

## Receptivity and the Will

Edward S. Hinchman

University of Wisconsin-Milwaukee

[The long version of this paper is forthcoming in *Noûs*]

The task of a broadly internalist approach to agency is to explain how an agent's all-things-considered practical judgment has necessary implications for action. The approach faces its chief obstacle in the possibility of either of two species of akratic break: between judgment and intention or choice, and between intention or choice and action. The standard internalist strategy for overcoming the obstacle is to argue that an agent's all-things-considered judgment needn't determine action in order to guide it, and that there would still be an 'internal' link between judgment and action if *global* akrasia were impossible. That would at least entail an important degree of necessitation: even if there is an akratic break in a given case, the agent couldn't *always* be akratic.<sup>1</sup>

It is easy to feel the force of the intuition that global akrasia is impossible – but why? If global akrasia is impossible, it seems that must be because of some feature of the individual case. What then is it about the relation between practical judgment and action that makes global akrasia impossible? In what follows I take on this challenge. I argue that the two species of

---

<sup>1</sup> See, for example, Sarah Stroud, "Weakness of Will and Practical Judgement," and Gary Watson, "The Work of the Will," both in Sarah Stroud and Christine Tappolet (eds), *Weakness of Will and Practical Irrationality* (Oxford: Oxford University Press, 2003). Little attention has been paid to the second akratic break, between intention and action, because until Richard Holton's recent work it had not been conceptualized as a kind of akrasia. Actually, Holton argues that it is *not* akrasia, but that's because he identifies akrasia with the other break. Holton argues that the second break is more what both philosophers and nonphilosophers have in mind when they speak of 'weakness of will,' a point I'm happy to concede. (See Holton's "Intention and Weakness of Will," *Journal of Philosophy*, 96 (1999), 241-62. As Holton notes, Amélie Rorty had earlier discussed the second break as a species of akrasia: see her "Where Does the Akratic Break Take Place?" *Australasian Journal of Philosophy* 58 (1980), 333-46.)

akratic break are not importantly different: in each case the akrasia manifests a single species of irrational self-mistrust. I aim to vindicate internalism by showing how rational agency rests on our capacity for a kind of trusting receptivity to the verdict of judgment.

To call the relation *receptivity* is to characterize it as fundamentally passive. To call it *trusting* receptivity is to ensure that the passivity is not incompatible with agency, since trust retains a crucial degree of control. I'll argue that the best way to meet the externalist argument from akrasia is to abandon the assumption that the will must be a locus of activity.<sup>2</sup>

## I

Jay Wallace presents this rationale for externalism about the relation between judgment and intention or choice:

Human agents have the capacity for a sophisticated kind of rational agency, insofar as they can reach independent normative conclusions about what they have reason to do, and then choose in accordance with such normative conclusions. This capacity presupposes that we are equipped with the power to choose independently of the desires to which we are subject. Once we have this power, however, it can be put to use in ways that are at odds with our own

---

<sup>2</sup> Throughout my discussion I'll assume that the relations between judgment and choice and between choice and action unfold sequentially – as if the relation between judgment and action had to take the form of a process. In the full version of this paper I drop that assumption and offer analyses of practical judgment, choice and intention that codify aspects of a single synchronic act. Those analyses enable me to isolate and explain the basis of judgment's rational authority. (The full version of this paper was recently accepted for publication in *Noûs*. I'm sure it won't appear till long after the present conference, however, and there should even be time for me to do some pre-publication revising in light of criticisms offered at the conference.)

practical judgments about what we have reason to do. That is, we can treat our disposition to do what we ought as a further desire from which we set ourselves apart, choosing to act in a way that is at variance with our reflective better judgment. This may be regarded as a hazardous by-product of the capacity for self-determination that makes rational agency possible in the first place.<sup>3</sup>

Let's call this species of akratic break 'incontinence.'

Richard Holton offers a similar pitch for externalism about the relation between intending, choosing, or more generally resolving and actually following through on the resolution:

Sticking by one's resolutions is hard work.... It certainly doesn't feel as though in employing will-power one is simply letting whichever is the stronger of one's desires or intentions have its way. It rather feels as though one is actively doing something, something that requires effort. My suggestion is that effort is needed because one is actively employing one's faculty of will-power.... [,] a kind of mental effort.... On this picture, then, the effort involved in employing will-power is the effort involved in refusing to reconsider one's resolutions; and the faculty of will-power is the faculty that enables one to achieve this.<sup>4</sup>

Let's join Holton in calling this second akratic break – of irrationally failing to stick by a

---

<sup>3</sup> R. Jay Wallace, "Normativity, Commitment, and Instrumental Reason," *Philosophers' Imprint*, Vol. 1, No. 3, 9-10.

