

Against Reductive Normative Realism (Excerpt from my 3<sup>rd</sup> Dissertation Chapter)

Stan Husi

This chapter surveys naturalistic or reductive normative realism. The view – conveniently named *naturalism* within the chapter’s confines – holds normativity to be a natural phenomenon. I shall discuss various natural constellations – normative engines – that have been proposed to explain normativity, and show why they must fall short in that aspiration. Desire and its relation to reason will be the most important element in this discussion, since it fuels the most powerful normative engine and provides the blueprint for most others as well.

In a nutshell, my worry recasts the familiar one: that there is a conceptual gap no substantive form of naturalism can bridge: from what is to what ought to be, from what is the response we are inclined to take to the one we should take, from what we cherish to what is appropriate to cherish, from what we praise and condemn to what is praiseworthy and worthy of condemnation, from what standards and norms we endorse and accept to which standards and norms we should endorse and accept, and so forth. There’s no natural route to get us from here to the Promised Land of shoulds and oughts; no normative engine is capable of delivering the final product of reasons. All we get are communities who find certain ways of living more agreeable than others, who, through a long process of trial and error, have mostly come to agree to abide by certain norms in order to get along. We may have internalized a considerable number of norms, we may get upset at each other for violating them, we may communicate and debate our shared attitudes, we may decide to apply sanctions to the callous and indifferent, and we may resolve to strive for further reform. But where, then, is that fact, that natural configuration, which explains

what ways of live are appropriately chosen as better than others, what norms we should agree to abide by, which norms we must internalize, when we are right and justified in getting upset at others, when it is truly called for to apply sanctions, and what further reform constitutes progress rather than just change?

The problem is not finding standards of correctness; that's easy, so easy, in fact, that we can immediately understand why many want more. Standards of correctness are provided by norms, viz. principles, policies, and prescriptive rules all of which encapsulate a directive element. The issue concerns the status of norms: that there are any privileged norms which possess objective authority. To fully appreciate the issue, we must keep distinct the formal aspect of being directive in character from the substantive one of directing with authority. The first distinguishes norms as norm, but it is the second we need to focus on. Philosophers harboring doubts about the normativity of morality, law, rationality, instrumental reason don't contest the presence of norms in these areas but their authority. Norms of etiquette still serve to illustrate the distinction best. The rule that one must answer in the third person to third person invitations clearly states a demand, but what is its normative force? (Foot, *Philosophical Review* 81, 1972: 308). Norms are easy to come by; their authority, their status as objective guidelines, however, must appear an altogether more consequential matter. Compare Korsgaard who writes "We live under the pressure of vast assortments of laws, duties, obligations, expectations, demands, and rules, all telling us what to do. Some of these demands are no doubt illicit ... just social pressure, as we say. ... I call the *normativity* of a law or a demand ... the grounds of its authority ... the way it binds you." (*Self-Constitution*, Oxford UP, 2009:2) The fundamental question, then, is whether out of the great plurality and diversity of norms any stand out as objectively authoritative. Naturalists say yes and then aspire to explain why.

## 1. Normative Naturalism

What is natural normative realism? Naturalism obviously affirms the existence of some normative reality. What renders the view distinctive is of what kind naturalism pronounces that reality to be. The answer, of course, is of the natural kind: since naturalism holds everything to be natural and normativity to be, normativity must be natural. Yet we need to know more precisely what that means. For this we can substantially draw on the previous chapter. There, in reconstructing non-reductive realism, we saw that for naturalism to emerge as a distinct position it does not suffice to acknowledge that all that is normatively relevant must be built up from natural bits and pieces in entirety. What we are after is normative status rather than that which bears normative status: what it is to have this or that status rather than what does have this or that status. Neither does it suffice for naturalism to emerge as a distinct position that normative status supervenes on the natural. Naturalism, in a nutshell, offers a view distinct from non-reductive realism only if it postulates that, in addition to *that whatever is normatively significant* to be natural configurations, *that it is normatively significant* also figures as a thoroughgoing natural phenomenon. Let's briefly recall why.

Remember, first, the important distinction between normative status and that which has such and such normative status. The previous chapter put great emphasis on the fact that non-naturalists need not deny that everything that has some particular normative status is itself a natural configuration. Naturalists consequently cannot distinguish themselves in virtue of holding solely that the normatively significant fully consists in certain segments of the natural unless they also hold that normative significance itself fully consists in certain segments of the natural. If the good life consists in happiness and health, profitable pursuits and engagement, close relationships and beneficial sociability, insight and understanding, then nothing in that list

need give pause to any firm naturalistic metaphysician. The suggestion that there are close relationships, for instance, hardly raises any special ontological concerns. What potentially may give pause to such a metaphysician, though – even with the aforementioned admission – is the idea that there is such a thing as *the good life* – a life particularly worthy of choice, a life we have reasons to strive for, admire, or envy: A proposition entirely distinct from one recognizing that there are lives we do happen to choose, strive for, admire, or envy. In other words, it is consistent to grant that there are e.g. close relationships without also granting that they bear the particular normative status of exemplifying value. The distinction between normative status and that which has it could be taken to separate classical normative ethics from meta-ethics; the former primarily being concerned with what has what normative status, the latter primarily being concerned with what it is to have normative status in the first place.

Consider supervenience next, the thesis that, necessarily, two situations differ in their normative status only if they differ in their natural configuration as well. Though widely considered the hallmark of naturalism, this thesis clearly cannot set apart naturalism from non-naturalism. This is readily apparent once we recall that even dualists such as Descartes were able to firmly and consistently pronounce the supervenience of the mental on the natural without thereby retracting their commitment to non-naturalism. Likewise, there is nothing incoherent in supposing the normative to be firmly non-natural *and* to supervene on the natural. This could be achieved in virtue of God's maintenance of a pre-established harmony between the normative and the natural; or it could just be like that, without explanation, but why suppose everything must necessarily receive an explanation? Perhaps our universe contains normative laws as primitive components just it contains basic physical laws as primitive components. Normative laws in that context would be laws that systematically correlate natural configurations with

normative status and which would thereby, by fiat, secure the latter to supervene on the former. Leibniz raised the puzzling question of why there is something rather than nothing – why there are laws to govern our universe – which is puzzling precisely because we are at a loss as how to answer it. Rather, it appears there simply is no answer as to why fundamental laws of nature obtain. Likewise, it could simply be the case that there are systematic normative-natural laws where again there simply is no illuminating answer as to why they obtain. That this would spell systematic frustration of our ambition to come to grips with the normative delivers no proof that there is no non-natural normative reality supervening on the natural.

