

Not by Reasons Alone

by

Kate Manne

B.A.(Hons) in Philosophy, The University of Melbourne

Submitted to the Department of Linguistics and Philosophy
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Philosophy

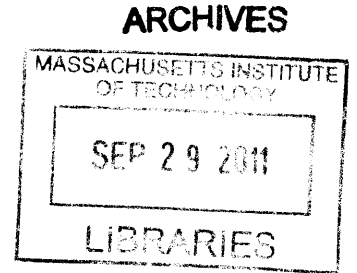
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2011

[September 2011]

© 2011 Kate Manne. All rights reserved.



The author hereby grants to MIT permission to reproduce
and to distribute publicly paper and electronic copies of
this thesis document in whole or in part in any medium
now known or hereafter created.

Signature of Author

Department of Philosophy
May 3, 2011

Certified by

Richard Holton
Professor of Philosophy
Thesis Supervisor

Accepted by.....

Alex Byrne
Professor of Philosophy
Chair of the Committee on Graduate Students

Not by Reasons Alone

by

Kate Manne

Submitted to the Department of Linguistics and Philosophy on May 3, 2011 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Philosophy.

Thesis Abstract

My thesis is a sustained argument that the practical reason is not a suitable master concept in ethics, let alone the only ethical notion we need.

In Chapter One, I defend the idea that desires have irreducible and pervasive normative significance. More particularly, I defend reasons internalism – the claim that desires are a necessary condition on practical reasons – by developing a new version of this thesis. In this version, desires serve to veto practical reasons, which are in turn grounded in objective values. I argue that this compatibilist picture provides an account of self-interested reasons which is intuitively superior to its purely objectivist rivals.

In Chapter Two, I argue that evaluative notions are distinct from prescriptive notions, which reasons talk is meant to encode. According to my account, it is partly constitutive of wickedness (an evaluative notion) that the wicked person lacks moral reasons – understood as the basis for potentially apt prescriptions – to mend his ways. For some people, I suggest, are deaf to moral instructions in something close to a literal sense. I argue on this basis that the distinction between evaluative and prescriptive ‘oughts,’ and the attendant possibility of iterating them, vindicates internalism about moral reasons too. A solution to Chisholm’s paradox is a welcome fringe benefit.

In Chapter Three, I begin to develop an alternative to prevailing reasons-based conceptions of ethics, by focusing on social relationships, such as friendship. I argue that agents can behave decently by being guided by implicitly normative concepts like friendship, which contain codes of conduct like “Friends help each other out,” and “You don’t snitch on your friends.” Such ‘do’s and don’ts’ and the corresponding concepts enable agents to behave well instinctively, even when they believe they have no reason to do so. I argue further that recognizing who someone is – i.e., the social relationship in which they stand to you – can be action-guiding and even mandating. I argue that this kind of social awareness is a viable alternative to positing intuitive responsiveness to reasons, and defend the idea that it underwrites a form of practical necessity worthy of the name.

Thesis Supervisor: Richard Holton

Title: Professor of Philosophy

Not by Reasons Alone

by

Kate Manne

Table of Contents

Introduction.....	5
Chapter One Resurrecting Reasons Internalism: The Veto Power View of Desire.....	10
Chapter Two Regimenting Reasons Internalism: Moral Reasons and Iterated ‘Oughts’.....	43
Chapter Three Reconceiving Ethics: Doing without Reasons in Practice.....	71
References.....	127

ACKNOWLEDGEMENTS

Material in these chapters was presented in 2010 at the Australasian Association of Philosophy (AAP) Conference, the workshop on gender and philosophy (WOGAP) at MIT, Boston University's Ethics Reading Group, and a colloquium at Southern Connecticut State University. I also benefitted greatly from giving talks in 2011 at Dartmouth College, Mount Holyoke College, Claremont McKenna College, Washington University at St Louis, and Cornell University. I thank these audience members for their helpful comments and feedback.

I owe a debt of gratitude to many teachers and friends who have helped me in writing this thesis. In particular, I would like to thank my committee members – Richard Holton, Sally Haslanger, Julia Markovits, and Rae Langton – for being superb mentors, philosophers, and people. I would like to thank each member of the MIT community for contributing to making it an unsurpassable environment in which to do philosophy, especially as a woman. The mentorship of Rai Gaita and Larry Blum has also been invaluable. Finally, I would like to thank my family for seeing me through. For my parents' love and love of ideas I will be forever grateful. This thesis is dedicated to Daniel, my husband, *sine qua non*. Virtually every idea in here bears traces of our conversations, our life, our love.

