Moral Agency *sans* Volitionism

Roman Altshuler

Abstract: Holding others and ourselves to moral norms requires that we understand agents as possessing a capacity for recognizing and acting on those norms. Taking our status as moral agents seriously therefore requires us to make sense of such a capacity. Wallace argues that a genuine ability to act in accordance with norms requires that at least some of our motives be up to us, meaning that as moral agents we must be capable of creating new items in the motivational sequences leading up to action. These active states, volitions, are sharply distinguished from the passive desires against which we exercise our agency. I argue that Wallace’s account is too extreme: the sharp dividing line between passivity and activity in our mental life should be softened. A dispositional view fares better at making sense of our activity as agents, while nevertheless satisfying the normative concerns motivating Wallace’s account.

Holding others and ourselves to moral norms seems to require that we understand agents as possessing a capacity for recognizing and acting on those norms. Taking normativity seriously therefore requires us to make sense of such a capacity. Responding to this challenge, R. Jay Wallace offers us a volitionist picture of freedom as an ability to intervene in the psychological sequence leading up to action.1 I argue that Wallace’s account is too extreme: the sharp dividing line between passivity and activity in our mental life should be softened. Instead, I argue that a dispositional view fares better at making sense of our activity as agents, while nevertheless satisfying the normative concerns motivating Wallace’s account.

I. Volitionism

If there are such things as norms to which we hold agents accountable, agents must have the ability to be motivated by such norms.2 Alternatively, if there are agents who can act in accordance with norms—moral agents—these agents must have a special capacity that allows them to so act. This has two

---

1 Wallace sometimes contrasts the normative capacity he has in mind with freedom of the will, which he presents as involving alternative possibilities. (Wallace 1994 1-7) Following Kant and Anselm, I will simply refer to the normative capacity Wallace has in mind as “freedom” or “free will,” leaving aside the further question of whether alternative possibilities are necessary to this capacity.

2 Wallace phrases his account in terms of reasons, but the argument retains its cogency regardless of whether we speak of the existence of reasons for action, norms that give us reasons, or norms that allows us to adjudicate between reasons. What matters is that there are some normative entities, and their existence presses on us an account of how it is that agents can act in accordance with those entities. I should add that by “entities” I do not mean to imply that there are normative objects or states of affairs as opposed to simply facts.
implications. First, moral agents must be capable of practical deliberation, understood as deliberation about what they have most reason to do.\(^3\) Second, they must be capable of being guided by the upshots of such deliberation. In saying that moral agents must be such that they can act according to norms, Wallace holds that the “can” must be taken in a non-conditional sense. A longstanding compatibilist tradition attempts to understand “can” conditionally. That is, to say that an agent A acts according to reason X is to say that A desires to act in accordance with X. Should it be the case that Y was a stronger reason than X and A recognized this fact, we can still say that—although his desire for X was stronger than his desire for Y—A could have done Y in the sense that he would have done Y had his desire for Y been stronger than his desire for X.

This conditional analysis has exercised an especially powerful influence on twentieth century compatibilism because it meshes neatly with the belief-desire theory of action, which holds that an agent A does X intentionally iff she wants D and believes that X is a means to attaining D. Such accounts, however, find it difficult to explain akasia. An akratic agent judges that, all things considered, he had better X, but he Ys instead. The conditional analysis can claim only that, had A wanted to X more, he would have. But Wallace argues that such an account is deceptive. It is not true that A genuinely could have Xed but chose instead to Y, because given A’s existing motivations he could not have done anything but Y. (223-224)\(^4\) If this is so, then it is equally false on the conditional analysis to say that A can act in accord with his judgment about what he has most reason to do. Rather, the situation seems to be this: given A’s desires and beliefs prior to his action, he necessarily does either X or Y on the basis of which of these he desires to do more. This is what Wallace refers to as the hydraulic theory of agency. If the hydraulic theory is right, however, it follows that A’s judgment about which he has more reason to do, X or Y, is relevant to what he will do only insofar as the belief inherent in the judgment interacts with the

\(^3\) Wallace vacillates between describing practical deliberation in terms of deliberation about what agents have reason to do, what they have most reason to do, and deliberation about what to do. I am not convinced that these are equivalent, and I suspect that the last is the most typical kind of practical deliberation. This distinction, however, should not make a difference to the argument developed here.