## 2. Desire, Pure and Simple

Desire fuels the most powerful normative engine. Since it provides the schema and blueprint for many others to come – e.g. the endorsement of norms and the agreement to install and abide by norms – we shall begin with desire. Understand desire to be that pro-attitude which maximizes the plausibility of the relevant response-dependent theory of normativity. Desire, for instance, need not have any phenomenological imprint and usually is not tied up with sexuality. Later I shall argue that the kinds of pro-attitudes that stand the best chance of carrying reasons are concerns embedded within our larger identities and projects. For now, the basic idea is that with the appropriate provision of desire comes reason and value; that practical reasons enter the stage with agents who harbor concerns and adopt ends. The universe is cold and indifferent, but we are not, and in light of our ends and what we care about some practical pathways emerge as more attractive than others. In the final analysis, it is supposed, reasons and values rest in desire, not the other way around. Spinoza wrote “It is clear ... that we do not endeavor, will, seek after

or desire because we judge a thing to be good. On the contrary, we judge a thing to be good because we endeavor, will, seek after and desire it.” (Spinoza, Ethics III.9 Scholium) And Simon Blackburn adds “Nature itself may be heartless and free of desire, but among the creatures it has thrown up are some which are not heartless, and not free of desires. We understand our values by understanding ourselves as valuing, and this we can do.” (*Ruling Passions*, Oxford UP 1998:50). I share a great deal of sympathy with what Spinoza and Blackburn have to say. In fact, reading their statements carefully, I fully agree. Spinoza talks about *judging a thing to be good* and not *a thing being good*; and Blackburn talks about *understanding our values* and not about just *values*. I too believe desire fuels our perception of and response to things as good, worthwhile, and reasonable. It is what anchors and sustains our practices. If that was all we seek to understand, I am on board.

But it is not. Usually desire is called upon to explain reason and value in their distinctive normative sense; not just our perception and thought about reason and value, but reason and value, period. And here I must depart. I find the desire-based account of normativity a view astonishingly close for being entirely off target. Normativity cannot be explained on the basis of desire. Desire-based accounts deliver something, and we may choose to call it *reason* and *value*. Given their wide employment to denote most everything, we may choose to allow those terms yet another usage. But labeling should not deceive us that the thing named entirely lacks anything distinctively normative. What we need is to look inside the engine and to see what it actually does. We shall see then what it explains is why we choose what we choose and why we prize what we prize. What we shall miss, however, is why we have reasons to choose what we choose and why we have reason to prize what we prize, or why doing anything of that sort can ever be appropriate or inappropriate.

There are desire-based accounts of reasons and there are desire-based accounts of value. Since both proceed in different manner they need be treated separately. I shall here focus on reasons. First, recall what we are after. What we seek is an explanation of how states of affairs can stand in the reasons-relationship to choices agents face in particular circumstances; how certain states of affairs can favor certain responses. The relationship itself needs explaining and not just what happens to stand in that relationship: which is just plain situations, agents, and choices. What favors what hardly gives rise to any profound puzzlement. *That* something favors what it favors, in contrast, is what it is hard to understand yet what desire-based accounts need to explain. Sporadically one finds authors wondering what the fuzz about reasons is all about. Reasons are just plain state of affairs and situations, they say, and what's problematic with that? Nothing, of course. Yet situations are just situations and are reasons for anything only in virtue of standing in the favoring relationship to whatever they favor. How, then, do desire-based accounts go about explaining reasons and the corresponding favoring-relationship?

Jonathan Dancy proposes an excellent précis of desire-based accounts of reason, where  $p$  is some state of affairs,  $A$  some agent,  $\phi$  some practical option or choice, and  $e$  some end or concern (*Practical Reality*, Oxford UP 2000:28):

If its being the case that  $p$  is a good reason for  $A$  to  $\phi$ , *this is because* there is some  $e$  such that  $A$  actually desires  $e$  and, given that  $p$ ,  $\phi$ -ing subserves the prospect of  $e$ 's being realized (or continues to be realized).

Dancy is no fan of desire-based accounts, but his characterization is remarkably close to Mark Schroeder's, who is a fan. Suppose  $r$  stands for some proposition (or state of affairs),  $x$  for some agent,  $a$  for some action (or option or choice), and  $p$  for some desire, end, or concern.

Schroeder writes (*Slaves of the Passions*, Oxford UP 2007:29):

For all propositions  $r$ , agents  $x$ , and action  $a$ , if  $r$  is a reason for  $x$  to do  $a$ , that is because there is some  $p$  such that  $x$  has a desire whose object is  $p$ , and the truth of  $r$  is part of what explains why  $x$ 's doing  $a$  promotes  $p$ .

That Dancy talks about state of affairs (or what is being the case) and Schroeder about propositions need not distract us here. These are variations in detail we can safely ignore for present purposes. The important upshot is that according to both accounts, reasons are complex quadruple relationships between (1) state of affairs, propositions or situations, (2) agents, (3) actions, options, or choices and (4) desired ends or concerns. Or, to put it differently, for something to be a reason is for it to stand in this complex quadruple relationship. Schroeder provides a neat example (2007:1):

Tonight there is going to be a party, and everyone is invited. There will be good food, drinks, friends, chat, music – and dancing. Ronnie and Bradley ... have been invited to the party. But while Ronnie loves to dance, Bradley can't stand it. ... So while the fact that there will be dancing at the party is a reason for Ronnie to go, it is not a reason for Bradley to go. ... Ronnie's and Bradley's reasons therefore differ. ... Moreover, it's not hard to see why Ronnie's and Bradley's reasons differ ... It is because of what they *like*, *care* about, or *want*.

Ronnie's having a reason thus consists in the quadruple relationship between a situation (the party), an agent (Ronnie), an end (love of dancing), and a action or choice (going to the party) such that the situation causes a particular choice to be the (best) option for an agent to bring about what he cares about.

Notice that in the present version it is not desires themselves that *are* reasons. Instead, desires figure as components and constitutive parts of the relevant reasons-relationship, though certainly important components and parts. Reasons surely are complex. Where there are reasons, according to the present account, there are situations and agents and ends and choices appropriately interrelated. If we wish to single out one of the above components as *the reason*, that should be states of affairs and situations. This choice would comport with common usage where we often call the relevant situation itself *the reason*: as when we say the dancing at the

party was the reason for Ronnie to go. This way of speaking is perfectly fine as long as we do remember that for reasons to be reasons – for state of affairs to be reasons – they need to stand in that particular favoring relationship. And to favor, according to the account in question, is to stand in a quadruple relationship that only features desires as one important component.