INTRODUCTION

“Often in practical philosophy the dominance of one view is the result of its rivals ceasing to make sense. It seems difficult to argue for the orthodox view, and impossible to argue against it. Its correctness is manifest. Rival heterodox views are not there to be refuted. They are condemned through their own unintelligibility. We can just understand what their supporters are saying, but it seems so pointless. It is mysterious how anyone might maintain such a view, unless they are blind to simple conceptual connections. To argue for the orthodox view can amount to no more than pointing out those connections.” ~Joseph Raz, “Liberating Duties” (1989), p. 5

It has become an orthodoxy in contemporary ethics that our subject matter is the practical reason. This orthodoxy is rarely argued for, or even acknowledged to be one picture among many. But I think this recent turn to reasons has done little to enhance our understanding of moral psychology, moral character, and the connection between ethical life and social practice – quite the contrary, in fact.

I know of few philosophers (and fewer ethicists) who share my qualms about reasons. The idea that the practical reason is the master concept in ethics, perhaps even the only primitive normative notion we need, has been widely touted. And one can quite readily understand why. In 1958, G.E.M. Anscombe argued that the moral ‘ought’ has dubious authority, even meaning, in a godless world. ‘What is a command without a commander, a law without a legislator?’ she forced us to ask. In 1977, J. L. Mackie complained that the idea that some actions have ‘to-be-doneness’ somehow built into them was, on reflection, extremely ‘queer’ (1977, p. 40). These two philosophers could hardly be more different in temperament. But their suspicions are interestingly of a piece. We learn, as children, that this or that action is wrong because it is a piece of deliberate cruelty. “But just what in the world is signified by this ‘because?’” we’d like to know (1977, p. 41). For the world is now, to adapt a line, the only thing that is the case.

So Anscombe and Mackie both worry that, in a secular and scientific era, core moral notions lose not only their grip but even their intelligibility. The turn to practical reasons is best understood, I think, against the backdrop of this concern. T.M. Scanlon writes, in the introduction to *What We Owe to Each Other*.

I do not believe that we should regard the idea of a [practical] reason as mysterious, or as one that needs, or can be given, a philosophical explanation in terms of some other, more basic notion. In

particular, the idea of a reason should not be thought to present metaphysical or epistemological difficulties that render it suspect. (1998, p. 3)

Why are practical reasons thought to be metaphysically and epistemologically unproblematic? As might be suggested by Scanlon's frequent omission of the modifier 'practical,' there seems to be an appeal to 'companions in innocence' going on here. Theoretical reasons are, after all, virtually unimpeachable. Reasons for belief are implicated as soon as one is in the business of giving and asking for them (in doing philosophy, among other things). Reasons for action thus look like a handy adaptation of a mercifully mandatory idea. Hence Scanlon's opening gambit:

My strategy will be to locate reasons, in the sense I will be concerned with, as the central element in a familiar form of reflection, and to call attention to structural features which I argue are common to thinking about reasons of all kinds: reasons for belief, for action, and for such attitudes as fear, resentment, and admiration. (1998, p. 18)

Scanlon's main ally in the turn to reasons is Derek Parfit. In his forthcoming magnum opus, *On What Matters*,

Parfit likewise opens:

We can have reasons to believe something, to do something, to have some desire or aim, and to have many other attitudes and emotions, such as fear, regret, and hope. Reasons are given by facts, such as the fact that someone's finger-prints are on some gun, or that calling an ambulance would save someone's life. (Forthcoming, p. 31)

So, on the one hand, we have a relation between facts with probative force (i.e., evidence) and beliefs, or propositions ripe to be believed (mind to world). On the other, we have a relation between facts with practical significance and actions, or states-of-affairs ripe to be brought about (world to mind). The relation in question in both instances is just a univocal, *sui generis* 'favoring' relation. Theoretical and practical rationality simply require being sensitive to this relation, and following through in belief and action respectively. What could be simpler or better as a strategy for retaining our grip on morality?

But a little reflection on the epistemological analogy tends to undermine the ambitions of those who, like Scanlon and Parfit, propose a reduction of ethics to practical reasons. For, far from just invoking an unanalyzed notion of evidence, epistemologists have to say something pretty intricate about ‘favoring’ a.k.a. justification, accommodate complexities like defeaters for evidence (which need not themselves be pieces of evidence, as with undercutting defeaters), account for the value of knowledge, answer the question ‘Why be rational?’, and explain the social dimensions of knowledge acquisition too. Positing theoretical reasons (i.e., evidence) is not supposed to take the pressure off in epistemology; why should moral philosophers expect to fare better, vis-à-vis practical reasons?