\(^4\) All references to Wallace will be by page number to essays reprinted in Wallace (2006).
agent's desires. But it is only A’s pre-existing desires to X or to Y that interact with the judgment to produce an action5; A has, in this scenario, no ability to act in accord with or counter to his best judgment. What he does is up to his desires, not to him. If the hydraulic theory is right, we lack genuine agency.

Few compatibilists still maintain the conditional account Wallace attacks, and he does not engage with the more recent, and far more nuanced versions of compatibilism put forward in the wake of Frankfurt and Strawson.6 This does not, however, lessen the appeal of Wallace’s approach. By attacking the core assumption of standard belief-desire approaches, he succeeds in bringing out what is at stake in the free will debates. Free will, on this account, is important not in the way self-expression or creativity is important. Rather, it is important because it is implicated in our practical deliberation, as well as our practices of holding others responsible. These practices presuppose the ability to guide ourselves in accordance with norms. But this ability is lacking on any account that assumes a hydraulic theory on which agents’ actions are a function of their given psychological states.

Nevertheless, Wallace maintains that the belief-desire account of action is not wrong. In explaining why A did X, we do rely on causal explanations, stating the motives that caused X. Reconciling this view with a recognition of ourselves as genuine agents, however, will require us to recognize a division in the notion of motive, or desire, that we take as the causal factor in agency. Wallace argues that we need to distinguish desires in the ordinary sense, as the passive states we simply find ourselves with—what he calls “given desires”—from states such as decisions and choices, things that we do, which he groups under the category of volitions. (e.g., pp. 6, 149-150, 173) The mistake of hydraulic theories is to think that our actions are caused by given desires; recognizing genuine agency will require

5 Technically, the point is not about whether A desires to X or Y, but whether A desires to D or E, where he believes that Xing will lead to D and Ying will lead to E. For the sake of simplicity, I will speak in terms of desires to X and Y.

6 Wallace himself, of course, puts forward a compatibilist view inspired by Strawson and, to some extent, Frankfurt. Nevertheless, his compatibilism seems to be a bit of an outlier, and his attack on the hydraulic theory seems aimed to apply to competing compatibilist approaches as well. For a discussion of these issues and whether Wallace’s account is consistent with compatibilism, see the review of Normativity and the Will by Gert and McKenna (2008).

7 The first suggestion is made by Fischer. See, e.g., (Fischer 2009). The latter is listed as a candidate for the significance of free will by Robert Kane (1996).
us to see that our actions can instead be caused by volitions, which are subject to our direct and immediate control and allow us to intervene in the causal chain between our given desires and our actions. On this view, an agent with a set of given desires who does X nevertheless has—leaving fixed the given desires to which he was subject prior to the action—the ability to choose instead to do Y. This volitional capacity accounts, for Wallace, both for our ability to choose independently of our given desires and our ability for *akratic* and other sorts of irrational action, as when we choose to do something we have reason to avoid and perhaps lack even strong given desires to undertake.8

II. Troubles for Volitionism

Wallace takes his account to be opposed to any form of psychological determinism. In order to be genuine agents, we must be able to intervene in the motivational sequence leading up to action, so that the sequence alone does not determine the action we undertake.9 However, even rejecting psychological determinism will not give us what we need to make sense of genuine agency. Wallace recognizes, for example, that desires, especially resilient ones, might focus our attention on particular considerations, making it more difficult for us to grasp correct reasons for action. (181) His strategy is to stress the “more difficult” clause: neither desires nor addictions make it literally impossible for us to grasp and act for correct reasons, since they interfere with but do not eliminate our volitional capacities. The line between interference and determination is a hard one to draw, however. If desires can focus our attention on particular considerations over others, they certainly increase the likelihood of our taking those considerations as decisive. Nor is it obvious that we have some ability to deliberate completely

---

8 On this point, Wallace differs from Wolf (1980), who argues that determinism is only a problem when it leads as away from good reasons. Being determined to good, she holds, is unproblematic. Wallace, on the other hand, wants to retain the ability to explain what seems, at least, to be a common ability on our parts: the ability to do something we neither particularly want to do nor take to be worthwhile, as evidenced for example by certain cases of procrastination.

