

THINKING, ACTING, CONSIDERING¹

Daniel Muñoz

This is an Accepted Manuscript of an article published by Taylor & Francis Group in the Australasian Journal of Philosophy in 2017, available online: <http://www.tandfonline.com/doi/full/10.1080/00048402.2017.1327547>

Abstract

According to a familiar (alleged) requirement on practical reason, one must believe a proposition if one is to take it for granted in reasoning about what to do. This paper explores a related requirement, not on thinking but on acting—that one must accept a goal if one is to count as acting for its sake. This is the acceptance requirement. Though it is endorsed by writers as diverse as Christine Korsgaard, Donald Davidson, and Talbot Brewer, I argue that it is vulnerable to counterexamples, in which agents act in light of ends that they do not accept, but are still merely considering. For instance, a young professional may keep a job option open not because she definitely wants or intends to take it, but just because she is considering taking it. I try to show (1) that such examples are not easily resisted; (2) that they present challenges specifically for Brewer, Davidson, and especially Korsgaard; and (3) that the examples also raise fresh, non-partisan questions in action theory. What is considering, exactly? How could it fall short of acceptance while still guiding behavior? How can we act for an end before thinking it through?

Keywords — intentional action, practical reason, philosophy of action, desire, considering

¹ Thank you to audiences at MIT's Ethics Reading Group, MIT's Work in Progress Seminar, and the University of Texas at Austin's 2017 Graduate Conference, where I was fortunate to receive comments from Caroline Christoff. Thanks as well to Austin Baker, David Balcarras, Brendan de Kenessey, Branden Fitelson, Kelly Gaus, Elliot Goodine, Caspar Hare, Abby Jacques, Rose Lenehan, Yael Loewenstein, Zelda Mariet, Alex Prescott-Couch, Brad Skow, Jack Spencer, and Steve Yablo. I owe special thanks to Nathaniel Baron-Schmitt, Jonathan Drake, Richard Holton, Leonard Katz, Sarah Paul, Tamar Schapiro, Patrick Quinn White, and two anonymous referees at *AJP*. Kari Rosenfeld's insights and encouragement during the process of revision were invaluable. I should also like to thank the editors of *AJP* for their suggestions, which inspired many salutary deletions. What remains of the paper is dedicated to Kieran Setiya. I could not ask for a more thoughtful teacher—or a more considerate advisor.

1. Introduction

What's the difference between an intentional act and a mindless event? One fine classic answer, inspired by Anscombe [1963], is that intentional acts are open to a special form of *explanation*, called 'rational' or 'motivating' explanation, whose purpose is to show how an action serves its goal (see also Vogler [2002]; Thompson [2008]).² Why is Kathy petting Sparky? In order to calm him down (goal: canine placation). Why is Yael taking the trolley? Because that's how to get to today's philosophy conference (goal: metaphysical mayhem).

But notice that some telic explanations have nothing to do with motivation [Frankfurt 1978; Collins 1984]. Why is my immune system producing antibodies? To ward off a repeat infection (goal: continued health). Why do your pupils dilate as the room darkens? Because your retinas need more light (goal: sustained vision). These aren't *motivating* explanations, simply because they aren't about what *moved an agent*. 'Moving' in the relevant sense means something like 'persuading' or 'rationally compelling'; persuasion needs a modicum of awareness and some change of heart; so, the 'moved' agent must be at least dimly aware of her goal and the means she takes to it, and the goal must be something she desires (however weakly) or thinks important (if only barely). Motivating explanations thus appear to be constrained by:

The Acceptance Requirement

A goal ('aim', 'end') rationally explains ('guides') an action only if its agent *accepts* the goal, in the sense of seeing it as somehow good or required, believing that it is worth pursuing, willing it, or having some desire or intention to bring it about.

This condition on rational goals sets them apart from the goals of mindless processes, and

² For now I ignore *responsive* actions, which are explained by features of situations, not goals.

thereby helps distinguish intentional acts from eye dilations, supernovas, and other agent-free phenomena.

The acceptance requirement is widely believed and seemingly obvious. I argue that it is nonetheless false: we are sometimes moved by ends that we do not yet accept, but which we are only *considering*. For example, I might take the long route home, which runs past Flour Bakery, simply because I am considering stopping in for a cup of soup (goal: a soupy morsel). This could be true even though I lack settled opinions on the prospect of soup-drinking. I may not anticipate hunger. I might not quite desire relief from souplessness. But even if I know what my soup-experience would be like, and even if perfect introspection uncovers no intention to achieve such an experience, I can take steps toward the bakery in service of the goal of getting soup. These steps would be more than arational spasms: they would be intentional, motivated actions, not reflex twitches or sub-rational compulsions or symptoms of an unwell mind. True, the prospect of a cup of soup would not ‘persuade’ or ‘compel’ me. And yet it would guide my movements, giving them purpose and collecting them into a course of action.

The big-picture moral is that the metaphors of persuasion and compulsion are too severe. Sometimes our motives are gentler—nudging, suggestive, tentative, playful, flirtatious. Still they move us. My goal is to argue for this point, and against the acceptance requirement, with a range of examples like the soup case. There are responses, which aim to show that there is no real inconsistency between our cases and the requirement. (The main response: one might insist in the soup case that there is an end which I must accept—*having the option of getting soup*.) I argue that the inconsistency is nonetheless real. (Probably, I’m only interested in keeping open the option of soup *because I am considering getting a cup*—so the unaccepted end of getting soup

still rationalizes my behavior. More on this in section 4.)