I argue in this thesis that, once the notion of a practical reason is subjected to critical scrutiny, the epistemological analogy is in one way *too* good to help the reasons centrists’ cause. For in ethics we similarly need to appeal to an array of other normative notions in order to develop the best account of practical reasons themselves – a task that turns out to be, *pace* Scanlon and Parfit, quite feasible. (See Scanlon, above, who Parfit echoes in claiming that “the concept of a reason is indefinable in the sense that it cannot be helpfully explained merely by using words;” Forthcoming, p. 31.) Thus, although practical reasons have an important place in ethics, I argue that they will be best understood insofar as they are not the only primitive normative notion. For one thing, we need to appeal to desires (and lack thereof) in order to formulate an adequate account of self-interested reasons and, relatedly, well-being (this is the main thrust of Chapter One). This claim is leveled against those who believe, like Scanlon and Parfit, that “desire is not a clearer notion in terms of which the idea of having a reason might be understood” (1998, p. 7). For another thing, we need to appeal to character evaluations like negligence and wickedness in order to formulate an adequate account of moral reasons and ‘ought’ talk (this is the main thrust of Chapter Two). This claim is leveled against those who believe, like Scanlon and Parfit, that evaluative notions can be understood in terms of reasons (1998, p. 95). Not so, I argue.

The upshot of these two chapters is that the program of reducing ethics to reasons gets things largely backwards: in order to best understand reasons, we need to make recourse to several other normative notions along the way. So, just as in epistemology, reasons are best treated as the *explandum*, not the *explanans*. This should not be terribly surprising – reasons are a very thin notion, after all, and therefore should not be counted on to do a lot of explanatory work.

According to my account of moral reasons, wicked people characteristically lack certain moral reasons to act. It follows that we need a story about moral sensitivity (and lack thereof) which explains what they are missing. And the story must improve on the now evidently dubious suggestion that they are insensitive to the moral reasons that – by hypothesis – fail to hold for them. In Chapter Three, I aim to develop the beginnings of a suitable alternative story. On the account I favor, moral sensitivity is not a matter of being sensitive to the reasons we have, or even impervious to an independent realm of abstract values. Rather, it is primarily a matter of not recognizing the social relationships in which we stand as (*inter alia*) human beings.

Attention to the social brings out a way in which the analogy between theoretical and practical reasons is not good *enough* for the reasons centrist's project. When it comes to epistemology, asking whether or not some fact is a reason to believe some proposition is typically unproblematic. That is, we can isolate the fact from the norm without undue conceptual difficulty. Admittedly, some epistemologists believe that questions like – 'Is the fact that his fingerprints are on the gun a reason to believe the butler did it?' – can only be answered against a tangle of other beliefs, theories, assumptions, and epistemic practices generally. But when it comes to ethics, the difficulty of separating the fact from the norm is a different order of magnitude. For the sorts of practices that give many important social notions their content can undercut and pre-empt the ethical questions that we might think of raising about them. As a result, questions like "Does the fact that it's your *wife* drowning give you a special reason to save her?" are uneasy, even arguably ill-posed (see Williams, 1981,

p. 18). For it is partly constitutive of moral-cum-social relations like marriage that *one does not let one's wife drown* over a 'competing' drowning stranger. Ethical chromatography is thus in my view a non-starter.

As well as these issues in metaethics, there is also a methodological issue at stake here. I think we need to focus less (at least at the outset) on rescuing ethics from supposed metaphysical obscurity and more on recovering the notions with which we navigate the social world. Notions like friendship, marriage, and even humanity – along with their sinister correlates, i.e., enmity, slavery, and sub-humanity – call for careful but non-squeamish excavation. So too do the related class of thick moral concepts (ripe for character assessment) and notions that recommend action less than they mandate it (e.g., 'have to' and 'must'). It remains to be seen what our ontological commitments will be at the end of the investigation, or what the relationship between these commitments will ultimately be. But as W.V. Quine has taught us, this is not to ignore metaphysics, but rather to roll up our sleeves and do it. And, as John McDowell has suggested, we might hope to thereby eventually answer the challenge to secular moral authority not from outside ethics, but rather from inside its resurrected walls.

So those would be my recommendations and ambitions for ethics going forward. In this thesis, I try to show that we have little choice but to seek out alternatives to reasons centrism in any case. For the best conception of a practical reason is, I will shortly argue, internalist. That is, reasons are only reasons *for* us when they can get a grip on us, by appealing at some level to our existing motivations. If that is right, then moral reasons cannot cover the field of ethics – there being truly callous people who are indifferent to the suffering of others. We will have to look to other moral and social notions in order to explain their defects. So we cannot do our best moral philosophy by means of reasons alone. This, in a nutshell, will be the arc of my argument.