Similarly, we best avoid speaking of desires as “generating” or “providing” reasons. This locution only suggests a rather misleading temporal picture involving two stages, where first there are desires and then thereafter, in some mysterious manner, they “generate” or “provide” reasons. Desire-based accounts essentially are accounts of what reasons and the relevant reasons-relationship are: take them apart, and this is what you find: a quadruple relationship connecting situations, agents, ends and choices, and certainly not a process of dubious creation. Schroeder, in his general analytical approach to understanding reasons, compares his account to the analysis of water as H<sub>2</sub>O. There as well it would be rather awkward an expression to say that the chemical complex H<sub>2</sub>O *generates* or *provides* water instead of just being water. If again we wish to single out one component as the provider of reasons, it would be as before state of affairs and situations. It is the dance-party which provides the reason for Ronnie to go. (Likewise, it is hoses and faucets which provides water and not H<sub>2</sub>O).

These are not merely terminological reminders but have larger ramifications. Reasons are world-wise and practical deliberation consequently is outward-directed rather than inward-directed. Deliberation engages the world and rarely amounts to a mere exercise in navel-gazing. Desire-based accounts do not face principal obstacles in acknowledging these fundamental facts so long as they stay clear on how precisely desire interacts with the other components and, most importantly, the world. What recommends certain choices instead of others first and foremost has to do with how things are. It’s the dancing at the party that makes all the difference for why

Ronnie should go. This is not to diminish the important role of desire, for without it no situation by itself does ever recommend anything. Yet as important the role desire plays, the role is better seen as that of a background enabler, which is that of a spotlight operator rather than that of a frontal stage actor. We look at our options through our desires instead of paying them direct attention. It is in light of one's desires that certain practical options become more attractive than others, where importantly the light springs from desires instead of shining on desires, and where what is illuminated in particularly favorable colors are choices rather than desires.

Only so conceived do desire-based accounts stand any chance of successfully capturing the deliberative perspective of agents. Agents are the unique consumers of reasons. To them reasons must be addressed and they must determine whether to take seriously what is so addressed. That's why it is such an important test case for accounts of reasons that they not misconstrue the perspective of us qua deliberators, the chief arbiters of reasons. Desire-based accounts would fail in this regard if they placed the primary focal point of reasons on desire. Fortunately, they do not. To desire is to take a stake in things, but what we take a stake in are things and not desires. Deliberation very rarely concerns what to desire rather than what to do given how things are. Here's an example. Desiring to have a piece of chocolate-cake is to look at that piece in a certain manner: finding it attractive, taking pleasure in the thought of eating it, and then figuring out how to get it. It is not that one has a neutral perception of the cake, and also notices a desire somewhere to eat it, on the basis of which one then draws the practical inference that one shall eat it. If this was the way desires came upon us, it would leave it rather obscure why we should care about them at all. Suppose I was to discover my desire for chocolate-cake, say by locating it on the computer-screen of some MRI device I am presently plugged into or by some process of psychoanalysis. That result may strike me as quite interesting; as yet another

curiosity about myself. But that fact certainly is not what drives me to get that cake; what does is my fondness of it – rather than my notice of that fondness – with the characteristic accompanying outward way of looking at the cake.

All in all, I share a great deal of sympathy with desire-based accounts. For convenience, we may call them *Schroeder-reasons* or just *S-reasons*. They comport with a substantial proportion of the term's common usage, as is nicely illustrated by Schroeder's example of Ronnie and Bradley. Furthermore, they capture situations of agents in a manner that reflects their own appraisal of these situations; what they find salient and significant about them and what choices will consequently appeal to them. By keeping close to agents' own deliberative vantage points S-reasons afford vast predictive powers. They are tied up with the intentional stance that unlike others is actually able to issue reliable long-term forecasts of what people are going to do. In this regard S-reasons' predictive utility may presently be unsurpassed. This is a pleasant benefit when we occupy a third-person perspective on other people. But desire-based accounts are in fact most attractive when we occupy a first-person perspective on our own deliberative situations. S-reasons capture what we wish our advisors to be concerned about, since we, the advisees, shall be able to recognize the recommendations they issue as truly helpful and non-presumptuous. This is not to reiterate the important point from above that capturing the first person deliberative perspective is paramount for any account of reasons. Rather, it brings us to the perhaps most important source of motivation for desire-based accounts: that they, unlike most competitors, refrain from imposing reason-demands on agents that incorporate aims they have not chosen as their own; that they refrain from hoodwinking agents into compliance with alien agendas in virtue of utilizing what could be seen as nothing but a particularly sophisticated form of bluff.

The motivation in question springs from the famous *internalism requirement*. The requirement places a necessary condition on practical reasons: whether some consideration qualifies as a reason for an agent depends upon its capacity to engage what that agent cares about. Bernard Williams, utilizing semantic ascent, offers this definition in terms of statements about reasons: The statement that “*there is a reason for A to  $\phi$  ... implies ... that A has some motive which will be served or furthered by his  $\phi$ -ing, and if this turns out not to be so the sentence is false.*” (*Internal and External Reasons*, in *Moral Luck*, Cambridge UP 1981:101) Williams calls reasons that meet the requirement *internal* and those that do not *external*, but since he does not believe in external reasons, what his definition effectively does is to place a constraint on practical reasons as such. His dichotomy between internal and external reasons does not offer a classification of two kinds of reasons so much as in effect to rule out *categorical* or *external reasons*, reasons agents allegedly have irrespective of what they happen to care about. What Williams should have said is that practical reasons are internal or cease to be reasons. The requirement proposes a litmus test each reason-candidate must pass. The requirement has received extensive treatment in the literature. Here I shall be rather brief and not defend a stance for or against; all I shall provide is a rough rationale for the requirement. The employment of the requirement within my larger argument does not rest on its validity so much as on the assumption that it figures as a vital element for desire-based accounts.