It might sound like I am advocating something toothless, a mere metaphor-swap. But the acceptance requirement is baked into a huge array of moral psychologies and cannot be excised without leaving bite marks. For Korsgaard's Kantianism, I will suggest that the bite might be large. But first, is it true that the acceptance requirement is widely believed? By whom, and for what reasons?

2. The Acceptance Requirement

The acceptance requirement, besides being a helpful mark of the intentional, ought to sound plausible in its own right. Of course agents won't be moved by 'goals' that they don't accept! If we were indifferent or hostile toward an end, why would we bother pursuing it? Why roll unless inclined? Basing an action on an unaccepted end would be like basing beliefs on the testimony of someone we distrust completely: it would be unintelligible.

If anything, calling the requirement 'plausible' is an understatement—as if it were a thesis under debate, not a platitude believed from the start. It must have sounded platitudinous to Davidson [1963: 685], who endorses it at the start of Section 1 of his first classic work in action theory:

A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action—some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable.

By 'rationalizes', Davidson means 'motivationally/rationally explains'. So he is saying that an action can be rationally explained only if there is something about it or its consequences that the

agent desires, values, or thinks worthwhile—in short, something the agent *accepts*.

Notice how tolerant this notion of acceptance is. There are considerable differences between ‘seeing something’ in an action (perceptual), thinking an action agreeable (cognitive, non-moral), thinking it dutiful (cognitive, moral), desiring it (conative), and prizing it (evaluative). Such attitudes may also be directed either toward the action itself or toward its goal, and might or might not rest on false beliefs. The acceptance requirement is not fussy about these distinctions. It says only that a motivating explanation must cite *something* definitely positive—in generously wide senses of ‘something’ and ‘positive’—in how the agent conceives of the action. This generosity is part of why the requirement can seem undeniable.

In fact, I don’t know of any deniers; instead the requirement is embraced by philosophers of every stripe.³ Here is Talbot Brewer [2009: 28], an Aristotelian and among the last people one would expect to find allied with Davidson (see also Brewer [2009: 29, 37–9]):

To be an agent is to set oneself in motion (or to try to do so, or to adopt the intention of doing so) on the strength of one’s sense that something counts in favor of doing so. That performing some action would bring about some state of affairs cannot intelligibly be regarded as counting in favor of performing the action unless one sees the state of affairs, or the effort to produce it, as itself good or valuable. Hence if we are to view persistent attempts to bring about some state of affairs as the doing of an agent, we must suppose that the agent sees something good or valuable about attaining or aiming at that state of affairs.

Brewer articulates a line of thought similar to the Anscombean line I opened with: start from an idea about agency, end up at the acceptance requirement. (Though I was talking about specifically intentional agency.)

³ Philosophers do, however, deny requirements involving more specific forms of acceptance. Stocker [1979], for instance, denies that agents must always see something *good* about their actions.

A third parallel thought-line runs through the moral psychology of Korsgaard's Kant. To act is to 'determine yourself to be a cause'; in order for you *yourself* to be the cause, rather than your incentives, you must be willing as something 'over and above all your incentives'; and 'in order to do this, Kant believes, you must will your maxims as universal laws' [Korsgaard 2009: 72]. What this amounts to: in order to act, you must 'universally' will a means-end pair, which not only involves accepting an end for oneself, but also taking the end to be the sort of thing that anyone should accept.⁴ This pushes the persuasion/compulsion metaphor into the realm of legislation—compulsion from authority.

Finally, we have the 'Humean' view that action is a body movement caused by a belief-desire pair—specifically, a desire for a state of affairs, plus a belief about which body movements would be effective means toward the desired state [Davidson 1963; Smith 1987]. This view gets us to the acceptance requirement fairly quickly, once we add that desires are the dominant member of belief-desire pairs—the real causes. A Humean line: motivating explanations *are* desired ends (or the desires themselves); desiring is a way of accepting; so, motivating explanations are all and only accepted ends (or desires for them).⁵

What good fortune: Humeans, Kantians, and Aristotelians have some common ground! I worry, however, that this convergence is due not to a shared discovery, but a common

⁴ A quick qualification. For Korsgaard's Kant, I cannot act unless I 'make a law for myself with universal normative force, a law that applies to my conduct over a range of cases', but this force needs only to be 'provisionally universal' (Korsgaard [2009: 73]; see also Kant [1785 (1998): 56–8]). This makes it sound like Korsgaard stops short of the acceptance requirement—one counts as acting so long as one has at least *provisionally* willed an end (provisional willing is *not* acceptance, since it indicates a less-than-determinately-positive outlook on the end). But what is provisional according to Korsgaard is not *whether* an end is willed, but *how far* the willing ranges over yet-unconsidered cases.

⁵ The parenthetical remarks are a courtesy to *psychologists* about motivating reasons (section 5).

oversight. When thinking about agency, we tend to imagine a certain Practical Hero: righteous, omniscient, patient, placid, physically able. By singling out and peeling away such layers of idealization—and tracing them to attitudes about culture, class, and gender—philosophers have come up with challenging insights about the nature of action. I’m suggesting that we peel another layer. The Practical Hero is an eerily *decisive* person, who not only knows what he wants, but has definite wants about everything he knows. When we turn toward less Heroically decisive (and John Wayne-ishly masculine) agents, the acceptance requirement begins to stretch the truth.

3. Guidance Without Acceptance?

Can agents be guided by unaccepted ends? Consider this example, somewhat more dramatic than my taking the long route home with soup in mind:

The Job Interview

Leona is considering working at Goldman Sachs. She knows what life at Sachs would be like—the salary, the social status, the complex impact of her work on the nation. Yet her political leanings make her unsure whether such a career would be desirable; she feels uncertain. Naturally, she wants to keep her options open until she can make up her mind. But if she misses today’s interview, the option of working for Sachs will close up. So, because she is considering taking that option, she goes to the interview.