<sup>4</sup> Richard Holton, "How is Strength of Will Possible?," in Stroud and Tappolet, *op. cit.*, 49. Again, Holton does not himself regard weakness of will as a species of akrasia. This disagreement is mostly just terminological – though I'm substantive disagreement with Holton's account of weakness.

resolution – ‘weakness of will.’<sup>5</sup>

Now here is the view I wish to put in place of these externalist views. I agree with Wallace that incontinence is a hazardous by-product of the capacity that makes rational agency possible in the first place. And I agree with Holton that avoiding weakness is by no means easy. What’s not easy, on my account, is trusting reasonably in the face of the twin phenomena of unruly desires and less than perfectly trustworthy judgment or intention. The difficulty lies in reopening deliberation *when but only when* the untrustworthiness of the judging or intending self requires it. Since the question is whether to redeliberate, it must be settled through the exercise of a non-deliberative form of intelligence. I hold that this intelligence is most fundamentally a counterfactual sensitivity to evidence of untrustworthiness in the judging or intending self: if you had evidence that your judgment or intention had been untrustworthily formed you would not simply have followed through on it. I put it counterfactually, but it of course should also cover the actual case, if there is in fact evidence that you are untrustworthy. By ‘evidence’ I mean good but not necessarily conclusive evidence. Even inconclusive evidence of untrustworthiness

---

<sup>5</sup> Corresponding to the two species of akratic break we could coin distinct species of internalism. One kind of akratic break, the one I’m calling ‘incontinence,’ occurs between judgment and intention. We might call ‘judgment internalism’ the thesis that an all-things-considered judgment about what one should do bears an internal relation to intention, choice, or other ways of committing oneself to act. And we might call ‘volitional internalism’ is thesis that intending or otherwise willing to  $\phi$  bears an internal relation to such a judgment. To call these relations ‘internal’ is to say that they’re in some sense necessary or non-empirical. We can assume that the internalist theses get determinate enough content from the way they are threatened by the possibility of incontinence, in which one judges that one should  $\phi$ , all things considered, and yet without changing one’s mind proceeds to do (or intend to do) something incompatible with  $\phi$ ing. The possibility of incontinence seems to entail that there is no necessary link in either direction between judging and willing: one can will contrary to one’s best judgment, and one’s best judgment can fail to engage one’s will. The other kind of akratic break, which I’m calling ‘weakness,’ occurs between an intention and its execution: one resolves to  $\phi$  at some time  $t$  but then, without forgetting or changing one’s mind, fails to  $\phi$  at  $t$  (or even attempt to). This threatens what we might call ‘resolve internalism,’ the thesis that intending, resolving or otherwise willing to  $\phi$  bears an internal relation to actually  $\phi$ ing (or at least attempting to). The threat in each case is that we cannot understand how such counter-normative agency is possible without positing a mediating faculty of the will with two features: to accommodate incontinence, it must not be necessitated by judgment; to accommodate weakness, it must be exercised with greater or lesser degrees of ‘strength.’ The faculty of the will is incompatible with these internalist theses.

can undermine trust. Insofar as you actively weigh positive evidence of trustworthiness, you act not from trust but from your own deliberative judgment – and the question remains whether to trust that judgment.

This non-deliberative form of dispositional thinking is not at all easy to exercise. One often trusts oneself when one shouldn't and fails to trust oneself when one should. But reasonable trust along these lines is what makes rational agency possible: without it, we would be either compulsive redeliberators or incapable of changing our minds. What makes agency possible is thus what makes these two species of *akrasia* possible: the rational requirement that one stop and redeliberate when but only when one's judging or intending self is untrustworthy.

