

PHILOSOPHY OF ACTION: 5 QUESTIONS

Patricia Greenspan

1. Why were you initially drawn to theorizing about action and agency?

Like many people, I was initially attracted to free will issues – at first embracing hard determinism, as part of a general rejection of doctrines associated with religion, though exposure to Kant’s views in my first philosophy course made me begin to consider nonreligious grounds for an indeterminist conception of free action. Of course, Kant also takes belief in God and immortality as presupposed by moral agency, but I was never much moved by those arguments. On free will, though, I thought seeing my acts as determined would give me a reason to expend less effort on them.

It’s hard to articulate this without committing what Mill called the “lazy sophism.” I don’t mean that I’d necessarily work less hard or the like. If determinism is true, and the causes of my accepting the doctrine and still working hard to accomplish something were in place, then presumably I *would* work hard. Nor will hard work be any less necessary to accomplish my aims. But why should I exert the effort it would take to *get* myself to work hard? This, too, may be needed in order to accomplish my aim, so the argument can keep going back. But on some level, an attitude of “let it happen” would also seem to be justified (at any rate, in non-instrumental terms), if the causes of my expending effort (or not) are already in place. Something at least seems funny here, though it takes some work to spell out coherently: I shouldn’t actually do things any differently, if I want to accomplish my aims, but I’d be warranted in adopting a different practical attitude toward what I do (roughly: “let it happen”), if

I genuinely took determinism to heart. Or perhaps the point is best put negatively: I'd lose the sense that I need to add something to the pre-existing situation – some *new* element of effort or the like, not already supplied by it. If this is possible, I think it might very well make a difference to what I'm likely to do, even if it wouldn't affect my means-ends reasoning about what I should do, or instrumental practical reasoning – even as applied to what attitude to take toward what I do. If it's *not* possible (as determinism would seem to imply), that means there's a sense in which I can't take determinism to heart. The implications of a belief in determinism would have to be kept compartmentalized from reasoning about action – essentially, Kant's point.

Of course, this is an argument at most only for *belief* in indeterminist free will, not for its truth. Later, though, in graduate school, what really got me hooked on free will issues was hearing a version of the Humean argument that indeterminism would undermine free will by making action random, arbitrary, merely a matter of chance, or the like. I had many other interests, so free will stayed on the back burner for quite a while, but the challenge of responding to various forms of this argument in the contemporary literature led me to general readings in philosophy of action. I particularly appreciated Davidson's work, on weakness of will and other issues as well as on reasons as causes, both for its clarity and for drawing mainstream attention back to these areas, though I tended to disagree with many of Davidson's views.

For me the basic challenge in this area and others has been finding ways to combine naturalistic inclinations – the sort of outlook that Davidson and others identify with Hume – with views on foundational issues in ethics that are usually associated with the opposite, roughly Kantian, approach. My interests in moral philosophy at around the same time centered on how

and to what extent one could combine the understanding of ethics as a human invention with the insistence on a hard or inescapable moral “must.” In fact, it’s occurred to me that the various subjects I’ve worked on over the years involve different uses of “must” in application to action – senses in which we might be said to *have to* do something, where what’s at issue may or may not be causation. My published work for the most part hasn’t been directly concerned with defending a position on free will versus determinism, but rather with exploring topics about motivation and normativity that sometimes are wrongly entangled with the standard free will issue. These include the explanation of cases of psychological compulsion, the rational and moral role of emotions, and in my current work, the nature and normative basis of practical reasons.

In nonmoral cases I’ve been inclined to debunk any straight claim that one “must” act, in a sense that might be interpreted in terms of “blind” causation by nonrational factors. But I don’t take the Humean line that the element of necessity is just in the mind of the observer. Instead, I essentially refer it to something about the agent’s view of her practical options. In cases of psychological compulsion, for instance, I take the claim that the agent “has to” act as elliptical for a claim that the “compelled” act is her only tolerable option. She has to act *or* suffer the consequences – some sort of psychological distress that she experiences as overwhelming. But her perception of the threat may well be idiosyncratic and distorted. So the point is not to defend such cases as fully rational, but just to dispel a common picture of them as explained simply by causal necessitation.

I also apply the general strategy of interpreting “must” in terms of intolerable options to normal cases, especially of “pressured” action from emotion. But in relation to Humean

arguments on free will issues, it provides another way of looking at the relation between character and action besides the all-too-easy resort to deterministic causation. In terms of the recent literature, I have in mind, say, the case of Luther's famous assertion that he could "do no other" than refuse to recant the 95 theses. What I think is really meant by such claims is not that alternatives are strictly impossible for the agent, but rather that he thinks he'd lose any coherent or tolerable view of himself (or "self-narrative") if he did otherwise.