What guides Leona’s action, on most realistic ways of filling out the case, is the end of becoming a banker. (Or perhaps it’s *being* a banker—‘end’ is notoriously ambiguous between act and consequence [Kolodny and Brunero 2013].) Banking at Sachs is the goal served by interviewing. But this goal doesn’t have to be something that Leona accepts.

On the contrary, the natural thing to say about Leona is that her ends are unsettled. It is not just that she lacks information about her prospects and preferences. Rather her preferences

themselves are wishy-washy—hesitant, undecided. This wishy-washiness is reminiscent of what happens when we suspend judgment on a proposition, in the sense that we have considered it, but do not believe it or its negation, and do not merely award it a middling credence.

Similarly, Leona *suspends her conative attitudes*. I don't mean that she performs some voluntary action, 'suspension'. I mean that (i) Leona doesn't want the job; (ii) she doesn't want *not* to have the job; and (iii) whether she gets job is not *irrelevant* or *neutral* from the point of view of her motivation. Intuitively, one's conative attitudes are suspended wherever one's desires and intentions are unsettled.

Here are some more cases, for variety's sake. A politician meets with the party's cognoscenti because she's considering running for senator. A young couple saves money because they're considering buying a house. A meat industry insider cracks open an ethics textbook because he's thinking about going vegetarian. In all three cases, agents base their actions on ends they don't accept. Their actions don't seem like a result of a settled outlook—the result of compulsion or persuasion by pre-established ends. If anything, the actions *contribute* to the process of working out an outlook, either by allowing for new options (saving money), revealing facts about the options (consulting cognoscenti), or putting the options in new light (reading ethics).

These cases and Leona's are richer than the original soup case (in which I take the long route home with soup in mind—not exactly life-changing stuff). But the new cases are less airtight, and so potentially less telling against the acceptance requirement. Too many questions arise. Focusing just on Leona, what are her options, besides a job at Goldman Sachs? Won't

there be *something* making the Sachs life seem more desirable than her other options, if only in a single respect? What about factual uncertainty—is it really so easy to imagine a decision like hers from the inside while factoring out so much doubt?

We can address these questions by making the case explicitly contrastive, with just one relevant contrast. Suppose that Leona is deciding between her current banking job and the Sachs job, and the only relevant difference is that Sachs employees enjoy a massive boost in social status (the Goldman name is better known). Leona is wary of this boost: she sees how it *could* be desirable, but (thanks to inchoate egalitarian stirrings) she also sees the boost as potentially bad or shameful—‘I’d get *that* kind of recognition, just for moving money around?’ This ambivalence is reflected in Leona’s desires, which are provisionally pro-Sachs as well as provisionally anti-Sachs. There is thus no dimension along which she definitely prefers either job, or thinks either one more choiceworthy.

But even this case bristles with distractions. Won’t there be some perks of a boosted status—improved self-esteem, the pleasures of impressing friends and strangers—that Leona would definitely desire? I am supposing that these are irrelevant to Leona’s motivation, but you might think that this supposition makes the case too unrealistic to be useful.

Let’s try a final case, with less room for distraction:

Two Sculptures

David, a young artist, must decide by tomorrow which of two sculptures to showcase in his exhibition. The two are similar except in one respect: the second was made with a certain experimental technique. David is sure that audiences won’t care much about the difference, but *he* cares, and he is not sure whether the technique is a gimmick or a good idea. So when asked, David hesitates to commit early to showcasing #2. He requests another day to think, rather than plumping for #2 or flipping a coin, because he is considering going with sculpture #1 instead.

David's views on the two sculptures aren't fully worked out: his desires and evaluative beliefs do not settle the question of whether or not to commit now to #2, since he does not yet accept the choice of either sculpture over the other. Still, David recognizes that there is something *potentially* more desirable about the end of choosing #1, and so this end can prompt him to request more time to think. The unaccepted end of choosing sculpture #1 thus rationally explains David's request; there is nothing else doing the 'real' explaining.

If the acceptance requirement were true, we would have trouble making sense of David's request. But it is easy to see what David is up to. His attitudes don't settle the question of what to do, but this doesn't stop him from acting anyway. He can be motivated even where his feelings and convictions give out. Just as it is possible to reason theoretically in the absence of fully worked-out beliefs, it is possible to reason practically, and execute one's reasoning, in the absence of a full stock of desires, intentions, and normative convictions.⁶

We now have a range of cases: my getting soup, Leona's choice of career, David's choice of sculpture, the politician-couple-carnivore variety pack. (Leona's case is the most interesting to me; David's and the soup case are the cleanest; and I hope they all reinforce one another.)⁷ In

⁶ The analogy with belief is hard to make precise (without saying controversial things), but here's the sort of case I have in mind. A considered agnostic, mulling over a metaphysical view, realizes that it entails atheism, and on this basis suspends judgment on the view.