The challenge, then, lies on both sides of trust: too little mistrust, but also too much. The hazards of too little mistrust are easy to appreciate, if hard to avoid: you follow through on a foolish decision, overlook how available options have unforeseeably changed, or otherwise fail to resist flawed judgment or choice and reconsider the matter. The hazards of too much mistrust are equally hard to avoid but also difficult to conceptualize in those terms. It's natural to worry when mistrust goes fully neurotic, leaving you collapsed in a heap of self-obsession. But it can seem a puzzle to interpret either species of *akrasia* in terms of excess rather than deficiency. Isn't the failure to intend as you best judge or to act as you intend *a lapse* – perhaps, as Holton puts it, a failure of 'will-power'? Well, ask the same question of failing to trust your judgment or volitional dispositions at all: can we explain the neurotic's collapse as his merely lapsing? It seems not: *accidie* is not the same as exhaustion. The *accidic* agent feels no motivation to pursue what he values not because he's too tired but because he's in some way alienated from his evaluations. However we explain *accidie*, it seems appropriate to emphasize the active element in this alienation. And the same is true of the two species of *akrasia*: incontinence cannot be

explained as the mere failure to be continent, and weakness of will cannot be explained as the will's failure to be strong.

If it seems paradoxical that a passive follow-through should express an agent's will, the reply is that not every form of passivity relinquishes control. An agent who chooses to act by trusting his judgment that he should  $\phi$  does not relinquish control over his  $\phi$ ing but does precisely what it takes to retain it. The work of the will here is to step aside and let action be guided by trustworthy judgment. This *re*-makes one's mind insofar as (a) one is now guided by a counterfactual sensitivity to evidence of untrustworthiness in one's judgment, triggering which would reopen practical deliberation, and (b) this sensitivity may misfire, rejecting judgment or intention as untrustworthy but *without* reopening deliberation. Akrasia would thus involve mistrustful activity – either rebellious choice or rebellious action – where rationality requires this species of trusting passivity. Akrasia is counter-*normative* agency because the sensitivity has misfired and has not led to redeliberation.

## II

Before elaborating my argument for this position, let me make my view vivid with an analogy.

Imagine you're learning to dance in ballroom fashion, with your instructor as your partner. She leads, and you follow. Here are two interesting features of the case: *her* judgments and choices determine what happens as the two of you move across the dance floor, but what happens includes actions that *you* perform in just the sense in which you would perform them were you to lead. She determines the pace, the direction, even the precise shape of the physical movements you make, but this isn't like the case in which in an uncooperative moment you go

limp like a rag-doll in her arms. The imagined actions are yours.<sup>6</sup> Nor is this like the case in which you learn by leading while she barks orders, or by imitating her movements, each of you with a separate partner. You perform the imagined actions, unlike these others, by letting your conduct be directly determined by the will of another. The two observations are related. It is your receptivity to the influence of the judgments and choices determining what you do that marks these doings as your actions.

Now imagine you're an overeater learning to perform that delicate dance called *dieting*. The challenge, of course, is that you have nothing to guide you but your own judgment. It leads, and you follow – or try to. As in the ballroom, your moment-by-moment impulses are *just all wrong* and need to be guided by a force apart from them. Here too, rational agency is receptive agency. It is your receptivity to the influence of your own judgment that marks the doings as your actions.

The two cases are not perfectly analogous, of course. In the interpersonal, the influence on you is not merely your instructor's judgment but more pressingly her intentions or will. This suggests a different analogy with the intrapersonal: that your dieting cannot proceed unless you are suitably receptive to the influence not merely of your judgment *that* you should refrain from

---

<sup>6</sup> Being receptive to another's lead does not amount to relinquishing control or guidance over your actions. Your receptivity will only go so far, and *you* determine how far. If your dance lesson begins to stray toward what feels a bit too close to judo, you may withdraw your receptivity for the simple reason that that was not the sort of thing you wanted to do when you made your will receptive to your partner's lead. We might seek a theoretical rationale for this in Harry Frankfurt's conception of action as purposive behavior (see "The Problem of Action," reprinted in his *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), especially p. 49). Just as, in Frankfurt's example, you may remain in control of your automobile even as it coasts downhill in neutral gear and you remove your hands dare-devilishly from the steering wheel, so may you control the actions you perform when you give yourself over completely to your partner's lead. In each case your control consists in the fact that the course of events is *subject* to your intervention or adjustment, even if you never do in fact intervene or adjust anything. In the former case you control the motion of the automobile, although it is not you who even indirectly produces that motion. In the latter case you control the movings of your body, although it is not you who moves it.

a second helping but of your choice or intention *to* refrain.