This way of looking at things is compatible with a causal view, but it doesn't require one. In short, character isn't (or needn't be) a deterministic cause of action, or part of one – nor, if we deny it that role, do we have to take it instead just as summing up the agent's actual pattern of action. It makes certain actions more likely, of course, but rather than just substituting a weaker notion of causation, I'd be inclined to see it as *constraining action rationally*, in part by imposing difficulties (upfront "costs") on some of our practical options, apparently making them poor choices. I'm sure there are further ways of working it out, but this general idea, of what might be called "constraining rationality" as an alternative to causation, is what initially got me going in philosophy of action.

2. What do you consider to be your own most important contribution(s) to theorizing about action and agency, and why?

I've applied the idea of constraining rationality to acts of a sort that are usually quickly dismissed as results of nonrational factors – unless they're distorted to fit a somewhat bloodless desire/belief model of instrumental rationality. The main example I've dealt with at length is action out of emotion. Here we also have what's now called "*constrained* rationality":

emotions are designed as “quick-and-dirty” heuristic responses, suited to cognitively limited agents like us, who often have to act without taking time to access and process all relevant information. We need emotion as a mechanism of “snap” response to partial evidence, and the rationality of an emotional response ought to be assessed in that light. So I argue that emotions are sometimes appropriate, and supply a reason for action, in cases where we wouldn’t have adequate evidence for a corresponding evaluative belief.

I have in mind everyday cases of emotion, not “blinding” rage or the like. You rebuke someone, avoid him, deny him a favor, or whatever, because you’re angry about something he did. In part this just means that you’re under the impression that he deserves it; some sort of retaliation is called for. I treat that evaluation as spelling out the content of your emotion: what your anger “says,” in effect, about its object, the person you’re angry with. Where anger is appropriate, it therefore seems to give a reason for action – in the sense of a rational ground, or what recent authors refer to as a “normative” reason: not just a motive explaining what you do, but also a consideration you could cite in favor of it.

Now, some might want to say that the emotion just refers to an evaluative judgment or corresponding desire that states your real reason for action. Its rational role is just to make salient some independent consideration, bringing it to consciousness and holding it there by loading it with feeling, or “affect.” But it’s important to my account that the affective aspect of emotion *reinforces* whatever reason is given by its evaluative content. The fact that anger is an irritating feeling – in itself, but also in persistently reminding you of something you think you ought to do but haven’t done yet – adds a further, self-regarding reason for doing something to alleviate or appease the feeling. I don’t mean that you’re likely to be thinking about improving your own state

of feeling when you act out of anger. But you're at least peripherally aware of how the emotion feels, and how retaliatory action might make it better. So anger gives you a further motive.

Even without supplying a further motive, though, the affective element of anger would still supply a normative reason for action, and we may need to add this to the evaluative reason to justify action in some cases, where the emotion isn't fully warranted in the terms that apply to belief, though it's still an appropriate reaction to partial information. I don't mean *morally* appropriate (as I've sometimes been misread as saying), but just rationally – given the heuristic purposes of emotional response. Anger might be appropriate, say, in reaction to what seems to be an intentionally harmful or insulting act, even though fuller information would show it to be based on some sort of misunderstanding.

So I've argued that it's some element of unpleasant emotion that's responsible for the illusion of "must" here, or emotional pressure; and that, besides constraining action, it supplies a reason of the sort that can *justify* action. I'm now working on a general conception of practical reasons that supports a deontological moral "must," but without the common notion that reasons as such require action. It will also leave an appropriate space for emotions as sources of moral motivation, without the Humean insistence that reason alone is incapable of motivating. But most important for my purposes, it will allow for an intermediate position on free will issues, by helping to answer arguments against libertarianism from chance, arbitrariness, and the like that turn on taking alternative choices as irrational.

Since I've also done recent work allowing for cases of moral responsibility without libertarian free will, the overall picture I'm filling out might be seen as a libertarian version of "semi-compatibilism": libertarian with respect to free will, but compatibilist about responsibility

– or at any rate, the core elements of responsibility, which suffice for its practical applications. I haven't yet published work that puts the picture together – it's only gradually come into focus – but perhaps it will ultimately be my main contribution to philosophy of action, since as far as I know it's a position no one else holds. All libertarians I'm aware of rest their case on taking free will as a requisite of moral responsibility. The term "semi-compatibilism" was coined by John Fischer to cover his soft determinist view, which acknowledges an incompatibilist sense of free will only to dismiss it immediately as irrelevant to practical reasoning. But this is what the argument I sketched at the outset of this interview would deny, if we focus not just on instrumental practical reasoning about overt action, but also on other sorts of reasoning used to justify practical attitudes. The attitude toward one's action that I think is tied to libertarian free will resembles what Robert Kane calls "ultimate responsibility," but it's importantly limited to *forward-looking* responsibility: the sort we attribute to ourselves when we take responsibility *for* some as-yet-undetermined state of affairs. Being the ultimate *agent* in a chain of causes seems to me to be ultimate enough for purposes of attributing credit or blame to an agent, so I think we can reconstruct backward-looking responsibility in a form that's compatible with determinism.