⁷ Bratman [1987: 38–9] and Holton [2008: 28–9] feature a pair of related examples. Here is Holton's version of the first: you want to move a tree off of your car and are deciding between removing it yourself and paying someone to do it. In the end, you make a cancelable appointment with a yard-worker and *also* load up your wheelbarrow with tools. You fully intend to move the tree, but only partially intend each means. This case is similar to Leona's: you load the wheelbarrow *because you're considering removing the tree yourself*. But there is a dissimilarity: whereas you have accepted the end of having the tree moved and are merely undecided on your choice of means, Leona is still making up her mind about her end itself. Though Leona's case is a natural extension of Holton's and Bratman's, her case alone flouts the acceptance requirement.

describing these cases, I have relied on the locution ‘because they’re considering...’ – why? First, because it does the job: it signifies that the agent is motivated by an unaccepted end (viz. whatever the agent is considering doing). Second, because nothing else does the job as well. ‘Leona is interviewing in order to get the job’ is (arguably) not quite true: it falsely entails that Leona accepts the end of having the job.⁸ ‘Leona is interviewing because she must do so to get the job’ is an improvement (definitely true!), but it doesn’t even suggest that Leona might not fully want the job. We do get this suggestion from ‘Leona is interviewing because she *might* want the job’. But ‘might’ can mislead us into thinking that we are dealing with something epistemic, a doubt about what Leona’s (non-provisional) desires are. What is relevant to the acceptance requirement is not doubt *about* desires, but something more like doubt *in* desires. This idea, I claim, is best captured by talking about what agents are considering.

4. Other Ends

To save the acceptance requirement, one would need a persuasive response to our counterexamples.

The most tempting response is that in each case, there must be some *other* end, which the agent *does* accept, that *really* explains the action. In particular, the most common response I get to Leona’s case is that what guides her isn’t the end of becoming a banker, but of keeping bridges unburned: she is aiming to *keep her options open*. No doubt, this is something that Leona

⁸ I won’t insist on this point, but I doubt that ‘One is doing A in order to do B’ really does entail a desire or intention to do B, though I think it entails that one is *at least considering* B.

desires (and so accepts).⁹ And it also explains her action: Leona goes to the interview because she is keeping an option open. So we have a rational explanation by an accepted end. Isn't that enough to square Leona's case with the acceptance requirement?

Not quite: there is a revenge problem. The end of keeping options open does explain why Leona interviewed. But what about the act of keeping options open *itself*? Normally, this is not something done for its own sake. If I want to avoid burning bridges, it's not because I care about the bridges: it's because I might want to cross them.¹⁰ In a similar way, Leona is only interested in having the option of working at Sachs because she is considering taking that option. Why else would her desire for options depend on how her consideration plays out? For suppose that Leona has finished considering. If she now *wants* the job, she won't care about having the *option* of taking it except instrumentally.¹¹ If she's now certain that she *doesn't* want the job, then she won't *at all* want the option of taking it. The upshot is that her feelings about having options are not fundamental. To claim that Leona wants to have options for its own sake is to misplace the force of her motivating ends.

What about the possibility that other, grander ends might lie behind Leona's action?

⁹ This might be false on some senses of 'desire', involving affect or a relation to our psychological reward system [Arpaly and Schroeder 2014]. Perhaps Leona feels cold and unenticed by the prospect of keeping options open. But she does desire options in the technical sense I have in mind, which involves *motivation*: roughly, desiring that *p* entails being motivated to bring it about that *p* if one can—whereas one who is considering doing something might not be motivated to do it. (This is consistent with, but weaker than, Dancy's [2000] claim that desire *is* motivation.) Thanks to a referee for pressing me to say more here.

¹⁰ An exception: sometimes, having options is desired finally because a lack of options would undermine personal freedom. But freedom isn't essential to Leona's case, and it certainly isn't relevant to the others.

¹¹ A *final* end motivates in its own right, whereas an *instrumental* end motivates only because it conduces toward something else [Korsgaard 1983].

Maybe she is really interviewing for the sake of *happiness, virtue, making the best decisions*, or *The Good*, and she is merely unsure whether these ends would be achieved by taking the job. I grant that most agents care about such ends; for all we've heard, Leona probably cares, too. But I doubt that *all agents always* care. Why would they? What could force someone like Leona to accept such ends, no matter what?

On the contrary, even our most natural desires can be extinguished by circumstance or depression. And yet someone whose desire for wellbeing is snuffed out might hesitate to act on a self-destructive impulse, not because she desires to avoid harming herself and her loved ones, but just because she sees avoiding harm as *potentially* important. The would-be self-harmer suspends desires for (or desires the opposite of) all grand ends in the neighborhood—wellness, happiness, beneficence—yet she needn't be powerless against her impulse. She might resist by provisionally accepting a nobler end.

Here's a similar case. Picture a saint who hasn't lost faith in her deity, but who no longer feels committed to serving it. The saint is still able to go through the motions at church—all the while guided by the end of service, not an ulterior motive like a fondness for hymns—because even though piety is no longer on her list of settled ends, it's still a leading candidate. She fears that a life of service might not be for her but still sees it as *potentially* for her. Thus the saint's attitude toward service falls short of acceptance, even though attitudes of acceptance may be fairly weak—like my intention to drink the light roast rather than the dark for my morning coffee (section 7). Still, the saint isn't *neutral* toward worship, as most of us are toward the goal of walking an even number of steps this year; nor is she *against* it. The saint suspends her conative attitudes. This sort of suspension, absent any argument to the contrary, should be

possible for any end no matter how grand, just as it's possible to suspend belief in any proposition no matter how sublime (minus oddballs like 'I never suspend beliefs').

The cases of the saint and hesitant self-harmer directly suggest that we sometimes suspend desires for even the most marvelous ends. What could be grander for the saint than religious service? For the would-be self-harmer, what end could lie beyond the wellbeing of herself and others?

So much for the idea that Leona and the rest must be working toward some accepted end that does the real explaining. But note that even if Leona's actions are explained by some accepted end, that wouldn't show that the unaccepted end of banking played no role. Some ends are both instrumental and final, and some acts serve more than one end. Leona might interview because she's considering the job *and* for some other reason.

5. Responsive Reasons

So far I've assumed that when an agent acts because she's considering doing something, she is guided by the end of doing that something. But what if the real explanation isn't even an end? The deep motive might be 'backward-looking', in the sense of Anscombe [1963]—or as I prefer to put it, the reason might be *responsive* rather than *telic* [Muñoz msc.].