Chapter One

RESURRECTING REASONS INTERNALISM: THE VETO POWER VIEW OF DESIRE

OR: You Have to Want It

“Fancy not being able to see that the real cause is very different from the mere sine qua non of any cause!” ~Plato, *Phaedo*, 99b

Bernard Williams famously espoused *internalism about reasons*, claiming that practical reasons require the backing of the agent’s desires (“Internal and External Reasons,” 1979; reprinted in his *Moral Luck*, 1981). More recently, Derek Parfit has championed *objectivism about reasons*, claiming that practical reasons are provided by desire-independent, objective values (in his forthcoming book, *On What Matters*). But these two positions have long been billed as incompatible, and a large literature has sprung up around this now entrenched debate in metaethics.¹ In this chapter I argue that the debate was ill-founded to begin with: it rests on a tacit metaphysical assumption that the only way in which a fact could be normatively relevant is by *providing* practical reasons. But this assumption is false. As well as facts which provide practical reasons, we need to make room for facts which *sanction* or – to put the point negatively – *defeat* practical reasons. (Epistemologists routinely invoke this distinction for theoretical reasons, so why should practical reasons be any different?) With this distinction in hand, it quickly becomes apparent that one can be an internalist *and* an objectivist: the idea being that practical reasons are provided by desire-independent objective values, but that *you have to want* the objective good in order for it to exert its potential normative force. This constitutes a novel view about the normative significance of desires – or, more generally, of what we care about –

¹ Parfit’s objectivist allies include Warren Quinn (1993), T.M. Scanlon (1998), Joseph Raz (1999), and Jonathan Dancy (2000). (The term ‘externalist’ is sometimes used instead of ‘objectivist,’ but this would be misleading given the tenor of my argument.) Recent sympathetic discussions of internalism include Stephen Darwall (2003) and Ruth Chang (2004); Mark Schroeder (2007) and Sharon Street (2009) have advocated views in the neighborhood too. There is also a sizeable literature on Williams’ original argument for internalism, including recent papers by Stephen Finlay (2009) and Keiran Setiya (forthcoming). But I will be leaving Williams’ argument to one side here, and striking out on my own. Finally, there are prominent conceptions of practical normativity which stand in interesting relations to Williams’. However, I wouldn’t be able to do justice in this space to the views of Michael Smith, Christine Korsgaard and Candace Vogler, among others.

according to which desires don't provide but rather rubber-stamp objective would-be reasons. Thus, I call it the *veto power view of desire*. This chapter aims to develop this view of practical reasons, and to defend it as applied to self-interested reasons. (I defend it as applied to moral reasons in the next chapter, showing that it does not lead to vulgar moral relativism or subjectivism – on the contrary, when properly understood, it turns out to be at least as morally serious as its rivals.) The pay-off in the current chapter is not only the resolution of a metaethical log jam, but also a compatibilist theory of well-being which should make everybody happy.

1: THE KEY DISPUTE, ILLUSTRATED

To locate the first-order dispute this chapter seeks to mediate, imagine that you are an 'agony aunt' with a popular advice column. One day you receive a letter from an impoverished painter, aged thirty-five. He explains that he has been struggling to make ends meet for the better part of a decade, ever since he left his humdrum but respectable job to dedicate himself to his calling. He is trying now to decide whether to continue painting, or to give up and buckle down to a 'real job' again. His artistic zeal has not waned; on the contrary, it is what gives his life a sense of purpose. But his current existence is pretty wretched, and he is not optimistic that things will ever improve. He has no spouse to consider, nor any children to support. Even so, he writes:

It is a gloomy prospect to have to admit to myself that my paintings will probably never be worth a penny. The public won't change, and likes only easy, pretty pictures. I still don't care about making money, but as I grow ever poorer, I do feel the pinch. So, if I continue painting, I will live almost like a monk or hermit, working for my master passion, and giving up a comfortable life. I shall be forced into poverty and social isolation. I will struggle to put a roof over my head, to secure even the absolute necessities, and to withstand the siege of failure which I suspect will last all my life.

Since I am getting on in years, I try to hold no illusions, and to reflect soberly on the wisdom of things before undertaking them. Sometimes I feel too feeble to keep fighting against my present circumstances, and I would surely have to be cleverer and richer and younger to actually triumph. I do not hanker after victory anymore, though.

Were I ambitious, I would doubtless feel differently. But I am utterly indifferent to success or failure. I started to sign my canvases once, but I soon stopped; it seemed too silly for words. For I care neither about fame nor fortune, nor even about my own material comfort. I find all I care about is dedicating my remaining energies to impressionism.

So, dear Abby, what should I do?

With a handshake, and ever yours,
Vincent

P.S. Enclosed is a painting of a bunch of flowers, and a still life of an old pair of shoes.²

So: passion or prudence, which should Vincent choose? And how would you go about giving him *advice* as to which path he ought to take? Being an unusually meticulous agony aunt, loath to shoot from the hip, you might make a start by drawing up two columns: one to list the pros of Van Gogh's ceasing to paint, another to catalog the cons. But even before you began assigning probabilities to the various possibilities, or weighing very different sorts of goods, you might find yourself hesitating about how to fill in the blanks.

