that photography must have been of greater value *to* the agent, in any sense independent of the fact that she chose to stress it. A rational agent might just decide to discount all but a particular consideration that is known to be lower in an overall ranking.

One might ask how often we appeal to a structure of reasons as elaborate as that involved in discounting critical reasons. Remember, though, that the charge of rational incoherency against libertarian free will is aimed against a counterfactual possibility—"doing otherwise"—which might rest on reasons we do not in fact have in mind. In fact, in the real-life version of Rome/Riviera, I chose to go to Rome. Doing otherwise may have been a out-of-character for me, but according to what I have argued here, it could still have made sense rationally. My defense of weight-reversal as arbitrary but rationally coherent applies even to cases where we act more or less automatically, without considering options at all, let alone prioritizing reasons in light of optional higher-order reasons. What matters is what we could have done instead and how we might have justified it—whether we could have set priorities justifying a different decision on that occasion, not whether we did. Except in rare cases—perhaps Luther's—where acting against our usual priorities would be impossible for us, we have the option of deviating "just this once."

### **3. Conclusion**

My account of reasons for doing otherwise preserves our sense of having options—what is sometimes represented as an experience of free will—in cases of everyday action. Undetermined action need not be limited to special occasions of self-creation. Posting freedom farther back in the causal chain leading to ordinary in-character action seems to me to be one

instance of a current tendency of event-causal libertarians to concede too much to compatibilists. Another instance is the limitation of freedom worth caring about to cases where it makes a difference to moral responsibility, as questioned in my first section. Yet another I might mention briefly is a picture of indeterminism as involving some sort of interference with a deterministic mechanism applicable to action by default.<sup>27</sup> All that indeterminism entails, of course, is the absence of a fully deterministic mechanism, not the presence of some further factor messing it up.

My discussion here assumes that we can attribute acts and choices to a rational agent without appeal to a special notion of agent-causation. Space does not permit a full treatment in this paper, but I should note that there are other forms of the objection from arbitrariness, luck, and the like besides the one I have dealt with here, which turns on rational incoherency. An indeterministic event-causal account may not be able to explain why an agent decides to do one thing *rather than* another. But I think what needs explanation is really something more general: how a behaving organism gets to be a rational agent, exhibiting the kind of overall coordinative pattern of thought and action that is orderly enough to tolerate a few practical swerves. At that level our explanation might be deterministic. All we would need in order to attribute undetermined acts to such an agent—and to make it odd to attribute them to luck—is an appropriate connection to other elements of a coherent self.

## Notes

1. E.g., Robert Kane, *A Contemporary Introduction to Free Will* (New York: Oxford University Press, 2005), p. 130; cf. also *The Significance of Free Will* (New York: Oxford University Press, 1998). Shaping our motives might be interpreted to accommodate a single instance of choice, though typically Kane represents “self-forming actions” as choices by which we “make ourselves into the kinds of persons we are” (p. 135; cf. p. 145).
2. Robert Nozick, *Philosophical Explanations* (Cambridge, Mass.: Harvard University Press, 1981), p. 294ff.
3. Alfred R. Mele, *Free Will and Luck* (Oxford: Oxford University Press, 2006), esp. ch. 5.
4. Cf. esp Bernard Williams, ‘Internal and External Reasons,’ in *Moral Luck: Philosophical Papers 1973–1980* (Cambridge: Cambridge University Press, 1981).
5. E.g., the work collected in Gerd Gigerenzer and Reinhard Selten (eds.), *Bounded Rationality: The Adaptive Toolbox* (Cambridge, Mass.: MIT Press, 2002); for a general summary see Gigerenzer and H. Brighton, “Homo Heuristicus: Why Biased Minds Make Better Inferences,” *Topics in Cognitive Science* (2009), 107–43.
6. Cf. George Ainslie, *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person* (Cambridge: Cambridge University Press, 1992).