A responsive reason to φ is a reason given by how things *already* are (or are set to be)—by a feature of the situation as it stands, like the fact that one has promised to φ , or more generally, facts about one's past acts and present relationships. A telic reason to φ , by contrast, is given how things *would* turn out if one were to φ —by features of outcomes, like the fact that φ -ing would make people happy, or that it would minimize wrongdoing. In short: telic reasons

tell us to pursue goals, whereas responsive reasons ask us to do what fits the situation.¹²

(A quick aside: if there are responsive reasons, then there is a responsive version of the acceptance requirement, as well as counterexamples analogous to Leona's. For just as agents can be ambiguously attracted to ends, they can feel ambiguously urged by features of the situation—the messy state of a relationship, a promise made under a cloud.)

How could responsive reasons save the acceptance requirement? Here's an idea, due to Kieran Setiya. What if we have responsive reasons to keep options open? Our reason could be the fact that we don't know which option to take. More plausibly, the reason could be the specific fact that the agent is considering doing this or that action—so, Leona's reason for going to the interview could be that she's considering taking the job. But now 'that she's considering the job' really *is* her reason, not just an allusion to guidance by an unaccepted end. We only thought such guidance was possible because we were confusing a typical responsive motive for an unusual telic one.

There are two problems with this response. First, it exaggerates the difference between the original case and one in which Leona does accept the end:

The Second Interview

Leona is done considering: she now intends to get the job at Goldman Sachs. To seal the deal, she needs only to attend a *pro forma* second interview, which she does.

Why does Leona go to the second interview? Well, now her goal really is to get the job, so she must be interviewing in order to do so. This is a telic explanation. She accepts the end of getting the job, and this end explains her action.

¹² My distinction is inspired by Raz [2011: 46], who claims that telic (his term: 'practical') reasons are grounded in value, whereas responsive ('adaptive') reasons are grounded in fittingness (see also [Müller 2011]).

But is this really so different from the original interview, back when Leona was only considering? Both cases seem telic. There is no change in Leona's focus from situations to goals. Granted, by the time of acceptance, the end of banking is Leona's only salient motive. But it does not seem like a brand *new* motive, replacing an old responsive one. Why posit such a sharp change? Consider what it feels like as *you* gradually slide from merely considering to fully intending. Do you swap out reasons, one telic for one responsive, or is there a single end to which you are slowly warming up?

The second problem is more complicated, but it starts from a simple thought. Isn't there something weirdly self-indulgent about Leona if her reason is that she's considering the job? (What's so important about *her* that her mere thoughts make actions worthwhile?) Usually when facts about our mental states motivate us, there's something unusual going on. For example: Sharon calls her family for support because she wants to start drinking again; Nathan seeks a psychiatrist because he believes he's being followed (which he recognizes, on some level, as paranoia); Ricky seeks help because he's considering something self-destructive. These agents are moved by facts about their own mental states because these states are bound up with things that matter—safety, sanity, health. Leona's mental states have no such notable connections; so, it's weird to say that the fact that she's considering the job is itself her reason to go to the interview.

Now the complication (this paragraph draws on [Dancy 2000], omitting discussion of 'psychologized' views). I have so far been assuming the truth of *non-psychologism* about motivating explanations. For non-psychologists, the real explainers of action aren't mental *states*, referred to by definite descriptions like 'my desiring coffee', but *propositions*, referred to

by that-clauses like ‘that the coffee is tasty’. But for *psychologists*, the real explainers are always mental states, especially desires and beliefs. Any explanation that alludes to a goal—‘he did it for the sake of victory’, ‘they did it because it maximized happiness’—is really a disguised explanation in terms of goal-directed conative states, like desires.

The point of this is that the two sides will differently interpret ‘because Leona’s considering taking the job’. For the non-psychologist, the explanation will sound responsive and oddly self-indulgent, since the bare fact that someone is considering an end isn’t usually so significant (though there are exceptions, like Ricky’s case above). There is nothing plausible about this explanation, so interpreted. The psychologistic interpretation is more promising: ‘because Leona’s considering taking the job’ is an appeal to a telic conative mental state—*considering*. This is a new type of psychologistic explanation (the usual sort appeals to non-provisional states like intending), but it is inconsistent with the acceptance requirement. For when we first stated the requirement, it was in non-psychologistic terms. In the psychologist’s idiom, it becomes:

The (Psychologistic) Acceptance Requirement

A conative attitude rationally explains an action only if the attitude is a form of *acceptance*—i.e. a non-provisional intention, desire, willing, or plan.

Considering, unlike desiring and intending, is not a way of accepting. So if Leona’s state of considering rationally explains her action, she remains a counterexample to the acceptance requirement.

6. Considering and Accepting

So far I’ve made a big deal of the difference between considering, on the one hand, and desire

and intention on the other. But what makes considering so special? To start: (i) considering involves suspension of conative attitudes—one feels hesitant but intrigued; (ii) considering isn't a static frame of mind, but a process or activity, which is why we naturally talk about considering in the progressive aspect; (iii) considering has a natural conclusion in (not action, but) the rejection or acceptance of an end; we say that we're 'done considering' not after *doing* the considered thing, but after *deciding* to do it. In sum, considering is an activity that aims at resolving our suspended conative attitudes.¹³

To be considering an end is more or less to be mulling it over, to be open to acceptance though not yet on the hook. But any end I am considering I can *provisionally accept*, conducting myself in its light, treating it as if it were my end, not (just) because I see some definite value in pretending, but because I see intriguing glimmers of what might be value in the end itself. This is the sort of thing that happens when one is trying on a new identity or, like the saint, clinging to a fading one. We explain an act by appeal to provisionally accepted ends by saying that the agent was 'considering doing' something or other.