For suppose that these declarations were not mere posturing on Van Gogh's part, and that he genuinely didn't care about his own material comfort. Then did the fact that his way of life was (by his own admission) uncomfortable indeed still *count against* his living so meanly? And suppose his indifference to worldly success was wholly sincere too. Then did his subsequent reputation nevertheless *count in favor* of his persisting in painting? Of course, things ended rather badly for poor Van Gogh – by most reports he went quite mad – which might admittedly give us pause in this particular case. But the general question this chapter takes up still stands. Namely, do you have self-interested reasons to pursue objective would-be goods (like fame and fortune), and to avoid objective would-be ills (like poverty and social isolation), even if you genuinely don't care about them at all? More especially, do certain *pleasures* we decline to have in life, and episodes of *suffering* we willingly embrace, still have the usual impact on our well-being, on what would be in our best interests, or

² This is my adaptation of two letters from Vincent Van Gogh to his brother Theo, dated August 1888, his biographical details as above. I have used poetic license here in merging parts of both of these letters, and then rearranging sentences to restore cohesion. But – with the exception of the penultimate line of course – these are all loose quotations, including the postscript. I intend the example to echo Williams' key illustration of reasons internalism: the character of Owen Wingrave, who had no desire to join the army, notwithstanding his family's fervent protests. But it is also intended to evoke the example of Gauguin, which appears in a different context in Williams' work, and whose relevance I go on to explain in *n.* 17.

on what would make a good life for us?³ Do they belong, in other words, in the agony aunt's two blank columns?

No, Williams would have said; yes, Parfit would instead insist. In this chapter, I want to persuade you (the agony aunt in training) to side with Williams the internalist, and to demonstrate that – *contra* Parfit – we can be internalists without endorsing the bleak view that *nothing matters*, objectively speaking: Parfit's ultimate stalking horse. To that end, I will suggest in the next section that internalism is intuitive in cases like Van Gogh's. I will then explain how internalism comes under attack, due to an inadequate metaphysics of reasons – which makes it look as if the view has implausible and downright nihilistic implications. I will show how the attack can be summarily deflected in section 3, by means of the veto power view of desires, with its superior metaphysics enabling us to maintain internalism while espousing objectivism all the while (so things certainly *do* matter, objectively speaking, on my view). In sections 4–6, I flesh out this novel proposal in enough detail to show (in section 7) why appending the internalist condition to objectivism in the proposed way carries the day over a purely objectivist position like Parfit's. By way of a preview, I argue that there is something very off about the (hard to resist) implication of pure objectivism that every objective good – like pleasure, paradigmatically – must be good for its recipient, irrespective of whether or not she wants it. I will draw on a rich range of examples in this connection – e.g., *schadenfreude*, illicit pleasures, ignoble pleasures, and pleasure imposed on someone against her will – which together serve to show that the veto power view has broad intuitive appeal. I will go on to consider the flipside of pleasure and desire, i.e., suffering and aversion, in due course as well. But I anticipate.

³ These are all, I will take it, ways of asking essentially the same question. The terms 'flourishing,' 'happiness' and *eudaimonia* are other notions in the vicinity, but they are too theoretically loaded for my purposes. The subject is also sometimes broached using terms like 'quality of life' and 'welfare,' which I'll eschew here. These terms are naturally reserved for the *policy goals* we should adopt, not the concept of well-being in my sense – i.e., what makes someone's life go well, from the perspective of the person living it. I'll also avoid the potentially misleading term 'prudential reason' here, since the concepts of prudence and pleasure (a popular candidate self-interested good) are virtually antagonistic, as we saw in Van Gogh's case.

2: INTERNALISM, LOST?

Why be an internalist about reasons? Cases like Van Gogh's tend to provide much of the initial impetus. For there is something very strange about the idea that someone has a standing reason to go after would-be goods, like bourgeois pleasures and creature comforts, which he really doesn't care to have. It thus seems fitting that our advice to someone like Van Gogh is constrained by what they want out of life. It isn't just that the putative reasons above aren't decisive: we can all agree that reasons are merely 'contributory' entities (the usual line being that you ought, all things considered, to do what you have *most* reason to do). But the intuition in play here is rather that these aren't even considerations to be weighed. Nor is it just that telling someone to go after goods he doesn't want would be mere browbeating, or a boring and pointless exercise: even if we could somehow convince him to do so, it would seem to miss the very point of advice. Something can't matter *for* you unless it matters *to* you, would be one way of putting the driving thought here.

So, internalism is intuitive – as long as we spell out the thesis carefully enough to let these intuitive attractions shine through. Suppose we go with the following formulation:

INTERNALISM: An agent can have a reason to φ *only if* she has some desire that would be furthered by φ -ing. (Or, to paraphrase, an agent's reasons *depend upon*, or *require the backing of*, her desires.)