I'm interested in both analogies, since I'm interested in giving a univocal internalist explanation of the two species of akrasia.

### III

I'll offer two arguments. First, my account captures both the logic and the phenomenology of akrasia. More generally, as I'll argue in the next section, it offers the best explanation of the normative guidance on which internalists rightly insist in the face of externalist postulation of a faculty of the will.

Consider this more detailed case. You've just started on your diet this week, and you're dining this evening at your in-laws. Your mother-in-law has concocted a fabulous cherry pie for dessert, whose potential she insists cries out for a scoop of ice cream. Did your diet-concluding or diet-intending self foresee this predicament? If not, why should you trust it now? If so, why should you trust a self that judged best or intended what, given your social awkwardness, will certainly amount to regrettable rudeness? Yet you know that there will always be such predicaments. The fundamental question is not how to avoid misweighing your best judgment as you form an intention, or how to maintain the strength of that intention, but whether to trust the self that drew that conclusion or formed that intention. Perhaps this was not the week to start your diet. Perhaps your diet needn't involve the renunciation of every dessert. Or perhaps it's time you simply forged ahead toward this goal and began awkwardly to learn the art of gracious refusal. There are of course these and other questions to deliberate, but a prior question is non-deliberative. It is a question of how to feel toward the earlier self whose judgment or choice

created the predicament in the first place.<sup>7</sup>

The need for reasonable self-trust is logical: it follows from the assumptions that you can't act until you close deliberation and that acting is not only the causal but the rational upshot of deliberation. You'll be redeliberating even insofar as you consciously wonder whether to trust your judging or your intending self. And if you're redeliberating – at any level – you can't be intentionally acting. (Of course, you can conclude deliberation without feeling certain; you need merely be as confident as the deliberative context requires.) As the rational upshot of a closed deliberation, action must therefore be subject to rational guidance that is non-deliberative. As non-deliberative, trust must function as a counterfactual sensitivity. If you search for reasons to rely on someone – their reliability in the past, the risks of not relying on them, etc. – you are to that extent not trusting them. We often rely on people for such reasons, and sometimes we rely on ourselves for such reasons. But that is a very different inter- or intra-personal relationship from trust. In fact it is not really a *personal* relationship at all, since we rely on things like ladders and bridges in just the same way.

But the account also captures phenomenology. However conflicted you may still be about whether to  $\phi$ ,<sup>8</sup> your commitment to act acquires only as much guidance as a concluded deliberation can provide, and your action in turn acquires only as much as that commitment can

---

<sup>7</sup> Note that phenomenology doesn't clearly distinguish the two kinds of akratic break. When you akratically took that second helping was it because you were weak in the face of temptation or because you formed a new intention out of accord with your best judgment? While there are clear enough cases on each side to make the distinction between incontinence and weakness important, there are many cases in the middle where either description is as good as the other. That is another reason, beyond the appeal of theoretical economy, to hope we can account for them in common terms. However we explain them, the two species of break should come out irrational for the same kind of reason.

<sup>8</sup> Again, you can conclude deliberation without feeling certain. You need merely be as confident as what's at stake in the deliberative context requires.

provide. Both provisions pose questions for your self-relations: should you trust yourself? It is never too late to rebel against the authority of that self, though if it is too late to stop the action the rebellion may not prove successful. We don't usually regard an agent's last-minute 'Egad, what am I doing?' as transforming the action into a mere compulsive spasm, but we do give him some credit for jumping off. (Of course, sometimes the exclamation is an exercise in hypocrisy, and we blame him both for the action and for this other action of trying to escape blame.)

If you don't trust your judgment or choice, there are two possible cases. In one, you reopen deliberation on the question whether or when to diet. In the other, you don't. Mistrusting your judgment or choice without reopening deliberation is irrational but not uncommon. That's what happens in each type of *akrasia*: you don't abandon your judgment that you should diet, or your intention *to* diet, but neither do you trust that judgment or intention. This is irrational because the attitudes of trust and mistrust are constitutively responsive to a rational norm. Trust is not the conclusion of a deliberation, but neither is it mere acquiescence. As I've argued, it is guided by a counterfactual sensitivity to evidence of untrustworthiness, triggering which would reopen deliberation. Mistrust therefore rationally requires deliberation. If you mistrust but do not reopen deliberation, your mistrust is pathological and therefore irrational.<sup>9</sup>