9 Just how we can reject psychological determinism while accepting physical determinism, given that psychological events and states are presumably constituted by or dependent upon physical ones, is left underexplained. Instead, Wallace turns to the question of whether psychological determinism seems plausible (163-164), but he discusses its plausibility in terms of folk-psychology rather than in terms of its connection with physical determinism.
independently of such desires, which may in fact be necessary to make at least some considerations salient for us and worth considering in the first place. Our desires may work not against, but through our deliberation, making it impossible for us to judge and act independently of them. (Katsafanas forthcoming) Nor is it clear that our deliberations or volitions are entirely up to us. We can make ourselves think, and we can make ourselves think about a particular subject, but what actual thoughts occur to us is not in our control. The occurrence of such thoughts is, instead, the means by which we exercise control. (Strawson 2003) Thus, there is a sense in which our thoughts—including both our judgments and our decisions—are tinged by a kind of passivity; they, like our desires, are something “one merely finds oneself with.” (174)

This thought should lead us to a reconsideration of Wallace’s distinction between activity and passivity in our agency. The division between passive given desires and active volitions is presented as a strict dichotomy among our motivating states, but there are reasons to doubt that we can simply leave it at that. Wallace notes that given desires do have a conceptual structure that “makes it possible for them to respond to our deliberated judgments about what we have reason to do.” (149) If so, then even given desires seem to exhibit a kind of activity. Thus, as some anti-volitionists have argued, our desires represent our evaluative judgments in such a way that it makes sense to hold people responsible for their desires and not only for their explicit judgments.\(^{10}\) (Arpaly and Schroeder 1999 ; Smith 2008, 2005) This is an appealing view of desires, because it does not insist on fracturing our mental life into two diametrically opposed sorts of states.\(^{11}\) Moreover, Wallace’s distinction between passive and active states does not make the distinction in terms of any features of those states, whether conceptual or phenomenological. He appeals entirely to the claim that active states must exist if there are to be moral

\(^{10}\) Of course that does not imply that we should hold people equally praiseworthy or blameworthy for all their conative states; the point is only that desires are not—simply by virtue of being desires—completely set off from our “real selves” and placed into a domain of involuntary states over which we have no control and for which we are in no way responsible.

\(^{11}\) Wallace seems to recognize that at least some motivational states “straddle the distinction” between passivity and activity. (177) He does not, however, consider the possibility that the distinction itself may be one of degree rather than kind. That is, it is possible that some motivational states exhibit more passivity and others more activity, without it being the case that there are simply three sorts of motivational states: active, passive, and straddlers.
agents. This consideration, however, establishes only that there must be states that are active in the relevant sense; it does not show that passive states must exist in opposition to them.

Wallace’s way of breaking up motivational states does have an intuitive appeal, however. For while normative requirements do not require us to posit the existence of passive given desires, it may seem as if our experience does. To make this point, Wallace argues that given desires exhibit a resilience to change in accordance with judgment. They do not automatically match up with our value judgments; if they do, that is a matter of luck. (175) However, the same is true of many other mental states: we are very capable, as Wallace notes, of holding on to beliefs we recognize to be inconsistent with our better judgment. Nor is there reason to doubt that our beliefs about what we have reason to do may sometimes fall into this class of what we might call resilient beliefs. Moreover, as I argued above, our judgments themselves may be something passive, something we merely find ourselves with. So we have some reason to think that our judgments may be just as passive—in Wallace’s sense—as our given desires.

If at least some of our judgments are no more in our control than our desires, we may note that those supposedly active entities—judgments and volitions—may sometimes turn out to be resilient not simply relative to our other beliefs, but relative to our desires as well. For example, let’s say that Kant’s grocer—not the prudent grocer, but rather than one who loves his customers—judges that he could make a killing by ripping off select shoppers and that he could easily avoid getting caught. Moreover, let us say he concludes, after long reflection, that ripping off his customers in this way would be the right thing to do (perhaps some of his customers are themselves thieves; in any case, we need not suppose here that the grocer’s judgment is correct). Yet due to his love for his customers, the grocer cannot bring himself to act on his own best judgment. Here we have a case of akrasia, no doubt, but it is akrasia of the sort Arpaly has worked out in great detail: the grocer is acting against his best judgment, but the action is more rational than it would have been had he followed his judgment. (Arpaly 2003) Or consider someone who comes up against “volitional necessities”: who judges that X would be the right thing to do, and in fact wants to do X, and yet cannot bring herself to do it. In some such cases we might think that it is the
volition that is resilient, passive, and given. The desires, lined up with the judgments, would then appear to be the active states. Wallace could respond, of course, that to call desires that stand opposed to our volitions active is to misunderstand the nature of desires; but that would simply beg the question.