3. What other sub-disciplines in philosophy and non-philosophical disciplines stand to benefit the most from philosophical work on the nature of action and agency, and how might such engagement be accomplished?

There's room for a multi-disciplinary field with the study of motivation as its unifying subject – "conative science," perhaps. While I tend to favor fairly abstract issues in my own work, there are recognized points of contact with psychology, psychiatry, neuroscience, the social

sciences, and law, as well as to moral philosophy and philosophy of mind – the various sub-disciplines inside and outside philosophy that concern themselves with rational and moral motivation and the assessment of human agency. I wouldn't get much work done if I tried to canvas the literature in all these areas, though I'm open to input at conferences and the like. However, the emphases and assumptions of experimental work often seem to be at odds with those of philosophers. What I read in psychology and neuroscience lately suggests that these fields could benefit from increased awareness of the subtlety and complexity of work in philosophy of action. They need particularly to be made aware of the inaccuracy, or even absurdity, of some of the claims to philosophic implications of empirical work that I keep encountering: claims to have refuted this or that philosophical view or historical philosopher with an experiment or two.

There's also a tendency within philosophy, though, to accept as established various positions or arguments on the passions or free will or other issues in the area on the basis of an over-simple conception of the alternatives, or of what's entailed by a scientific world view. Before one considers possibilities of "engagement" between different areas, I think an essential first step toward getting out word of serious work in philosophy of action is to give the subject a recognized place in the philosophical curriculum. This isn't aided by making discussions of it too quickly dependent on literature in other areas, as useful as that may be at a more advanced stage. The area needs to be set on its own feet, preferably with dedicated course offerings – rather than always being treated as parasitic on ethics or philosophy of mind, the areas where most of the relevant historical work was carried on – if it's to be made more salient to people in other fields. At a minimum, it should be mentioned by name in connection with relevant readings in

introductory courses, so that students know it exists as a recognized field of contemporary research. To allow for dedicated courses, at a level likely to attract students headed toward other areas, we really need suitable introductory texts for the general area. Robert Kane's recent *Contemporary Introduction to Free Will* provides a taste of what's needed, and I've used it successfully in a 200-level course, though it's limited to free will. Anthologies (or those still in print) don't seem adequate; those I've sampled bring together such diverse approaches that the pieces don't connect into a coherent whole. We need something capable of serving as a solid foundation for a general course in philosophy of action that would help make the area more salient to people in relevant fields.

At a more advanced level, sponsoring more cross-disciplinary conferences and blogs, and collecting some of the better contributions to them for publication, might be helpful in encouraging serious engagement between different disciplines. I mean "cross-disciplinary" to contrast with "interdisciplinary," which often involves bringing a thousand unrelated flowers into bloom; what's wanted is work from different disciplines focused on a common topic, conceived specifically enough that diverse treatments are clearly treatments *of* the same thing. Perhaps the easiest way to initiate this is to get participants to address work from other areas on their own subjects of interest – though in immediate terms, that might have to be a bit antagonistic. Personally, I'd love to see something in print that collects philosophers' critiques of recent work outside philosophy purporting to have philosophic implications. But I'm not in favor of making philosophy of action into philosophy *of* the relevant empirical disciplines. We're not in the position of philosophy of physics, say, where the science is developed enough that it's appropriate for philosophers to study the subject mainly *by* studying scientists' work on the subject.

Engagement will also be deeper and less one-sided if people in other areas gain some serious background in philosophy of action, which is most likely if the area is brought to their attention as undergraduates.

4. What do you regard as the most neglected issues in contemporary work on action and agency that deserve more attention?

I think that the role of emotions within practical rationality is still largely neglected, even as the general topic of emotion has come into its own outside philosophy of action. The dominant question in the emotions literature has been what emotions are, with the answer shifting between feelings or physiological reactions on the one hand (the views familiar from Descartes and William James) and on the other hand evaluative judgments (the twentieth-century “cognitivist” view held by Bedford and Solomon, among others). But the important question for philosophy of action is how emotions bear on action. Taking emotions as cognitive states is of course one way of assigning them rational bearing on action, but there are others worth considering, and in any case the tendency among cognitivist authors has been simply to assimilate their role to that of judgment. I think they have a distinctive role to play that’s gone largely unexplored. Many philosophers cling without question to the old dichotomy between emotions and rationality, even if some of them find reason to favor the emotional side over the rational.