It may seem that the second *akratic* break does not have this structure, since weakness of will requires merely an unreasonable revision of one's intention, not the deeper irrationality of

---

<sup>9</sup> In this respect trust functions like any emotion. Anger, fear, love, and compassion are pathological and therefore irrational if they do not, other things equal, lead you to act in certain ways in appropriate circumstances, where the emotions themselves determine criteria of appropriateness. If your reliance is trust, then you are necessarily irrational to the extent that you do not deliberate or redeliberate when confronted with evidence of untrustworthiness. As we've seen, you may nonetheless wind up relying on the person, or even on yourself, but that reliance will, unlike trust, be deliberated and cannot perform the role of trust in mediating the relation between deliberation and action. Where you mistrust, you can only deliberate how to proceed.

rejecting it while failing to reconsider.<sup>10</sup> No doubt many cases in which the agent unreasonably redeliberates are worth chastising as ‘weakness.’ But the irrationality that puzzles is not just any unreasonable failure to follow through on an intention but a failure to follow through *without reconsidering* the matter. I’m arguing that we begin to understand such cases by observing that the non-deliberative sensitivity that makes agency possible can be triggered without reopening deliberation.

#### IV

Now for the argument from explanation. As Gary Watson notes, externalist accounts of agency run the risk of conflating incontinence with existentialist ‘radical choice’: “On the externalist view... going against reason must always come down to a choice among possible commitments.”<sup>11</sup> Watson goes on to argue that we need to posit a faculty of the will to explain the possibility not of counter-normative agency but of normative uncertainty or indeterminacy, as for example when your reasons do not determine means to your end. But even here, he holds, your will is guided by judgment at least to the extent that “[w]hen intention fails to be guided by judgement, it fails to operate in its executive capacity – it fails to operate *as* a will.”<sup>12</sup>

I agree with this last claim, but I don’t think Watson or any other internalist has explained why it’s true. Until we understand just what it is for judgment to guide or to fail to guide

---

<sup>10</sup> Holton pursues this thought at length in “Intention and Weakness of Will,” *op. cit.* The debate between his position and mine depends on our different views of intention. Holton accepts the outlines of Bratman’s view. In “Trust and Diachronic Agency,” *Noûs* 37 (2003), and in the full version of this paper I give grounds for preferring a view that highlights the intrapersonal side of intention that I’m emphasizing here.

<sup>11</sup> “The Work of the Will,” *op. cit.*, 181. By ‘weakness’ in that passage Watson means what I’m calling incontinence; i.e. he’s only discussing the first akratic break.

<sup>12</sup> “The Work of the Will,” 183.

intention we won't understand what it is for an intention to operate or to fail to operate as a will. And if we don't understand *that*, then we won't understand akrasia. So until we have this explanation, the externalist is entitled to argue that only the postulation of an *externalist* faculty of the will, and of will-power, can explain the possibility of akrasia. We thus return to the two pressing questions for the internalist: How does judgment guide choice (or intention)? And how does choice (or intention) guide action? The question in each case is not *whether* it does but *how*.

My account has the virtue of providing these explanations. Rational guidance has two features: rational authority and motivational efficacy. A rational guide is simply a rational authority that as such motivates. On my view, judgment is rationally authoritative for choice (or intention), and choice (or intention) rationally authoritative for action, insofar as each is trustworthy in relevant ways, where such trustworthiness registers in the trusting not as a positive attribution but as the failure to trigger a counterfactual sensitivity to evidence that it should not be attributed. Again, it's just this that makes agency possible in the first place, since agency requires, in addition to the deliberative weighing of reasons, a rational receptivity to reasons taken as weighed – in other words, not only activity but this special sort of passivity. Much remains to be said about the specific nature of this authority.<sup>13</sup> My present claim is only that internalists need to explain these species of non-deliberative rational authority, and that my account at least points in the right direction.

Its explanation of rational authority entails an explanation of how that rational authority

---

<sup>13</sup> I don't claim that the authority *is* the species of self-trust just described, merely that it *registers* as such trust. In the full version of this paper, I suggest that the authority is aretaic and that a fundamental role for virtue thus lies in making agency possible.

can motivate. For trust already includes a motivational propensity: you trust someone – including yourself – only to the extent that you are disposed to let the person influence your motives. So if it is something internal to trusting receptivity that explains how your judgment authorizes your choice or intention, and again how your choice or intention authorizes your action, we need add nothing to get an account of motivational efficacy.