The core intuition behind the internalism requirement is that agents must be capable of seeing a point in taking seriously the demands that are applied to them. The internalism requirement incorporates a profound skepticism with regards to all forms of substantive demands and standards that are forced upon agents without their approval: dictates that they must do something even though they cannot quite see why. James Dreier captures the spirit nicely:

When we give a justification, we are either explicitly citing or adverting to some norms. But we can't just cite any old bunch of norms. Which norms count toward justification? The problem is that if we simply cite a bunch of rules, the agent may well ask, what are those rules to me? She may ask for a reason to follow them. And we can't just shrug this off. Suppose that someone cited the laws of India in support of moral principles. We ourselves recognize that this sort of justification is useless. ... If we cite the laws of India and our subject asks what reasons she has to follow them, we understand what she's asking. She's again asking for reasons. She doesn't see any force in the rules we've cited. (*Humean Doubts about Categorical Imperatives*, in: *Varieties of Practical Reasoning*, ed. Elijah Millgram, MIT Press 2001:41):

It comes all down to reasons, then. And reasons, the requirement pronounces, cannot be imposed on agents from the outside as it were lest they lose all force and point. If someone, after careful reflection, were to remain entirely unmoved by some consideration, its capacity to figure as a reason is seen as systematically undermined. The question of why accept a consideration as a practical reason is always a fair one to ask. And the buck stops with agents on whom it is to answer this question. Their indifference towards a consideration, says the requirement, must mean the denunciation of it as a reason by the only arbiters and judges there are. Considerations that have aspirations to reasonhood must thus find a foothold in the concern-structure of the relevant agents to whom they are proposed, for either they will connect to what they care about or they simply will get disconnected. If considerations fail to *engage* agents – by incorporating aims that are considered *alien* by those agents – the requirement strikes out these considerations as mere reason-pretenders falling short of qualifying as true practical reasons. Richard Joyce writes (*The Myth of Morality*, Cambridge UP 2001:80-81):

Normative reasons claims – claims concerning what it is rational for an agent to do – must be something that potentially engage the agent to whom they are applied. This doesn't mean that the presentation of a true normative reason claim immediately results in the agent being motivated; rather, it means that the agent cannot sensibly both acknowledge that something is a normative

reason for him to act and ask “But so what?” Any adequate theory of normative reasons must make out reasons to be precisely those things that forestall a “So what?” response.

How does the internalism requirement lend support to desire-based accounts? The requirement certainly does not imply the truth of desire-based accounts. Various authors, in fact, consistently endorse the requirement and yet reject desire-based accounts. The former places a necessary condition on practical reasons, whereas the latter provides a sufficient condition for reasons. The former constrains what qualifies as reasons, without necessarily saying what reasons are or to offer an account. The latter, in contrast, does say what reasons are and does offer an account. Still, how the requirement is able to lend crucial support to desire-based accounts is not hard to see. According to the requirement, principles retain normative force only if its subjects, the relevant agents, are prepared to accept them. And accepting a principle is to exhibit some willingness to comply, which requires a motive, a desire to do something. David Velleman offers a similar reconstruction (*The Possibility of Practical Reason*, Oxford UP 2000):

Suppose that reasons for someone to do something must be considerations that would sway him toward doing it if he entertained them rationally. And suppose that the only considerations capable of swaying someone toward an action are those which represent it as a way of attaining something he wants ... These assumptions, taken together, seem to imply that the only considerations that can qualify as reasons for someone to act are considerations appealing to his antecedent inclinations – that is his desires or dispositions to desire. (170-171)

The internalism requirement demands reason must have the capacity of engaging the concerns agents have, and desire-based accounts take care of this by precisely making reasons a function of those concerns. Desire-based accounts, hence, explain reasons in a manner that fits exactly the job description placed by the requirement. Desire-based accounts find support in the requirement because they offer the best explanation of the requirement itself. The relevant logic here is abduction, not deduction, that is to say inference to the best explanation. It is unclear

whether those who accept the requirement and yet reject desire-based accounts can do equally well. Their alternative way of accommodating the requirement smacks of being ad hoc. Suppose that alternatively, reasons were grounded in concern-independent considerations, such as objective values. Dancy, among others, has argued for this, while also holding on to the requirement. The question he and his compatriots face is why, given that we are now supposing reasons spring from value rather than concerns, some value here is capable of empowering reasons while the exact same value there is not, solely in virtue of extrinsic and unrelated facts concerning what certain agents happen to care about and what they happen to be motivated by. The question, in other words, is how to explain the relevance of such concern/motivation-related contingencies for reasons given that what reasons are, according to this alternative picture, has nothing to do with those concern/motivation-related contingencies. It does not suffice to insist that reasons must be capable of motivating agents. The question is precisely why this should be so if the source of reasons had nothing to do with motivation altogether. To date I am not aware of any satisfactory answer to this question.

The requirement, then, lends vital support to desire-based accounts. Yet the role it plays for desire-based accounts actually goes further. The requirement does not just figure as one witness among others to speak on their behalf. Without the requirement desire-based accounts are exposed defenseless with regards to the myriad of competitors who seek to empower features other than desire with reasons-related capacities. This can most dramatically be illustrated with regards to desire itself, only desires of others. According to desire-based accounts, it is not just desires as such that figure as crucial components for the reasons agents have; it is only agents' *own* desires. This is structurally somewhat puzzling. Why should only desires agents have themselves be capable of figuring in the relevant reasons-relationship and not also desires held

by their agential neighbors? After all, desires are desires, and if two desires had in common everything but their location, it is not immediately obvious why their geography should make all the difference. It sounds rather awkward to say that this desire on your right can undergird someone's reasons while its identical twin desire on your left cannot. Consider an example. Suppose Bert and Berta share a room that Berta wants to be warmer but Bert does not; Bert is in fact indifferent. Berta and Bert sit side by side with the thermometer in equal reach. Now what desire-based accounts imply is that Berta has a reason to get up from her seat and turn up the heat and that Bert has no such reason; disregarding, for them moment, any concerns Bert may have for Berta and that he also lacks any other social/moral sensitivities; imagine Bert and Berta suffer from systematic mind-blindness, a particularly heavy form of autism. How come, then, that whether the exact same desire can figure in Bert's reasons depends on its location? The answer must be that this is so because only if the desire is Bert's own do we have any guarantee that Bert will be moved by it and see a point in doing something that furthers it. In other words, because out of two otherwise identical desires the internalism requirement preempts the motivationally disconnected one to count at all.

This completes my reconstruction of desire-based accounts and of what they have going for them. Unfortunately desire-based accounts render reasons entirely void of distinctive normativity. Once we attend carefully to the details of desire-based accounts, I believe this becomes next to obvious. We must not fall prey to an awkward reverse consciousness suspecting what is readily apparent really must be deceptive instead. The crucial point is the relevant relationship between situations, agents, ends and options is blatantly non-normative or rendered thus only in virtue of externally normative principles that transcend the compass of desire-based accounts. Let me explain.