III. Toward a Non-Volitional Agency
So just why do we need to posit free will, or a capacity to judge and act in accordance with norms? Wallace's claim is that this capacity is presupposed in practical deliberation. Most animals are likely capable of deliberating—deciding which course of action to take when presented with an option. But human deliberation involves an ability to decide not simply what to do, but also what we have reason to do, or what norms we ought to follow, and an accompanying ability to act in accordance with those norms. Unless we held ourselves to have such an ability, the task of thinking about norms and reasons would be pointless; it is only on the condition that we hold our deliberation to be potentially effective—capable of making a difference to how we act—that we engage in it. (151) But note that the deliberation need only be potentially effective. For sometimes it does seem as if we decide to act contrary to our strong given desires, and then find a sort of motivational inertia preventing the decision from having any effect.

Deliberation does not lose its point, or its purchase on us, simply because it sometimes fails; we would lose reason to deliberate only if we knew that it was always pointless to do so. This requirement of practical deliberation has to be distinguished from another reason sometimes raised in defense of volitionism and occasionally alluded to by Wallace: that when we act, we often feel that we could have acted differently. Thus, when I procrastinate in an akratic way, I may feel that I can or could have

---

12 Wallace stresses that the capacity is presupposed both in our deliberation, and in our practices of holding others responsible. I focus on the first of these, in part because I think holding others responsible has no justification other than that provided from our first-person deliberation, and is thus not an essentially third-personal stance, as Wallace insists (124). Leaving out the point about responsibility, however, should make no difference to the argument here.

13 My point is not that coming to have certain beliefs would cause us to simply stop deliberating altogether. Here I am only stressing that if deliberation commits us to postulating a special practical capacity, that capacity need not be one that guarantees that we can always act on the outcome of the deliberation.
avoided the procrastination. But this feeling does not require us to postulate any ontological capacity to support it: it is perfectly consistent to suggest that at the moment I act, I do not in fact have the ability to act otherwise, and yet I feel as if I do. It is not, however, consistent to claim that I can guide myself in accordance with norms, and yet that I lack any ability to do so.

The possible range of ontological commitments implied by our normative commitments, then, is not exhausted by the sort of capacity Wallace posits. On the volitionist account, our capacity gives us the ability to intervene in the sequence of psychological states leading up to action. It is, we might say, a synchronic capacity: a standing capacity to create a new motivational state, a new event, at any moment in time. But the capacity could serve the same normative purpose if we conceived of it in diachronic terms. Imagine instead that our actions—as well as our judgments about what we have reason to do—proceed from dispositions. Here I do not want to commit myself to any particular way of sketching out what such dispositions might be like. We may take them to be dispositions of character, along Aristotelian lines, or we might take them to be reasons-responsive mechanisms, following Fischer and Ravizza (1998); we might even, for the sake of simplicity, think of them as standing plans or principles to which we are committed (Bratman 1999; Korsgaard 1996). On something like this account, we would

---

14 It is interesting to note in this regard that, while Wallace rejects agent-causal theorists’ metaphysical views—namely, that agents are substances capable of intervening in the causal order at the physical level—he retains a view of agents as defined by the ability to undertake radically undetermined actions.