Considering thus plays an explanatory role similar to that played by intention, though intention is more committal. This suggests an analogy between considering an end and what Bratman [1992] calls *accepting* a proposition—since acceptance plays something like the role of belief, but doesn't entail belief.¹⁴ To accept a proposition is to take it for granted in the context of

¹³ Notice that this first feature (suspension) might come apart from the others (working toward resolution). One might suspend judgment on *p*, for example, without wondering whether it's true, or trying to make up one's mind about it. More generally, it would be a good idea to ask which combinations of these features (and the four from Bratman that I discuss below) would form interesting, natural, or familiar 'clumps'. I owe these points to conversations with Kieran Setiya and Richard Holton.

¹⁴ I hope it's clear that I'm not assuming that there is only one sense of 'acceptance', on which we

practical deliberation. By default, our deliberation's 'cognitive background' — the propositions whose truth (or high probability) we take for granted — consists of what we believe. But given various practical pressures (e.g. not to overspill our busy brains) we must adjust our background, removing ('bracketing') some propositions and adding ('positing') others ([Bratman 1992: 4-8]; see also Ullman-Margalit [1983: 162] on 'presumption rules', as well as Cohen [1992]). The adjusted background is the set of propositions we accept.

What makes accepting different from believing? Bratman [1992: 8–9] cites four things:

Relativity

Reasonable acceptance, but not reasonable belief, is context-relative; when writing a letter home, the soldier doesn't accept that he will survive the day, though he does accept this when planning his evening.

Aims

Reasonable belief is shaped only by evidence; acceptance is sensitive to practical pressures, like high stakes and limited time.

Voluntariness

Acceptance, but not belief, is voluntary.

Agglomeration

Belief, but not acceptance, is subject to an ideal of agglomeration.

But notice that considering and intention *share* most of these properties:

Relativity

Reasonable intention and reasonable consideration are *both* context-independent. If one stops intending or considering something, that indicates a change of mind.

Aims

Intention and considering are *both* shaped by the value of their object (if anything).

Voluntariness

Intention and considering are *both* under direct voluntary control (given that deciding is

'accept' both ends and propositions. We have two technical uses, introduced for different reasons, though I think they are importantly similar in that both signal non-provisionality.

voluntary, and that decisions issue in intentions and polish off consideration).

Agglomeration

Intention, but not considering, is subject to an ideal of agglomeration: I may consider doing A and consider doing B without considering doing A & B both.

The analogy between considering/intending and accepting/believing, then, is not perfect. And the imperfections don't stop there. Recalling the features of considering we started with, (i) end-considering, but not proposition-accepting, is *provisional* by virtue of involving suspended attitudes; (ii) considering, but not accepting, is *processual* rather than stative; and (iii) considering *concludes with* intention or desire, whereas accepting a proposition does not, and often should not, 'conclude with' belief in it, and certainly not with disbelief.

7. Graded Attitudes

Agents like Leona base their actions on ends that they don't reject, accept, or take a neutral stance toward. These agents are just considering their ends, which I take to involve the suspension of conative attitudes like desire and intention. A competing account of such cases would say that they involve *graded* conative attitudes—the agents desire or intend to a certain degree, perhaps represented by a real number. If Leona's attitude toward the job is graded, and graded intentions are a form of acceptance, then her case meets the acceptance requirement, after all. But does Leona really have graded intentions, and should we count them as acceptance?

When one hears 'graded attitudes', a certain sort of gradient comes to mind: *precise* degrees of an attitude's *strength*. I doubt that human beings really have intentions (desires, etc.)

with precise degrees of strength.¹⁵ But let us grant that such attitudes are real; still they don't fit Leona's case. What's distinctive about Leona isn't that her attitudes are *weak*. A weak intention is one that can be easily overturned—e.g. I intend to drink the light roast for my afternoon coffee, though I would give up the intention if the dark roast went down in price by \$.50, or if someone offered a free mug of tea, or if a gentle breeze invited me to go fly kites all afternoon. My intention to get the light roast is modally fragile, and there is little I would sacrifice to secure its object. But there is nothing tentative or provisional going on here. I am not just considering—I am weakly but *decidedly* pro-light roast.

Leona's stance toward the job, however, is indecisive. She is still mulling it over: thus she will only accept it as her end after some 'warming up'.¹⁶ What's more, Leona's attitude is not easily overturned or modally fragile. She might be willing to sacrifice quite a lot to keep her options open (suppose that she loathes being interviewed—and yet she goes). It might take a lot more than a gentle breeze or a faint sweetening of her other options to discourage her attendance. So her tentative attitude toward the job is very unlike my weak intention to get the light roast.

Although it is true that Leona's attitude comes in degrees, they are not degrees of strength, but *provisionality*. For this reason, Leona's case is still anathema to the acceptance requirement: provisional acceptance is not a kind of acceptance. (Fake barns aren't a kind of

¹⁵ Two potential objections (adapted from Holton's [2008] case against the psychological reality of credences): (i) graded intentions don't show up in ordinary thought; and (ii) they are not easily squared with all-out intentions, which do.

¹⁶ If my intention to get the coffee strengthens, one would not say that I thereby warm up to anything. To warm up to x is to feel *more determinately* (not *more strongly*) pro- x . My thanks to an anonymous referee for convincing me that my earlier thoughts on 'warming up' were half-baked.

barn.) A further question is whether intentions have *precise* (as opposed to imprecise, or indeterminate) degrees of provisionality, which could be modeled by assigning each attitude a real number.¹⁷ A precise model might assign 1 to any perfectly non-provisional pro-attitude, and assign less provisional attitudes a number closer to 0. (An imprecise model would instead assign each attitude a set of reals.)