Take situation S, agent A, end E as well as options X and Y. Now suppose that as a matter of empirical fact, A will more likely realize E if he were to X rather than Y, and everything else is truly equal. Given how A's overall cognitive-motivational system is set up, and given that he really cares about E and is not stupid, we would expect A to prefer X over Y. We would certainly expect A to gravitate towards X rather than Y and be puzzled if otherwise. The one-million-dollar question, however, is this: can we detect anything distinctively normative in that complex relation that explains why A *ought* to do X, *should* do X, *must* choose X over Y, would become *subject to warranted criticism* if he did not in fact choose X over Y, and so on? Can we detect anything that explains, based on that relationship, why preferring X over Y is genuinely the *appropriate* choice and why preferring Y over X is genuinely the *inappropriate* choice?

We cannot. What we can detect is this: if A chooses X over Y, he will more likely realize his end E. If A chooses Y over X, he will less likely realize his end E. This appears to be a simple empirical observation of the kind that psychologists, sociologists, or anthropologists might advance with regards to some study subject. It fits the perspective of some neutral observant of the strategic situation of two combatant parties: If Napoleon moves there, he will risk his right flank. If he retreats, he will compromise the benefit of the victory the day before. Granted, we may employ reasons-language to capture this and other empirical observation. This may be especially illuminating since it reveals how the relevant agents appraise the situation. It would be perfectly fine to use the vocabulary of S-reasons to do all this. Calling S-reasons *reasons*, however, does not render them reasons in any distinct normative sense, just as saying the reason why the volcano erupted was that magna was building up involves nothing normative either.

It is worthwhile to stress, then, that my present contention is not with what are often called *motivating reasons*. I am not trying here to come to grips with what could be classified first and foremost a psychological phenomenon. It is true that people usually exemplify systematic patterns in their reasoning and behavior where goals and means become suitably interconnected; and further that those cognitive/conative patterns will be subject to and evolve in response to various ecological pressures. Effective practical reasoning certainly enhances our capacity to successfully deal with our environment; it contributes to self-preservation and flourishing in a hostile world. And it is also true that with systematicity in patterns comes understanding and prediction. Given that people tend to connect means and ends, we usually are able to cite the relevant sets of beliefs and desires that illuminate why they acted as they did; beliefs and desires that effectively present their choices in a favorable light and promise to make sense of them. The issue of motivating reasons, and of what sets them apart from both purely normative and purely explanatory reasons, is tricky. But to the extent to which we can distinguish between motivating and normative reasons – or at least between two roles reasons can play, a motivational and a normative role – my concern lies with the latter and not necessarily the former. Whether our agent A is under some psychological pressure to prefer X over Y and whether he is under some normative pressure to prefer X over Y are commonly supposed distinct matters. And here I do not contest that S-reasons, or our responsiveness to S-reasons, figure as crucial components of a complex psychological capacity – to reason and plan for the future – that has enabled our species to take over the world. I am mildly optimistic about the prospects that a decently plausible psychological story shall eventually be forthcoming explaining why we exhibit these kinds of systematic patterns in reasoning and acting linking ends and means. For all I know the story will include various evolutionary and cultural details.

But this concerns questions distinct from those we are presently occupied with. My focus here entirely rests on whether agent A has distinctively normative reasons to prefer X over Y, instead of merely and fortunately being subjected to cognitive habits that have served our species well. Or, to put the question slightly differently, whether the sense in which he does have such reasons is a distinctively normative sense, as opposed to a purely motivational and psychological sense of reasons.

Returning to this question, it is safe to observe that it is at least not obvious whether S-reasons are normative, and this alone spells trouble. Desire-based accounts of reasons face great opposition. The important point now is not to take sides in this debate, but notice that there are sides to take. And what opponents of desire-based accounts deny is certainly not the empirical observations from above: namely that doing X rather than Y will better serve A's end E. What they deny is that this, or this alone, determines A's normative situation. It appears, then, that in addition to whether we have the sort of complex quadruple relationship in front of us there is also the issue of whether that relationship captures what reasons are. For if there was no such further question, then specifying that relationship would be all there is to say, and the impression that there is a further issue one can potentially disagree with would amount to nothing but confusion. Yet we should not be that comfortable with passing verdict of confusion on entire traditions of philosophy.

It appears, in short, that there are two coherent hypotheses, neither one ruled out on logical or conceptual grounds alone:

- (1) Given S, doing X instead of Y will better serve A's end E and (hence) S is a reason for A to X.
- (2) Given S, doing X instead of Y will better serve A's end E but S is not a reason for A to X.

To put some meat on this, take some ax-murderer who wants to indulge in a particular brutal act just for the fun of it, and learns he could best achieved this by causing a blood-bath in the local kindergarten. Does he have a reason to do so, given that this is what he wants? It's controversial, to say the least, and that's all we need to focus on for the moment. Now, given that proponents of both hypotheses (1) and (2) acknowledge the fact that, given S, doing X instead of Y will better serve A's end E, there must be something else they contest, and therein precisely lies the problem for proponents of desire-based accounts. It seems they have mentioned everything there is to mention. What further fact is there in their desire-based arsenal? Since all parties agree on the empirical/relational facts that supposedly carry reasons, the question of which of (1) and (2) is true should be closed and settled, but it clearly is not.

Now, the easiest way to rule in favor of (1) over (2) is to introduce an additional principle. A principle of practical reason, rationality, or prudence that dictates that one should, given that in situation S, doing X rather than Y better serves ones ends, do X. Or perhaps it is as simple as that it is a good thing (bears truly commendatory force) if people get what they want. Kant suggested the principle that "Who wills the end, wills (so far as reason as decisive influence on his action) also the means which are indispensably necessary and in his power." (*Groundwork*, Paton translation) What precisely figures as the best candidate for this auxiliary principle is controversial and involves various complications. For now we can put all complications aside and stipulate that principle P establishes the requisite normative connection between situations, agents and ends on the one side and choices on the other. P would explain why e.g. it would be irrational for A to prefer Y instead of X given the facts are as they are supposed to be. The controversy, then, would concern principle P and its status. Unfortunately,

complementing desire-based accounts with one further principle P puts enormous strains on them; it threatens to undermine what motivates such accounts in the first place.