15 I am arguing here that the volitionist account faces serious problems, ones that a dispositional account allows us to avoid. Ultimately, however, the viability of a dispositional account will rest on our ability to determine just what dispositions of this sort come to, and clearly not all proposals will be equally suited for the task of offering an alternative to volitionism. For example, Fischer and Ravizza’s reasons-responsive mechanisms are, presumably, acquired during some period of time (usually during the socialization stage of childhood and adolescence) and remain stable after that point; a similar view about dispositions of character can be found in Jacobs (2001), and Kane (1996) sees us as shaping our character through occasional, free “self forming willings,” while the majority of our actions are simply functions of character so formed. But as I note here, views of this sort will be insufficient to offer an alternative to volitionism, since they do not allow that our dispositions be direct expressions—as well as sources—of our ongoing agency over time. That view requires an ongoing feedback loop between our dispositions and the acts and deliberations undertaken on their basis. Similarly, Korsgaard’s view, as I note below, seems to fall to the other extreme: it allows us to alter or acquire dispositions at any point through an act of identification with a principle. Other views of dispositions, perhaps more promising with regard to the aim of offering an alternative to both volitionism and the hydraulic theory, can be found in recent Sartrean attempts to explain dispositions in terms of our projects; on this, see especially Webber (2006). Ultimately, however, I suspect such an account will have to be reworked in terms of projects understood in a Heideggerian sense, which I tentatively work out in Ch. 5 and 6 of Altshuler (2010).
not have the ability to create new volitions at any moment in time. Instead, our volitions would stem from some interaction between our perception of our surroundings, or our grasp of the situation in which we find ourselves, and our dispositions to respond in certain ways to situations and surroundings of this kind.

To see how this might work, we can consider one way of hashing out the Aristotelian distinction between continence and virtue. The continent person might, for example, consider picking up the wallet she finds on the ground and pocketing the money, but she will not act on this consideration. The virtuous person, on the other hand, will not even have such a consideration occur to her. This sort of thought is behind, for example, Smith’s suggestion that a thought to pocket the money might be blameworthy, even though it is not very blameworthy provided the agent does not seriously entertain it or act on it. (Smith 2005) The point of such an account is that it does not give us immediate and direct control over our motives as the volitionist would have it. This does not, however, mean that our motives are completely outside our control. Instead, it places all our motives—including our choices and decisions as well as our given desires—under the control of our character, or the mechanisms or dispositions that constitute it. Such an account avoids the problems raised by the stark division of motives into ones that are given and ones that are immediately up to us.

But does a dispositional view of this sort help to account for our moral agency, as Wallace conceives of it? To do so, it would have to fulfill two requirements. First, part of the point of positing a capacity of acting according to our conception of norms was to stress that we, as agents, can respond to norms; that is, that we and not desires operating in us are in control of what we do. On the dispositional view I am proposing, this means we must have some control over our dispositions. It is not enough to say simply that our dispositions, because they coordinate our activity over time, serve to constitute our practical identity and thus “stand in” for agential activity in producing actions.\footnote{This is a paraphrase—with regard to dispositions—of the view Bratman takes with regard to plans. (Bratman 2001)} That sort of view is fully compatible with the sort of hydraulic theory Wallace correctly dismisses. But it is a mistake to think that the only alternative to the hydraulic theory is an ability to create psychologically undetermined motives.
Such an ability is obscure enough; surely it is not more obscure to suggest that our character—or the dispositions and mechanisms that constitute it—may be the vehicle of our agency.

Second, the other part of the point of postulating a special capacity was to account for our practical deliberation. Our deliberation must be capable of having some effect on our action; otherwise it would be pointless. This means that whatever capacity we posit to explain the role of normativity in our lives, that capacity must involve an ability to act in accordance with norms or reasons. As already noted, the capacity need not guarantee that we are at every instant capable of responding correctly to right reason: a thoroughly vicious agent, for example, may be genuinely unable to see the force of such reasons, or—seeing it—may nevertheless lack any ability to motivate himself to act in accord with it. But this does not mean that the vicious agent lacks the ability to act in accordance with practical deliberation; it means only that he lacks this ability at the moment he acts. But we can reasonably say that he can act in accordance with the dictates of correct deliberation, provided that the disposition that at the moment prevents him from doing so is one that was up to him, and one that is still up to him to change. Even the vicious agent’s practical deliberation may then be effective, not because he has the ability to decide and act correctly at any moment, but because the deliberation itself is partly constitutive of the disposition that now prevents him from acting correctly. Thus, correct deliberation may not have a direct and immediate influence on his conduct at the moment; it may, however exert its influence on his future conduct by reconstituting his character.