Such ways of articulating internalism have the virtues of simplicity and generality. So as to isolate the best position in the vicinity, however, we need 'desire' to be taken here in the broad and rather idealized sense that the agent *could* now become motivated to φ were she to be sufficiently well-informed. (We don't want to say that a spendthrift doesn't have a reason to save for retirement if she's under the misapprehension that money grows on trees, and so doesn't have an *actual* desire to start building a nest egg.) More precisely, I will follow Williams from now on in saying that an agent has a desire to φ iff there is a "sound deliberative route" from her existing motivations to her being motivated to φ , were she disabused of any relevant false (non-

normative) beliefs and given further germane information (1995, p. 35).⁴ It is also vital to read ‘reason’ here in such a way that internalism does not end up being trivially true. Hence, the term ‘reason’ refers throughout to facts which favor or disfavor a course of action, irrespective of the agent’s false beliefs or faulty reasoning, and even if her ignorance or error blind her to the normative force of those reasons. Indeed, even though Van Gogh could not possibly have known about his eventual celebrity, this fact could still have given him a reason to keep painting in the sense intended here. This is to construe reasons as *normative reasons*, not so-called *motivating reasons*, which incline the agent to act and subsequently explain her action. Obviously, if internalism holds, then there is an intimate connection between the two notions. But to avoid confusing the issue or begging any questions, I will express claims about the agent’s motivations for acting as claims about her driving *rationale* from now on.

When properly spelled out, internalism about reasons is intuitive, I claim – and I will continue to beat the drum with many other examples over the course of this chapter. Why is internalism so controversial, then? Let’s hear from Parfit, explaining why he is a firm opponent of ‘subjective’ theories, according to which “our reasons for acting depend on facts about what would fulfill or achieve our present desires or aims, or on facts about what we choose” (p. 44).⁵ He thus counts Williams-style internalism as a species of subjectivism which “appeals to the desires or aims that we would now have, or to the choices that we would now make, if we had carefully considered all of the relevant facts” (p. 44). Here is Parfit’s central objection to theories in this whole family:

⁴ I leave it open that we might also have to help her to be more imaginative, as Williams suggested (1981, pp. 104-105), or to relieve her of depression, neuroses or bias – i.e., to have her undergo a course of ‘cognitive psychotherapy,’ to borrow a notion of Richard Brandt’s (1979). Thus, ‘desire’ is a really a term of art here, even encompassing – in Williams’ view – “such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent” (1981, p. 105). But while admittedly broad, this usage of ‘desire’ is not so unnatural. We have something like Williams’ notion in mind when we talk about what people genuinely or truly care about, or what they “really want, deep down.”

⁵ Manuscript dated 28 April, 2008; all subsequent references to Parfit are to this manuscript, unless otherwise stated.

According to subjective theories, if we had fully informed desires to hit our howling baby, or to smash some malfunctioning machine, these facts would give us reasons to hit our baby and smash this machine. If what we most wanted and chose was to frustrate all of our future desires, this fact would give us a decisive reason to frustrate all these desires. If what we most wanted and chose was to waste our lives, and to achieve other bad or worthless aims, these facts would give us decisive reasons to waste our lives, and to try to achieve these bad or worthless aims. (pp. 75-76)

But, as Parfit quite reasonably holds, “we would have no reason to hit our howling baby, to waste our lives, and to try to achieve other bad or worthless aims. So subjective theories are built on sand,” he subsequently goes on to conclude (p. 77). Upping the ante, he adds, by way of a summary: “On subjective theories, *nothing matters*. We should reject the arguments for this bleak view” (p. 14).

Parfit instead goes on to champion:

OBJECTIVISM: An agent’s reason to φ is always *given by* (or – equivalently – *provided by*, or *generated by*) the facts that would make φ -ing objectively valuable, independently of her desires.⁶

Setting aside moral values for now, Parfit thinks that, when it comes to self-interested reasons, “on any plausible theory, hedonism is at least a large part of the truth” (p. 40). Parfit allows that there might be additional self-interested objective goods on the list besides pleasure (e.g., love, friendship, knowledge, and the appreciation of beauty). But he thinks that we certainly have self-interested reasons to pursue what we would *like* – e.g., enjoy, find satisfying, or take pleasure in – and to avoid what we would *dislike* – e.g., suffer from, be displeased by, or find painful. Of course, *liking* and *desiring* often go hand in hand: but examples like Van Gogh’s show that the two can indeed diverge. And we will be seeing many more cases in which motivation and satisfaction again come apart before the chapter is out.