### 3. Desire cum Principle P

Principle P, we now suppose, has the power of settling the dispute between hypotheses (1) and (2) from above; in other words, it establishes that given agent A finds himself in situation S where doing X rather than Y better serves his ends, A has a normative reason to do X; or, at any rate, something reasonably close to that. As mentioned there are various difficulties with finding an adequate formulation of P. For present purposes, however, I shall put all these complications to a side. What I am concerned about is adding a normative principle to desire-based accounts in the first place, regardless of its content. The only point I need to stress, then, is that P is a *normative* principle. It comes with its own provision of oughts and shoulds which claim true normative force. Principle P is not a mere rule of thumb, a statement of some regularity, or a specification of our cognitive tendencies and expectations. Rather, it issues claims about what we should do, are required to do, or have reasons to do and so on. It introduces a practical norm the authority of which can be questioned and scrutinized. And therein precisely lies the problem.

What explains that normative status of P? If P has normative authority – if it is capable of establishing the relevant normative connections – this constitutes a most interesting normative fact in need of explanation. And now desire-based accounts face an uneasy situation. They must choose between two options: to provide an explanation or not to provide an explanation for the normative authority of P. And as I shall argue, neither choice quite works. The first choice would

be certainly preferable, since desire-based accounts aspire to explain reasons and authority, and if for that purpose they invoke a principle with its own normative authority, then we want to know why they are entitled to appeal to that principle and suppose it to carry normative force. The second choice would be to suppose principle P correct and authoritative, primitively, without further explanation. This would turn desire-based accounts into versions of non-reductive realism of the sort discussed in the previous chapter. It would propose and leave unexplained one normative principle, albeit one that potentially could claim particular intuitive support. I shall comment on this second choice first.

Accepting the normative authority of P as an unexplained primitive obviously amounts to giving up on the initial aspiration of naturalism. Because of this the second choice seems a non-starter qua naturalistic candidate. Still, some patience may be in order. If desire plus principle P is the closest we can get to explaining reasons in naturalistic fashion, and if this exception is motivated, it may not be such a terrible thing to excuse one single principle from the relentless demands of explanation. After all, the principle, together with the usual set of empirical/motivational facts, promises considerable explanatory power, as every practical reason henceforth would receive a principled explanation. Unsurprisingly, then, a number of proponents of desire-based accounts – Humeans as they usually call themselves – have urged their compatriots to accept this single and local exception from explanatory demands. James Dreier in particular has argued for this. He proposes the simple reading of principle P – or (*M/E*) in his idiom – such that “if you desire to  $\phi$  and believe that by  $\phi$ -ing, you will  $\phi$ , then you have a reason to  $\phi$ .” (2001:38) As mentioned before, the simple reading probably needs further amendments. For now, however, let us ignore these difficulties and follow Dreier in that

According to [this] reading, it is really a normative claim. It says, in effect, that you ought to perform the necessary and sufficient means to your desired ends. You might not do this. You might, at least on occasion, find yourself lacking the motivation to perform the necessary and sufficient means to some end you desire. This would be a fault of yours, a failure of rationality. Glossing over some distinctions, we might say that your failure would be a failure of instrumental reason. (2001:35)

Dreier then proposes that Humeans accept principle P, or (M/E), in the form of a categorical imperative, an imperative that “a person has reason to follow it that is independent of what she desires.” (2001:37). In other words, Dreier proposes to accept the normative authority of principle P irrespective of what people desire:

(M/E) has a kind of ground-level normative status. I think it also counts as a categorical imperative. Of course, the particular reasons that (M/E) generates are all hypothetical reasons. But (M/E) itself is not hypothetical. Its demands must be met by you, insofar as you are rational, no matter what desires you happen to have. That is why ... I think Humeans are mistaken to say there are no categorical imperatives at all. (2001:42)

These are astonishing words. Humeans have always dreaded the notion that there are demands that *must be met by you, insofar as you are rational, no matter what desires you happen to have*. There are many conceptions of practical rationality which involve robust and substantial demands: that one must promote one’s own well-being; that one must accord weight not only on present aims but also future aims; that one must strive for flourishing or sociability or perfection. And most importantly, of course, that, on pain of deep practical inconsistency, one must respect others if one respects oneself, and surely one must respect oneself. Humeans have regarded all this with profound skepticism. What they quintessentially contest is the attempt of substantiating the authority of ambitious norms on the basis of some robust and substantial conception of practical rationality. This precisely is the characteristic Kantian move, where a failure to comply with morality, for instance, is analyzed as a failure to comply with practical rationality as such.

Hence, the inception of desire-transcendent forms of practical rationality represents all that Humeans have traditionally been against. Dreier is fully aware of this singularity:

I will argue there is something special about exactly the kind of norms of rationality Humeans accept. This special status confers a kind of necessity on the Humean norms that we may properly doubt can accrue to other sorts of norms. The request for justification, I will argue, is intelligible as a demand for reasons bearing just that kind of necessity. And we may properly doubt that the demand for moral justification can be satisfied. So we may properly doubt that moral imperatives are categorical, but we must allow that some imperatives are categorical. (2001:37)

Unfortunately, granting this exception involves costs for desire-based accounts that far exceed the mere admission of a gap in explanation. It amounts to no less than the rejection of the most profound rationale for desire-based accounts itself: the aforementioned internalism requirement. If the normativity of principle P is taken to be primitive – that is not to be explained in terms of desire – the amended version of desire-based accounts cum principle P inevitably places itself in stark opposition to the internalism requirement. An agent may now respond to the dictates of P, namely that he must choose the means that best promote his ends, in exactly the same manner of philosophical puzzlement as he may responds to e.g. the dictates of some moral dictate, call it principle M, namely that he must help prevent the death of a starving child. Since according to the amended version principle P claims authority over agents regardless of whether it has a foothold in those agents, there is exactly the same sort of structural gap opening up between what reason-demands are proposed to agents and which in fact are guaranteed to engage those agents; between *engaging* reasons and *non-engaging* reasons. The bottom line, then, is that one cannot consistently use the internalism requirement to fend off some other principles on grounds of their potential disconnection to desire while at the same time give principle P a free pass. If principle P is structurally exactly alike those other principle, including moral principle

M, in that it too involves a claim to authority not accounted for in terms of desire, it ipso facto must share the same fate when it comes to whether it can live up to the internalism requirement.