Wallace repeatedly rejects Korsgaard’s constructivist alternative to moral realism. (Wallace 2006, 2004) While I will not go into the full details of that rejection, I want to suggest that a version of the constitutivism Korsgaard couples with her constructivist approach can provide an alternative to volitionism. The hallmark of constitutivism is the view that rational norms have the source of their normativity in their embeddedness within the structure of volition. In willing, Korsgaard argues, agents are necessarily committed to certain norms, particularly norms involving the effectiveness of their willing and the unification of themselves as willing agents. Other norms, on her view, are derivable from these
two and, since these two norms are necessarily implicated in all volitional activity, they and the norms derived from them are inescapable from the perspective of the agent. (Korsgaard 2008, 1996; Ferrero 2009) Korsgaard’s own brand of constitutivism adds an important feature: in acting, agents necessarily constitute their selves, or their practical identities, in the process of willing.\(^{17}\) Wallace objects that Korsgaard’s account fails to account for akrasia: if agents must always act on norms to which they are committed, they then lack the ability to violate their norms. And indeed Korsgaard does launch a puzzling attack on the very possibility of particularistic willing, i.e., willing an action without committing oneself to a principle.\(^{18}\) (Korsgaard 2009)

But the culprit in Korsgaard’s view here is, in fact, her volitionism. She holds that agents necessarily, in each action, constitute themselves according to norms they accept as universally binding; and this means, also, that agents are always, each time they deliberate on an action, capable of taking up and identifying themselves with new norms. If we avoid the volitionism, however, we can still retain the core of constitutivism. Through their volitional agency—including their deliberation and action—agents constitute themselves as having certain dispositions, which in turn guide their actions. A single decision by an agent is unlikely to completely overturn the agent’s character and radically alter her pattern of behavior; this is exactly what makes volitionism problematic.\(^{19}\) But the idea that agents, in deliberation and action, aim at a unified character is important: through deliberating and acting in certain ways, agents constitute themselves as entities who in the future will be drawn to deliberating and acting in similar

\(^{17}\) The most notable other constitutivist, David Velleman, has repeatedly rejected the claim that agents constitute themselves. (Velleman 2002, 2006) I think, however, that for the present purposes the disagreement is merely apparent. Velleman is committed to the view that self-understanding is a constitutive aim of action. But surely if our self-understanding guides our rational deliberation, then the reverse applies as well. Insofar as our practical deliberation (and its outcomes) in turn shapes our self-understanding, which then shapes further practical deliberation, it is clear that in action agents do constitute their own practical identities or agential selves. That another, metaphysical, level of selfhood is not thus constituted is beside the point.

\(^{18}\) See 84 ff. in Wallace (2006) for a discussion and rejection of this view. I do not mean to imply that the point about akrasia is Wallace’s only, or even most significant, argument against Korsgaard’s approach, which he has addressed repeatedly in a number of publications and talks.

\(^{19}\) This is not to deny that agents sometimes do act completely out of character. But there are familiar arguments, originating in Hume’s *Treatise*, to the effect that such actions must be random in a way that seems to rule out control, freedom, and responsibility.
ways. Thus, they constitute themselves over time, thereby constituting both their deliberative activity and their desires. When agents act in *akratic* ways, this view of constitutivism can easily account for that fact: the agents have not yet achieved a unified character, so that some of their volitional elements are still not aligned with their overall judgment. But insofar as the agents’ deliberation and action constitute their dispositions, they have partial—if not direct and immediate—control over their volitional states and the outcome of deliberation.

Such an account allows us, contra volitionism, to avoid the thought that the will is a purely active capacity lined up against the passive elements of our mental life and capable of creating new elements at any moment. Instead, we can imagine our volitional lives as both passive and active, with agency itself being a temporally extended process that involves slowly—through deliberation and action—transforming the resilient machinery of one’s mental life into a unified volitional capacity. Despite lacking the ability to create new motivations on the spot, however, this capacity can serve the explanatory purposes required by our normative commitments. Insofar as agents deliberate about norms and act on them, they constitute themselves as the sorts of agents whose desires line up with their judgments. This process may be difficult, relying on habituation and the gradual bringing to bear of one’s mental framework on the most resilient items within it. But the fact that our desires do not always respond to our judgments should not mislead us into believing that any match between them is nothing but a matter of luck.
References