7. John Martin Fischer and Mark Ravizza; *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge University Press, 1998), pp. 51–54.
8. I also acknowledge a forward-looking notion of responsibility, but assuming that it is related to the notion of liability to praise or blame, it depends on resolving to view someone later as blameworthy for failure to live up to his responsibility and thus refers to the backward-looking notion. Similarly, we can ask whether an agent was free to do otherwise than he did at some point in the past, but answering the question depends on looking forward from the standpoint of an earlier time.
9. See “Responsible Psychopaths,” *Philosophical Psychology* (2003), 417–29.
10. P. F. Strawson, “Freedom and Resentment,” *Proceedings of the British Academy*, 48 (1962), 1–25.
11. Mele, *Free Will and Luck*, esp. pp. 93–100; cf. Mele, “Soft Libertarianism and Frankfurt-Style Scenarios,” *Philosophical Topics* (1996), 123–41.
12. The result is said to be a “soft” version of libertarianism, meaning one that does not require indeterminism for freedom and moral responsibility, but instead is content with the claim that indeterminism is essential to a more desirable version of freedom and responsibility. However, if we distinguish a sense of freedom that goes beyond what is needed for responsibility, the resulting view would seem to count as “hard”—hard libertarianism proper, in my terms above.
13. Cf. Harry G. Frankfurt, “Alternate Possibilities and Moral Responsibility,” *Journal of Philosophy*, (1969), 829–839.
14. Mele acknowledges this possible application of his argument; see *Free Will and Luck*, p. 96.
15. Randolph Clarke, *Libertarian Accounts of Free Will* (Oxford: Oxford University Press, 2003), pp. 5–6.
16. Fischer, “Responsibility and Self-Expression,” *The Journal of Ethics*, 3 (1999), 277–297, esp. 287.
17. I owe this example to Stephen Frank-Emet.
18. If an agent in this sort of case might welcome God or Fate as the controller, must parents be the problem, after all? As with creativity, we can grant that romantic conceptions might ascribe value to being moved by forces that bypass the will, also including the Unconscious. But let us assume that the agent in this case means to be deciding on the basis of rational deliberation.
19. Fischer, *The Metaphysics of Free Will* (Oxford: Blackwell, 1994), pp. 6–7.
20. See esp. Jonathan Dancy, “Enticing Reasons,” in R. Jay Wallace, Phillip Pettit, Samuel Scheffler, and Michael Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Oxford: Oxford University Press, 2004), pp. 91–118, and Joshua Gert, *Brute Rationality: Normativity and Human Action* (Cambridge: Cambridge University Press, 2006). For my own account, at various stages of development, see “Asymmetrical Reasons,” in M. E. Reicher and J. C. Marek (eds.), *Experience and Analysis: Proceedings of the 27th International Wittgenstein Symposium* (Vienna: oebv&hpt, 2005), pp. 387–94; and “Practical Reasons and Moral ‘Ought’,” in Russ Schafer-Landau (ed.), *Oxford Studies in Metaethics*, Vol. II (Oxford: Oxford University Press, 2007), pp. 172–94, and the first section of “Making Room for Options: Moral Reasons, Imperfect Duties, and Choice,” *Social Philosophy and Policy*, 27 (2010), 181–205; for more on the connection to moral reasons, cf. “Craving the Right: Emotions and Moral

Reasons,” in C. Bagnoli (ed.), *Morality and the Emotions* (Oxford: Oxford University Press, forthcoming).

21. Such claims about rationality depend on the assumption, outlined earlier, that the agents in my cases accept the reasons attributed to them. Here a rational requirement that I go to Rome would also assume that I was aware that the reasons against failing to go there were not adequately answered. Let me remind the reader, too, of my assumption that discounting a reason on inadequate grounds may make an action wrong-headed without making it irrational—in the sense, remember, of “rationally incoherent”. Higher-order reasons in favor of discounting may themselves be answered by critical reasons, so that my view would not justify, say, prioritizing relaxation above attention to serious health concerns. It would simply allow that practical reasoning with that conclusion is coherent.
22. Joseph Raz, *Engaging Reason: On the Theory of Value and Action* (Oxford: Oxford University Press, 1999), pp. 96–105ff.
23. Cf. the treatment of “exclusionary” reasons in Raz, *Practical Reason and Norms* (Princeton: Princeton University Press, 1990).
24. Kane, *Contemporary Introduction*, p. 16; cf. pp. 36f. for discussion of a case similar to Rome/Riviera.
25. Cf. Richard Holton, *Willing, Wanting, Waiting* (Oxford: Clarendon Press, 2009), esp. chs. 4–5.
26. Raz, *Engaging Reason*, pp. 104ff.
27. See, e.g., the analogy to background noise in Kane, e.g. *Contemporary Introduction*, p. 136, and to a roulette wheel as a randomizing mechanism in Mele, *Free Will and Luck*, p. 114.