Are our attitudes really so fine-grained? Could anything but high-octane sci-fi ensure that Leona's intention (or desire) to get the job has provisionality .049999998, rather than .049999999? I doubt it. But really the question of precision is irrelevant to the case against the acceptance requirement. What is relevant is that (i) Leona's attitude is best understood as provisional to some degree; (ii) degrees of provisionality are not simply degrees of strength; and (iii) whereas weak intentions and desires count as acceptance, provisional ones do not.

Lots of questions remain about what provisionality is and why it should matter. Let me mention just two. First, could we understand provisional desires as indeterminate in polarity? This might involve representing their strength by a partly negative (anti-) and partly positive (pro-) interval, like $[-.25, .25]$. This would reduce provisionality to strength, but provisional attitudes would still count as wishy-washy rather than as forms of acceptance.

Second, is there any *normative* difference between degrees of provisionality and degrees of ignorance? No doubt there is a psychological difference. Contrast (i) acting on a .7 provisional desire for soup, as in the familiar soup case, with (ii) acting on a .7 credence that one has a non-provisional soup-desire, as in a case where my desires are opaque to me, until a

¹⁷ For some recent work regarding imprecise doxastic states, see [Bradley 2015; Schoenfield 2015].

trusty guru declares it .7 likely that I could go for a tomato bisque. But I suspect that these sorts of cases are normatively similar in some key ways. Just as agents must discount risky prospects (relative to sure things), they must care less about satisfying their wishy-washy desires (relative to settled ones). And in both cases, I suspect, agents have at least some leeway in determining how steep the discount is to be (for discussion of risk-discounts, see [Buchak 2014]).

There is, however, something more authentic about acting on one's felt provisional desires rather than a guru's best guess. Wishy-washy as my soup-desire may be, at least I know it's how I feel. (I'm not just deferring to *their* testimony to work out *my* goals.) This suggests that considering and other provisionally pro-attitudes may have a role to play in theories of transformative experience, where authenticity looms large [Dasgupta msc.; Paul 2014]. Warming up to an end, after all, is something that *you* must do. It is not as though you just consider things awhile then find yourself with new desires *ex nihilo*. Your desiring emerges from your considering, and there is a great difference between attitudes arrived at in this way and those due to neural glitches, unavoidable deference, or irresistible conditioning.

8. Thinking, Acting, Considering

The acceptance requirement, I have argued, is false. What does that mean for its proponents?

Aristotelians can probably relinquish the requirement without giving up their core beliefs. Consider Brewer's [2009: 27] reason for endorsing the requirement:

Our task is to understand how [a certain behavior] counts as action expressive of a conception of the good rather than as mere goal-directed behavior of the sort found throughout the animal and plant kingdoms, and we can only do this if we gain some glimmer of the conception of the good that inspires these actions.

Brewer's task is to set agency apart from other telic processes. The key difference—and his core belief—is that agents are 'inspired' by their 'conception of the good'. But why should inspiration require acceptance? An agent might see an end as *potentially* good, and thereby feel inspired to act, without compromising her status as an agent. Leona is no beanstalk growing toward the light. She is a conflicted actor inspired by, yet suspicious of, the prospect of a certain way of life. Even though Brewer denies that such an agent is possible, the possibility sits well with his conception of desire. For Brewer, desires are 'appearances' of value 'in the space of practical thinking', though they sometimes present only 'an *inchoate* sense of how or why it would count as good to act as they incline us to act' (Brewer [2009: 29], emphasis added). To this, Brewer should add that *considering*, too, involves an inchoate sense of value.

What about Humeans like Davidson? On their view, actions are movements of the body caused by a final desire for a state of affairs (alongside a belief that the movement can bring about the state). But Leona doesn't finally desire to keep her options open; she wants options only because she's considering taking the job. Humeans have two ways to make room for this. First, they might relax the requirement that every act be caused by a *final* desire, and allow that Leona's desire to keep options open is merely instrumental to the job she is considering taking. Second, they might relax the requirement that every act be caused by a final *desire*, and instead say that *considering* (plus a belief) can be a cause of action. I have quibbles about both proposals, but neither is disastrous or deeply anti-Humean.

For Kantians like Korsgaard, the acceptance requirement may be harder to let go. If we can act for the sake of unaccepted ends, then we can act without universally willing a means-end pair (even if the universality is merely 'provisional'; see fn. 4). This contradicts the

conclusion of Korsgaard's [2009: 72–6] argument against 'particularistic willing', which aims to show that 'universal willing' in accordance with the categorical imperative is constitutive of action—a major plank in the argument of *Self-Constitution*. I am not sure how Korsgaard would respond to this problem. She might soften the acceptance requirement, along with the conclusion of the argument against particularistic willing. It will be hard to do this without giving up on the core idea that good action involves universal willing, but I am not entirely pessimistic. We just need to find an intermediate state: Leona is not passively *pushed* by an incentive—that is, 'a motivationally loaded representation' presenting its object as 'desirable or aversive in some specific way' [Korsgaard 2009: 109]. But neither is she actively *legislating* an end for all rational beings as such (in situations like hers). The challenge is to find a way to count Leona more as active than the ragdoll subject to incentives, but less domineering than the rational legislator.¹⁸

As you might have noticed, the basic challenge here is not unique to Korsgaard. Davidson must find a mental state between desire and indifference; Brewer must show how agency could draw on incipient conceptions of the good. We seem to be dealing with three aspects of a general, non-partisan problem in the theory of action—but what?