⁶ Thus, like most theorists in this area (including Williams and Parfit) I use the terms – or, perhaps better, metaphors – ‘give rise to,’ ‘provide,’ and ‘generate’ interchangeably here, and so as to subsume but be broader than ‘constitute.’ As I noted earlier, I am also following standard practice in treating a second class of notions as more-or-less interchangeable: ‘be necessary for,’ ‘depend upon,’ and ‘require the backing of.’ For some purposes, there may be metaphysical distinctions that would naturally be picked out by different expressions in one of these two equivalence classes. But since this is all just loose talk here, they are distinctions without a difference for present purposes.

Let's trace the dialectic here. Parfit's core objection to internalism about reasons, and subsequently his main motivation for his own positive proposal (objectivism about reasons), constitutes alleged counter-examples to internalism in the form of foolish desires. Now, we might take issue with some of Parfit's examples here, since it is not clear that someone's 'desire' to waste her life or to frustrate all her future desires could withstand critical scrutiny, without losing its claim on intelligibility.⁷ Nevertheless, people *do* want the darnedest things, even when they are harboring no illusions. Take Van Gogh, for example: he wanted to cut off a piece of his own ear in a kind of self-dramatizing flounce, at least according to popular legend. It doesn't get much more foolish than that.⁸

But these alleged counter-examples to internalism, in the form of foolish desires, will only cut against the position under the assumption that the only kind of normative force that desires could have is *reason-providing* force. In other words, Parfit is here trading on an additional – ubiquitous, but typically unstated and unargued-for – assumption, viz.:

THE TACIT ASSUMPTION: an agent's reasons could only *depend* on her having certain desires insofar as these desires themselves *provide* these reasons.

Parfit also reveals the assumption behind the assumption, when he prefaces a discussion by singling out “facts that have *normative force or importance*, in the sense that these facts *give us* reasons” (p. 408, his italics).

⁷ Compare Parfit's famous example of 'Future Tuesday Indifference' (1984, pp. 123-124, and forthcoming, p. 53). We are asked to imagine a man who is indifferent to the prospect of agony on future Tuesdays, even though he knows it would be just as horrible then. But I am skeptical that this example meets the minimum constraints on desires (and lack thereof) which are required to render them intelligible. In fact, I think that the intelligibility of bare indifference to extreme pain is even more doubtful than the intelligibility of *bare desires* (i.e., wanting something without wanting it *for* anything), which many theorists doubt. I explore these issues in work in progress.

⁸ Admittedly, one might contest that this desire was sufficiently well-informed. For examples of foolish desires that do meet this test more clearly, consider *recalcitrant* desires. E.g., a perfectly healthy woman may come to believe that “fat is a feminist issue” (to use Susie Orbach's famous phrase), but continue to harbor a desire to be gaunt regardless. For another example, some desires have a calmer expressive or performative basis. Consider Thomas Hill Jr.'s case of the deferential wife (1973). This woman chooses to suffer in silence, in conformity to traditional feminine social norms, rather than asserting herself within her family. Consciousness-raising may have convinced her that there is no real value in keeping mum – so to speak – but she still finds herself ineluctably drawn to playing the role of martyr. Anyway, I agree with Parfit here: there can be desires that survive full information, but are nevertheless really foolish.

Now, notice that the tacit assumption implies that if desires are a *necessary* condition on practical reasons (as Williams-style internalists hold), then they would have to be a *sufficient* condition as well. The assumption is thus what makes internalism and objectivism look like flat-out incompatible positions. Indeed, Parfit simply assumes as much in several key passages in *On What Matters*.⁹

But the incompatibility is illusory, I believe. For despite tending to go unquestioned, the tacit assumption is eminently questionable. Moreover, jettisoning it shows that internalism is not to be so easily demolished.

3: INTERNALISM, REDUX – THE VETO POWER VIEW, TAKE ONE

What gets perpetually overlooked here is that, in claiming that desires are necessary to ground practical reasons, internalists need not rely on the assumption that desires themselves *provide* these reasons. There is another equally straightforward way of going: namely, internalists may hold that the *absence* of a suitable desire on the part of the agent *defeats* reasons that would otherwise hold for him.¹⁰ And this inversion may make all the difference to the plausibility of the position. For we can now deny that the fact that an agent really *wants* to do something stupid – to cut off his own ear, say – would give him a reason to do it. Whereas the fact that

⁹ For example: “According to objectivists, practical reasons all derive their force from the facts that make certain things relevantly good, by giving us reasons to want these things, and to try to achieve them. *According to subjectivists, we have no such object-given reasons.* Some subjectivists even claim that it is we who make things good” (pp. 44-45, my italics). In this passage, Parfit unhesitatingly ascribes to subjectivists (and thus internalists) the *denial* of objectivism, i.e., the claim that *no* reasons are provided by desire-independent values.