We can explain akrasia as what happens when you mistrust the source of rational authority over your motives without reconstituting that authority through further deliberation. Your motives thus come unglued from what authorizes them, but in a way that is intelligible – since you would not have been irrational had you taken that extra step and reopened your mind. You judge that you shouldn't have dessert tonight but fail to choose accordingly. Or you intend to pass on dessert but fail to act accordingly. We can imagine a variant on each case that would not be akratic: it occurs to you that your sugar cravings are unexpectedly interfering with your concentration, and this evidence of your untrustworthiness in judging or intending (since you didn't foresee the problem) leads you to wonder if tonight is really a good time to begin your diet. That is, you come to mistrust your judgment or intention, and the mistrust points a path forward, practically speaking, by reopening deliberation. In the akratic version of the case, by contrast, though you mistrust the judgment or intention your mistrust does not reopen deliberation. The mistrust deauthorizes the judgment or intention – in the sense that you do not treat it as authoritative – but you do not reconstitute that authority by redeliberating and reaching the conclusion that you should not start your diet tonight. It is intelligible that you should fall into this species of irrationality simply because changing your mind has two distinguishable steps. Though they frequently happen simultaneously, we can distinguish the step of mistrusting your judgment or intention from your first step toward a different deliberative conclusion.

Sometimes, after all, you mistrust your judgment or intention without knowing how to go about approaching the question differently. Sometimes, that is, you continue to judge that you should  $\phi$  or continue to intend to  $\phi$  despite the fact that you mistrust that judgment or intention. It ‘feels’ wrong to begin your diet tonight, for example, so you don’t – even while continuing to judge that you should or intending to. Your intelligible mistake is to let your mistrust take you only halfway toward a change of mind. It’s as if the mistrusted judgment or intention were someone else’s, and you could fulfill your agential responsibilities simply by resisting it. ‘I don’t trust that influence,’ you seem to be saying – without realizing that the influence is your own.<sup>14</sup>

We can likewise explain the possibility of choosing under judgmental uncertainty or indeterminacy in terms of what is not only actually the case but a truism: that trusting your judgment takes you only so far. You judged that you should grab a box of Oat Flakes; you didn’t judge which box. You judged that you should knock the robber to the ground; you didn’t judge precisely how. Still, when just that box winds up in your cart and the robber winds up on the ground nursing just that bruise, it was because of your choice. We can agree that judgment does not – indeed, cannot – authorize every aspect of what you do when you act while nonetheless

---

<sup>14</sup> The normativity of trust relations doesn’t in general preclude such a refusal of trust. Simply refusing another’s invitation to trust without deliberating what to do instead can be perfectly reasonable. (‘Hey, get off that couch and let me give you a leg up to this bird’s nest.’ ‘Um, no.’) But when the one inviting your trust is you yourself, you rationally can’t just refuse. You can rationally refuse only by reconstituting the authority by appeal to which you issued the invitation. The difference does not lie in any deep self-other asymmetry but simply in the fact that the invitation in the second case is your own. You have to be responsive to it because you *are* responsive to it insofar as it is yours. We aren’t talking about a case in which the judgment or intention operates subconsciously, or in which you’re confused about your identity. From this angle, your mistake in akrasia would lie in a form of self-deception about your identity: you pretend to yourself that your own judgment or intention can be rejected as if it were that of another. You don’t like the fact that you have committed yourself to dieting this evening, so you pretend that the commitment is a coercive intervention from a perspective not your own. (This pretence needn’t rise to the level of self-deceptive belief, so this wouldn’t amount to explaining practical akrasia in terms of doxastic akrasia.) I don’t claim that the pretence need be in play. This is merely a way of making the pathological trust relation vivid.

insisting on an internal link between judgment and action. We thereby explain how the exercise of judgment makes choice or intention function as your will. Your will is what determines your actions as such, not what determines your every movement when you act. (In like manner, choosing the box on the left or to grab the robber's right leg doesn't resolve every question of muscular contraction.) Insofar as choice or intention fails to be guided by the authorization of judgment, either because you choose or intend counter-normatively or because of normative uncertainty or indeterminacy, it isn't functioning as a will. Counter-normative agency is thus an explicable by-product of agency, while choosing where judgment gives out merely reveals that agency unfolds at a level that is less fine-grained than the levels at which we can be described as either doing things or choosing to do things (just as the level of agency is much less fine-grained than the level at which we can be described as contracting our muscles).