The important point is entirely structural in character. It is also subtle, and we must ensure not to get sidetracked by various irrelevancies. This is not an exercise in psychology. As a matter of empirical fact, it would be highly unusual if agents were to respond to the recognized ways of furthering their ends with thinking “so what, what is that to me?” Few will do so, unless they are trying to score a philosophical point. If someone were indeed to respond in that way, we would start to wonder and probably would have cause for concern. We might start counting the days he will stay around. It would be crazy for him to disregard basic means-ends rationality that secures his further existence. But then, few will respond in that way when presented with commonsensical moral principles either. If someone were indeed to respond in that way, we similarly would start to wonder, and now would have even more cause for concern: now we might start counting *our own* days to stay around. It would be no less crazy to disregard basic moral principles that secure our *coexistence*. Generally speaking, the issue is not what *is psychologically likely* to happen but what is structurally *possible* to happen. Even if it so happened that every creature in the universe accepted principle M, it still would violate the internalism requirement, because it lays claim to authority in a manner irrespective of desire. Hence, it is entirely beside the point whether principle P or some other principle such as principle M comports to a greater extent with common empirical psychological patterns. The point, plain and simple, is that by assuming the normativity of principle P to be primitive and independent of our desires, this version inherits exactly the same structural features that enabled the *engaging/non-engaging* gap to open up for other more ambitious normative domains such as morality. It is because both principles P and M lay claim to authority irrespective of desire that it

becomes possible that what an agent desires and what he must do come apart. On what basis, then, are we entitled to award principle P the status of primitive authority that we are not also entitled to award to the basic moral principle M?

It shall prove most instructive if we keep alive the comparison between principle P and principle M, which represents the prototype of what the internalism requirement was devised to strike out. Consider an example of Dreier's that can easily be adjoined by a second and relevantly similar one. Consider Ann's case first.

We tell her that she ought to take a prep course for the LSATs. She asks why. We point out that she wants to raise her chances of getting into a competitive law school, and she can raise her chances by taking the prep course. She admits as much, but she still isn't motivated to take the prep course. So we cite the rule [of means-ends rationality, or principle P]. Now suppose that Ann agrees that this rule does indeed instruct her to take the prep course ..., given what she believes and desires, but she shrugs and doesn't accept the rule. (2001:38-39)

Ann obviously exhibits a severe form of practical irrationality. She would be the kind of person we are likely to get worried about. Don't plan any trips with Ann! Her practical reasoning is off, and we should fully concur with Dreier that *we must conclude that there is something wrong with Ann*. Now consider Antoinette's case:

We tell Antoinette that she ought to give her leftovers to the starving child outside. She asks why. We point out that the child will die otherwise, and that her donation involves virtually no costs to herself. She admits as much, but still isn't motivated to make the donation. So we cite the rule that one ought to help prevent a gratuitous death of an innocent child if one can do so without anything but marginal costs. Now suppose Antoinette agrees this rule indeed instructs her to help, given her situation, but she shrugs and doesn't accept the rule.

Antoinette obviously exhibits a severe form of immorality. She would be the kind of person we are certain to get worried about. Don't plan to do anything with Antoinette! Her

attitudes are repulsive, and so again *we clearly must conclude that there is something wrong with Antoinette.*

Ann, then, flouts principle P and Antoinette flouts principle M. What's the difference? The content of the two principles, of course, but nothing more. With regards to the relevant structural feature, namely the relationship of their presumed authority to desire, they are exactly on par. And since we are not presently concerned with the contents of normative principles at all, the answer should simply be: nothing, period. For sure, it is an open possibility that we conclude on substantial grounds that one principle is valid and the other is invalid. After all, they differ in what aspects of reality they enable to carry reasons; principle P places considerable weight on the concerns of agents and what furthers them while principle M places considerable weight on the capacities of agents to prevent gratuitous deaths. But this difference in content is irrelevant for present purposes and anyways, thus far we have not encountered any substantial argument for or against each principle. It is certainly not the injunction that Antoinette ought to hand over her leftovers that any participant in the present debate would want to take issue with. To paraphrase Bernard Williams, there are many things we want to say to her: "that [s]he is ungrateful, inconsiderate, hard, ..., nasty, selfish, brutal, and many other disadvantageous things. ... it would be better if [s]he were nicer..." (*Internal reason and the obscurity of blame*, in: *Making sense of humanity*, Cambridge UP 1995:39) But neither do I wish to take issue with means-ends rationality. Three cheers to means-end rationality! Both principles, then, involve norms most would find intuitively compelling and few would wish to discard. The question does not concern our de-facto inclinations towards both principles, but that they both lay claim to authority in the same desire-independent manner, and that consequently, as far as the internalism requirement is concerned, they both stand or fall together.

The amended version of desire based accounts, then, attempts to quarantine what has the capacity of giving reasons to agents – exclusively, namely, what has the right kind of relation to what they care about – only to invite back in a principle that brings with it all the features the quarantine was erected to keep out in the first place. The line the amended version seeks to draw demarcating acceptable from unacceptable principles is inherently unstable: by drawing it where it wishes it to be drawn it must reemploy an element plainly falling on the other side and hence censured off hands. Yet desire-based accounts cannot have it both ways. They ultimately shall prove incapable of withstanding the dual pressure from left and right, which is simultaneously applied by those, on the one side, who seek to acknowledge the authority of additional desire-independent principles, and those, on the other side, who even question the authority of the single principle P: Philosophers such as Hume, who famously contested the authority of means-end rationality, such as contained in principle P, when he wrote *Tis as little contrary to reason to prefer even my own acknowledg'd lesser good to my greater.* (*Treatise* 2.3.3.) To respond to the skeptics who insist each principle must be accounted for and who are prepared to jettison all wanting an explanation, proponents of the amended version must plead to allow one exception, and to accept principle P despite its disconnection to desire. Yet thereby they inevitably become vulnerable on their other flank.

And indeed, those who wish to advance more ambitious principles to carry normative force are quick to point their finger at precisely this instability. Christine Korsgaard notes that “... once this kind of irrationality is allowed in the means/end case, some of the grounds for skepticism about more ambitious forms of practical reasoning will seem less compelling.” (*Creating the Kingdom of Ends*, Cambridge UP 1996:321) Here, then, do we find two representative figures questioning the soundness of the amended version of desire-based

accounts. Korsgaard, coming from one end, will press to allow in principle M as well, given that principle P, which is already supposed inside, no less violates the internalism requirement than does principle M. And Hume, coming from the other end, will press to disallow principle P as well, given that principle M, which is already supposed outside, no less violates the requirement than does principle P. Both Hume and Korsgaard appear consistent in a way in which the amended version does not. Desires-based accounts, then, cannot consistently refuse to accept principles such as principle M on grounds of its independence to desire and yet at the same time accept principle P even though it exhibits just the same sort of independence of desire.