The problem has to do with a tempting yet too-simple model of deliberate agency: the Two-Phase View. In short, the view is that thoughtful action is thinking, *then* acting. Episodes of agency foliate into two discrete phases. First we deliberate, chewing over various ends and means; then, after accepting an end, we execute our deliberation by taking means toward the

¹⁸ One clue might lie in recent work on inclination (e.g. [Schapiro 2009]).

ends we have accepted. The purpose of phase one is to arrive at acceptable ends; phase two's purpose is to realize one's ends, thereby making an honest agent of oneself.

The Two-Phase View, for all its elegance, is obviously tailored to an ideal scenario: the Practical Hero thinks awhile (deliberation/practical reasoning), then carries out a resolute plan (execution/action). But sometimes we intentionally realize ends without prior deliberation—phase one never happens, though it's as if it did. Other times we act intentionally with no purpose in mind at all, as when I idly drum my fingers—phase one never happens, and there is not even an end in place as if it had. These are well-known sorts of non-ideal cases. We basically understand them, and we rightly don't let them scare us away from the Two-Phase account of simpler cases.¹⁹

Cases like Leona's present a new kind of non-ideality: phase two begins *while phase one is still underway*. Ideally, the two phases are sealed off by the moment of acceptance: an end is adopted after phase one, and phase two starts only thereafter. But Leona's predicament shows that this isn't always true—that the acceptance requirement is false. The two phases can't be as discrete as the Two-Phase View would have it; their relationship is more fluid and reciprocal.

So here is our problem. Acceptance guides our actions because it represents the conclusion of our thoughts. But what kind of state could guide an action whose basic point has yet to be settled? (If the answer is 'considering', then what exactly *is* considering?) This question

¹⁹ Another ideal: an action's value is transparent to its agent and specific to one episode. Brewer [2009: 39] shows that this idealization is not always met: certain actions ('dialectical activities') have only 'opaque' value, which may 'unveil itself incrementally' over the course of many episodes (see also [Arpaly 2003; Brewer 2009: 73–5, 96]). (One of Brewer's examples is doing philosophy.) My point is different, since 'opaque' values are still non-provisional. The phenomenology of unveiling is thus unlike that of warming up. In warming up to an end, your mind undergoes a change from considering to accepting; when your end is 'unveiled', you see that *this* was what you wanted *all along*.

isn't exclusive to any one tradition in action theory. The answer could come from anywhere.

REFERENCES

- Anscombe, G.E.M. 1963. *Intention*, second edition. Oxford: Blackwell.
- Arpaly, Nomy 2003. *Unprincipled Virtue*. New York: Oxford University Press.
- Arpaly, Nomy and Schroeder, Timothy 2014. *In Praise of Desire*. Oxford: Oxford University Press.
- Bradley, Seamus 2015. Imprecise Probabilities, *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2015/entries/imprecise-probabilities/>.
- Bratman, Michael 1987. *Intention, Plans, and Practical Reason*. Center for the Study of Language and Information.
- — — — 1992. Practical Reasoning and Acceptance in a Context, *Mind* 101/401: 1–16.
- Brewer, Talbot 2009. *The Retrieval of Ethics*. Oxford: Oxford University Press.
- Kolodny, Niko and Brunero, John 2013. Instrumental Rationality, *Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/win2016/entries/rationality-instrumental/>.
- Buchak, Lara 2014. *Risk and Rationality*. Oxford: Oxford University Press.
- Cohen, L. Jonathan 1992. *An Essay on Belief and Acceptance*. New York: Clarendon Press.
- Collins, Arthur 1984. Action, Causality, and Teleological Explanation, *Midwest Studies in Philosophy* 9/1: 345–369.
- Dancy, Jonathan 2000. *Practical Reality*. Oxford: Oxford University Press.
- Dasgupta, Shamik msc. How to be Authentic.
- Davidson, Donald 1963. Actions, Reasons and Causes, *Journal of Philosophy* 60/23: 685–700.
- Frankfurt, Harry G. (1978). The Problem of Action, *American Philosophical Quarterly* 15/2: 157–162.
- Holton, Richard 2008. Partial Belief, Partial Intention, *Mind* 117/465: 27–58.

- Kant, Immanuel 1785 (1998). *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. 1983. Two Distinctions in Goodness, *Philosophical Review* 92/2: 169–195.
- — — — 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Müller, Anselm 2011. Backward-looking Rationality and the Unity of Practical Reason, in *Essays on Anscombe's Intention*, ed. Anton Ford, Jennifer Hornsby, and Frederick Stoutland. Cambridge: Harvard University Press.
- Muñoz, Daniel msc. *Responsive Reasons*.
- Paul, L.A. 2014. *Transformative Experience*. Oxford: Oxford University Press.
- Raz, Joseph 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.
- Schapiro, Tamar 2009. The Nature of Inclination, *Ethics* 119/2: 229–256.
- Schoenfield, Miriam 2015. The Accuracy and Rationality of Imprecise Credences, *Noûs* 50/4: 1–19.
- Stocker, Michael 1979. Desiring the Bad: An Essay in Moral Psychology, *Journal of Philosophy* 76/12: 738–753.
- Smith, Michael 1987. The Humean Theory of Motivation, *Mind* 96/381: 36–61.
- Thompson, Michael 2008. *Life and Action*. Oxford: Oxford University Press.
- Ullman-Margalit, Edna 1983. On Presumption, *Journal of Philosophy* 80/3: 143–163.
- Vogler, Candace A. 2002. *Reasonably Vicious*. Cambridge: Harvard University Press.

Massachusetts Institute of Technology