My thesis, then, is twofold: your choice or intention functions as your will when it expresses an executive authority that derives from its trusting receptivity to the verdict of your judgment, and your behavior manifests your will when it manifests a trusting receptivity to the influence of your choice or intention. Choice or intention registers thus receptively the authority of trustworthy judgment, making it motivationally efficacious. My main argument is that this thesis best explains the possibility of the two species of akratic break.

## V

It may seem wrong to say that you mistrust your judging self when you fail to choose or intend as you judge best, since you needn't deem your judging self untrustworthy. Let me conclude by explaining why mistrusting someone, including yourself, is not the same as believing the person

untrustworthy.

Consider this case. You deliberate and reach the judgment that you ought to get right to work this morning. Then without reconsidering the matter you choose to linger over the newspaper, intending to keep reading merely, as you put it, “till I finish this cup of coffee.” Perhaps you do form the intention to get to work but ‘find yourself’ nonetheless giving in to the temptation to linger. The coffee is of course almost cold, and you know you have no disposition to finish it. But you continue to judge, all things considered, that you ought to work this morning.

On my account, your mistake lies in the fact that you fail to trust your judgment but without doing what follows from that mistrust, namely reopening deliberation on what to do this morning. Of course, one reason why you don’t reopen that deliberation is that you don’t believe you made a mistake when you conducted it earlier: you continue to judge, all things considered, that you should work this morning. Given this judgment, how could failure to reconsider be a mistake? The problem is that you’re failing to let yourself be governed by that judgment. That is, you’re failing to trust it. Since it’s your judgment, you’re therefore failing to trust yourself. You don’t, of course, *believe* yourself untrustworthy. In fact, you believe yourself trustworthy. But that belief is irrelevant to whether you trust.<sup>15</sup>

Compare an interpersonal case. ‘Go on up,’ your climbing instructor assures you. ‘I’ll

---

<sup>15</sup> Such doxastic judgments may or may not inform trust and mistrust, whose criteria of appropriateness, like those of other emotions, govern the subject without figuring directly in the subject’s judgmental attitudes. Just as love presents its object as lovable without entailing any judgment of lovability, so trust presents its object as trustworthy without entailing the judgment that she is. Just as you can without doxastic inconsistency love or hate those whom you do not believe worthy of the emotion, so can you trust or mistrust where you do not believe the object deserves it.

use this rope to keep you from falling if you lose your grip on the rock.’ You trust him, deeming him highly competent, so off and up you go. Five feet up, however, you stop. You can’t bring yourself to go further. It seems you don’t trust him after all – but without deeming him untrustworthy. You can’t imagine a more competent climbing partner and wouldn’t know how to go about redeliberating whether he’s worthy of your trust. Still, you can’t bring yourself to trust him. Your judgment simply doesn’t govern your will.

This sort of thing happens so often that we rightly don’t find it at all paradoxical. Trusting another requires more than the judgment that he or she is worthy of your trust. Yet we puzzle over the intrapersonal analogue: how could you fail to trust your judgment when you don’t deem your judgment untrustworthy?

My proposal is that we model the intrapersonal on the interpersonal in two steps. First, note that you might judge and even feel perfectly confident that your climbing skills are up to the route you’ve mapped and yet – perhaps spooked by a recent fall – find yourself unable to trust them. Here you fail to trust yourself in exactly the way you failed to trust your climbing instructor. Second, make the competence not a physical competence figuring in how you will execute the action but a mental competence figuring in how you deliberated whether to perform it in the first place. The parallel nonetheless holds. Just as you can mistrust your own physical capacity without deeming yourself unworthy of that trust, so can you mistrust your own judgmental capacity without deeming yourself unworthy of that trust. After all, you could as well mistrust your instructor’s judgment – say, on whether this route is within your competence, or simply on how distracted by conversation to let himself get while spotting you – as his physical ability to keep you from falling.