Perhaps one reason why emotions have not been linked in a serious way to philosophy of action is that serious treatment of emotions within ethics has often been tied to approaches stressing the virtues, following Hume and Aristotle – as against the image of a Kantian approach as unwilling to be tainted by contingent emotional influences, or of a utilitarian approach as

treating emotions as extrinsic motivators, morally significant only to the extent that they count among the good or bad states resulting from action. Virtue ethics takes character and motives rather than acts as the primary subject of moral evaluation. But emotions motivate acts – and if the work done by Robert Solomon, myself, and other recent authors on the degree of control we have over emotions is taken seriously, emotions can also be assessed *as* acts, or at any rate as resting on acts, insofar as we can control them at least by indirect means. They can be seen as raising free will issues, given the common tendency to take them as causes of action in a stronger sense than the one that seems to apply to unemotional judgment and desire.

In this connection, I'd add that another relatively neglected area is free will *apart from* the question of universal determinism. The intensity of the debate over general theoretical positions on free will has tended to obscure some particular topics that deserve attention in their own right. An example is the explanation of cases of psychological compulsion. These mainly come up in philosophy of action as problems for compatibilist accounts. I agree that they need to be handled in some way that's independent of determinism, so compatibilists have made an important contribution just in disentangling them from common determinist assumptions. But it's not clear that they can be handled adequately by simple application of a general compatibilist theory. Gary Watson discusses issues of compulsion in a set of essays dealing with addiction in *Agency and Accountability*, though his eventual explanation of addiction shifts it out of the category of compulsion. The issue of course has practical implications for medicine and the law, but in purely philosophical terms such cases raise questions of motivation and responsibility that deserve more focused consideration.

Speaking more generally, perhaps one should think of the neglected issue that stands

behind both of those I've mentioned as supplementing causation in the explanation of action. Davidson importantly defended treating reasons as causes, but he and others moved all too quickly to thinking that they *must* be causes, and that fitting actions under the causal rubric was enough to explain them. The only recognized alternative seems to be a simple sort of rational explanation involving reference to future aims with no motivational "push from behind." But there are degrees of motivational "push" that need to be distinguished, to handle both normal cases of action from emotion and abnormal cases like compulsion. No doubt there are further examples I haven't thought of.

5. What are the most important open problems in philosophical theorizing about action and agency, and what are the prospects for progress?

I take it that "progress" in philosophy doesn't involve moving toward consensus but rather recognizing significant distinctions and drawing out implications, as needed to decide among competing positions. In those terms enormous progress has been made on the general issue of free will versus determinism in recent years – so much that I'm not sure it can continue at quite the same rate, though that issue will always remain both important and open.

At this point I think that issues of practical rationality are coming into the spotlight. So far, much of the discussion has been focused on the relation of reasons to an agent's ends or desires, but it's broadening out a bit, and I hope it eventually includes more work on alternatives to the simple desire/belief model of rational motivation. Typically, desire/belief is interpreted as a causal model, though it can be understood without reference to causation. However, since the belief it refers to is understood as spelling out the means to some desired end, it's limited to

instrumental rationality. To date I've seen only fragmentary attempts to characterize distinct forms of practical reasoning, following either Aristotle or Kant, and the model is taken for granted as the basis for much work in cognitive science. It's appealing in its simplicity, but cognitivists in ethics question the necessity of desire, and cognitivists about emotion question the inclusion of evaluative states like emotions in the category of desire. Besides making room for different mental states, though, we also ought to look into the *social* aspect of reasons – as I attempt to do in my current work by interpreting reasons in terms of potential criticism. Justification is *to* someone, after all – if only an imagined someone, or an inner critic, in an exchange made intelligible by social interaction. Along with recent work on agency in a social context, it's worth exploring alternatives to the primary focus on the first-person standpoint in understanding practical rationality.

I'd also welcome further work on interpreting practical “ought” – and “wrong” and related moral notions – in terms of reasons. This seems to me to be the way to go in attempting to “demystify” deontological approaches to ethics. Making sense of normativity may require a step beyond naturalism, but reference to reasons is part of our ordinary nonmoral justification of action, so it might be able to yield an approach to ethics that's non-consequentialist but more down-to-earth than current alternatives. Prospects for progress would be good with enough people working on the project, which is essentially that of finding a sensible middle-ground between extremes. In recent years the emphasis has been more on defining and defending sharply opposed alternatives on general issues, but I get the impression from students coming up now that this trend may begin to reverse itself soon. So I have hope.¹

1. And gratitude: let me thank Karen Jones for very helpful comments on initial drafts.