Desire-based accounts, I conclude, are not entitled to accept the authority of principle P without explanation. This precludes the second option of how they could incorporate principle P. The first one, which we initially declared preferable anyway, is to provide an explanation for the authority of principle P. Do desire-based accounts, then, have the resources to offer an account of the authority of principle P? Unfortunately, they have not.

The only element desire-based accounts can draw on to explain the authority of some principle is desire. When, and in virtue of what, we like to know, does an agent have a reason to comply with a principle? The answer, desire-based accounts must say, is when, and in virtue of the fact that, the agent has accepted the principle; when, in other words, honoring the principle features as one of his concerns. Suppose an agent accepts some principle and now faces a choice where he can either do something that honors the principle or alternatively do something which dishonors it. He must have a reason to do what honors the principle, we hear the imaginary proponent of desire-based accounts saying, because it furthers something he cares about, namely to honor that principle.

But now we squarely find ourselves in a circle. We introduced principle P precisely because we recognized that the mere empirical correlation between what someone desires and what furthers his desires cannot carry distinctively normative reasons. A gap remains, and principle P was called upon for the sole purpose of closing that gap. Hence in the attempt of explaining principle P's own authority one cannot presuppose what principle P was employed to explain in the first place, namely that with the appropriate provision of desire comes reason. In other words, if what is at stake is the reasons-related capacity of desire, it doesn't help to add yet another desire. If the transition from what furthers desires to what we have normative reasons to do requires an extra step, we make no strides by co-opting an extra desire for taking the relevant step.

Return, then, to Ann's case from above. Ann wants to get into law school and recognizes that she will more likely be able to do so if she takes a prep course for the LSATs, but then refuses to do so all the same. We propose principle P to Ann, and she acknowledges that principle P directs her to take the course under those circumstances. She still has no inclination to take the prep course. Now, we concluded there is something wrong with Ann, but wondered what exactly is missing. "What Ann is missing can't be any desire" Dreier keenly notes. We cannot get Ann to take the course by just adding yet another desire:

The desire that is supposed to bridge the gap between believing that a rule requires her to  $\phi$  and being motivated to  $\phi$  is the desire to comply with the rule. But suppose that Ann's mental inventory were supplemented with a desire to comply with the rule, in this case to comply with (M/E). Could this complete the picture? Were she to desire to comply with (M/E), would she then be motivated to take the LSAT prep course? By hypothesis Ann ... fails to be motivated by the acknowledged means to her desired ends. So adding a desire (complying with (M/E)) does not in her bring about the motivation to perform an acknowledged means to her end of doing well in the LSAT. ... This futile attempt is exactly what we would be engaged in if we were to try to bring

Ann to take the LSAT prep course by giving her a desire (complying with (M/E)) that would motivate her to take the prep course. So what Ann is missing cannot be a desire. (2001:39)

Dreier frames the argument in terms of motivational force, but it works just as well in terms of normative force. If Ann generally doubts she has reasons to do what furthers her ends, she will not regard her newly instilled desire to comply with (M/E) to conclude in a reason either. Supplying her with the goal of honoring principle P only entails that she now has a reason to comply with the principle if we are entitled to suppose that she has reasons to do what best satisfies her goals, which, of course, is nothing but a restatement of principle P, the very principle at issue. This addition in desire thus cannot change the situation. Previously we dealt with a single desire of Ann's, and we were flabbergasted by her unwillingness to proceed to means she knows will implement that single desire. Now we are dealing with two desires of Ann's, one about getting into law school and one about honoring principle P, and Ann again exhibits the same degree of unwillingness to proceed to means she knows will implement these two desires. The number of desires, however, is not the problem; rather, the problem entirely lies in the transition from desire to reason, be it one or many, and thus we are simply moving in circles here.

Dreier calls this line of reasoning the Tortoise argument given its resemblance to the famous story by Lewis Carroll. There we encounter a clever tortoise unwilling to draw an inference of the modus ponens sort and who is adjoined by a rather dull Achilles attempting to entice her to do the same by offering one after another inference of the modus ponens sort. If the Tortoise stubbornly remains disinclined to accept modus ponens – or doubts she has reasons to do so – then presenting yet another modus ponens argument for accepting modus ponens will not make the slightest difference, neither for her inclinations nor her conception of what she has

reasons to do. Likewise, if we already question that desire-based reasons contain distinctively normative force, then providing a desire-based reason on behalf of the authority of desire-based reasons cannot do the trick either. If we don't get from desire to reason, we don't get from desire to reason, period, regardless of whether the case at hand is of the concrete sort on a first-order level or involves principles on a higher-order level.

The fundamental problem, then, is not (1) that by accounting for the normative authority of principle P by way of desire we inevitably would limit the scope of its application; that only those agents who harbor the relevant class of higher-order and principle-related desires could truly be said to have reasons to implement what they acknowledge are means to their desired ends. We could come to accept that only Ann in the second case but not in the first case has reasons to take the prep course, even though both are fully aware that taking the course increases their chances of getting what they want. Nor is it (2) that the authority of principle P would become escapable in a rather blunt manner, for all one needed to do in order to relinquish its burden is to abandon the requisite higher-order and principle-related desire. Perhaps we cannot just discard desires directly, but a good recipe for achieving this indirectly would be to violate the principle with great regularity such that its psychological hold eventually wears off. Nor is the problem (3) that it becomes hard to see how principle P could retain much independent force and point; that one would either have the requisite motivations to honor principle P, in which case one would also have reasons to do what one is inclined to do anyway; or one would not have the requisite motivations to honor principle P, in which case one would not act contrary to any reasons. The absence of the corresponding inclinations would conveniently excuse one from being held answerable to the principle. It is not immediately evident, to say the least, how anyone could ever come to *dishonor* principle P in the first place. Put together, what (1) and (2)

and (3) cause are rather marginal worries by comparison. The fundamental problem is that even with respect to someone who does have the requisite principle-related desires and who does renounce all plans to escape it and who is motivated to honor it – the perfectly practically rational agent – we still have no account of why he should have any reasons to choose those means he acknowledges further his ends.

The amended version of desire-based accounts thus collapses. Accepting the authority of principle P without further explanation undermines the very rationale for desire-based accounts in the first place. And the attempt of providing an explanation of the authority of principle P has proven futile. The principle-adjoined version of desire-based accounts does not represent progress over the principle-free version.