A pithier way to make the point is to note that if you mistrust your judgment on whether you should  $\phi$  it isn't surprising that you might at the same time judge that you are trustworthy. To mistrust your judgment in this respect entails mistrusting your judgment that you are trustworthy in this respect. If you are not receptive to your own judgment on the question whether you should  $\phi$ , you can't expect to make yourself so by reflecting that you *judge* that you should be receptive. This is what we should expect if the problem is inappropriate receptivity: you can't make yourself appropriately receptive by trying harder. The problem is that you're trying too hard – manifesting discounted worries as anxiety and generally second-guessing yourself – and you need to find a way to let this activity subside.<sup>16</sup>

---

<sup>16</sup> Continuing the theme of second-guessing, let me sketch anxious replies to some other objections. First, we might also wonder about cases in which you fail to be guided by your own judgment only because you are gripped instead by an arational force such as a compulsion. What if when you give in to the temptation to linger over your newspaper, what you 'find yourself' doing is not reading but obsessively rehearsing an embarrassing memory of which an article has reminded you? You'd like nothing more than to relax and let yourself be guided by your judgment that you need to get to work, but as long as this force has you in its grip you simply cannot. Is it correct in this scenario to say that you 'mistrust' your judgment? It seems not: *you* are not doing much of anything in an agential sense except trying to resist this force that is acting upon you. (I assume that the obsessive memory not merely masking an irrational mistrust in your judgment, providing you with a convenient excuse: "Hey, I'm not lingering but in the grip of a compulsion here!" I assume that your obsessive memory is more like a compulsion to count the 'the's on the page than like daydreaming.) So if you fail to intend as you judge or to act as you intend in *this* way, it does not count as the philosophically perplexing sort of akrasia – that is, as the sort I am trying to explain. (Or we could reserve 'akrasia' for the perplexing sort of break and say that this is not akrasia at all.) There is nothing philosophically perplexing along these lines about compulsive action or ideation. The mental activity or behavior in question simply does not qualify as choice, intention, or action.

Still, do I not owe an explanation of the distinction on which I am relying, between a compulsive or otherwise arational failure to trust and the irrational mistrust distinctive of akrasia? Well, I have just given it. At least, I have explained my entitlement to rely on the distinction. We begin with a puzzle over how it is possible for you to choose or intend contrary to your own best judgment, or to act contrary to a choice or intention informed by that judgment. A grasp of that puzzle presupposes a grasp of the distinction in question, since the puzzle presupposes that you are not merely gripped by a compulsion or some other arational force. Noting this feature of the puzzle does to some extent explain the distinction, since it articulates a dimension of the arational/irrational distinction in play here. It is not, of course, a full explanation of that distinction. But in the dialectical context at hand it is all the explanation one owes.

One might worry that my account renders all akrasia merely arational, since I have not explained how akratic choice, intention or action could qualify as such (as opposed to something like compulsion), given that it fails to be informed by actual judgment. (As I argue in the full version of this paper, it is informed by an 'as if' judgment: you do act as if your best judgment supports you.) Here again I need merely clarify my explanandum. I regard the phenomenon of akrasia as most fundamentally manifested in the akratic *break*, which I explain as an irrational failure to trust your judgment. This failure is most strikingly present when the akrasia occurs by omission,

---

which is why I'm now focusing on a case in which the akrasia is most naturally described as a deficit: you judge that you ought to do something (get to work) other than what you're doing (reading the paper), and the question is how you could fail to treat that judgment as a rational guide. In such a case we don't need to explain how you could continue to read the paper, since that's what you were doing anyway. What we need to explain is how you could fail to at least try to get to work. To the worry that I haven't explained how the akratic pie-eating in my earlier example could count as anything but a compulsive spasm, I reply that I am not here giving an account of attributable agency in general. That akratic pie-eating falls short of a paradigmatic instance of attributable agency does not entail that it falls short of such agency altogether. An account of akrasia should explain how each species of akratic *break* is possible. It needn't add up to a theory of attributable agency.

In the full version I also address a deeper objection: it may seem that my account ignores present-directed intentions and generally imposes a temporal ordering among the elements of rational agency which doesn't define them. Can't I simultaneously judge that I ought to  $\phi$ , choose to  $\phi$  and  $\phi$ ? The acts needn't, it seems, unfold in a temporal sequence. Answering this objection requires a recodification and elaboration of my position, which in turn provides the framework for an account of the aretaic basis of practical-judgmental authority over the will.