¹⁰ Williams’ original statement of his position in the opening of “Internal and External Reasons” (1981) comes interestingly close to this way of putting things. Williams writes that, on the internal interpretation, sentences of the form ‘A has a reason to φ ’ “...imply, very roughly, that A has some motive which will be served or furthered by his φ -ing, and *if this turns out not to be so the sentence is false*”; whereas, on the so-called external interpretation “there is no such condition, and the reason-sentence will not be *falsified by the absence* of an appropriate motive” (1981, p. 101, my italics). He then states the view using a biconditional, but only for illustrative purposes. Later, in an attempt to clear up the misunderstanding this engendered, Williams wrote: “the internalist view of reasons for action is that this formulation provides at least a necessary condition of its being true that A has a reason to φ : A has a reason to φ if he could reach the conclusion to φ by a sound deliberative route from the motivations he already has... It is a further question whether the formulation provides a sufficient condition of an agent’s having a reason to φ ” (1995, pp. 35-36). He went on to say that he *did* think it was probably a sufficient condition as well, but that this was a separate question, and one he wouldn’t take up there. Thus, I think the current proposal is in line with, and might even be taken as a way of filling in, Williams’ original proposal – more so than positions which say that desires provide reasons, which makes it difficult to see why Williams took such pains to disentangle the foregoing necessity and sufficiency claims.

he really *doesn't* want to do something potentially sensible – to increase his income, for example – would nevertheless *undermine* the reason he would otherwise have had to act accordingly. On this novel internalist conception, as I will develop it, an agent's would-be self-interested reason to φ is provided by the facts that would make the outcome objectively valuable, just as Parfit insists. However, the agent also has to want that outcome in order to have an extant self-interested reason to φ , just as Williams insisted. Otherwise, the would-be reason is defeated, and its latent claim is *silenced*.¹¹ Hence, I will christen this 'the veto power view' about the normative force of desires. To summarize (at a first pass) the view on offer by means of slogans:

THE VETO POWER VIEW: Desires are the *sine qua non* rather than the *source* of reasons; desires do not *provide* reasons, they *sanction* them; or, to put it more prosaically, *desires rubber-stamp reasons*.

A fuller and more precise statement of the view, which makes the mirror image claim about aversions and reasons against, will follow in section 6 (after I've discussed aversions in section 5).

I believe that the veto power view is a plausible account of practical reasons, quite generally speaking. But this is a bold claim, so first to clip my own wings. I will limit myself in this chapter to defending the view as applied to self-interested reasons alone. For although the internalist claim as applied to our own well-being is widely regarded as intuitive, it is thought a more unlikely story about moral reasons, at least given certain theoretical commitments. I hasten to say that vulgar moral relativism or subjectivism would be anathema to me as much as anyone. However, I am also inclined to think that the domain of moral assessment is considerably wider than the realm of moral reasons as such. And internalism about moral reasons does tempt me, on grounds I lay out in Chapter Two. But even if internalism about moral reasons would be anathema to you (because, say, you have more rationalist leanings), you could still accept the internalist condition on self-

¹¹ This is John McDowell's notion. In discussing the virtue of courage, McDowell suggests that we "stop assuming that the virtuous person's judgment is a result of *balancing* reasons for and against." Rather, the aspects of the situation that call for courage "[are] apprehended, not as *outweighing* or *overriding* any reasons for acting in other ways, which would otherwise be constituted by other aspects of the situation (the present danger, say), but as *silencing* them." (1998, pp. 55-56, my italics).

interested reasons which I am proposing in this chapter. In fact, there are familiar grounds both intuitive and theoretical for thinking (*pace* Parfit) that self-interested and moral norms differ in terms of their dependence on desires – owing to the special relationship we bear to ourselves, and in which our autonomy or self-governance at least partly consists.

I have already suggested that my version of internalism is better than ‘pure’ internalist or so-called Humean accounts – i.e., views according to which desires are held to provide practical reasons. For on my view, we rule out the whole slew of counter-examples involving foolish desires which Parfit and his objectivist allies often take to be positively fatal to the position. And the veto power view conforms to the spirit as well as the letter of internalism, retaining its central intuitive attractions, as we will see over the course of the chapter.¹² So in order to defend the veto power view, two main argumentative burdens now fall on me. Firstly, I need to show that my view is better than its hybrid predecessors – that is, views which posit both objective and subjective conditions on well-being. Secondly, I need to show that my view improves upon ‘pure’ objectivism by appending the internalist condition to objectivism, which thus earns its theoretical keep. I will take these tasks in order.

4: A CLEANER HYBRID

The veto power view vindicates both internalism and objectivism – with the internalist condition serving to shrink the often teeming set of potential self-interested reasons generated by objective values down to a more manageable size. But insofar as it thus entails a hybrid view about well-being, my view is not unique.¹³ Take

¹² It is interesting to note in this connection that the examples used to support broadly internalist views have a tendency to be negative. Along with Williams’ original Owen Wingrave case, take this nice example of Stephen Darwall’s: “Suppose you are a middle-aged daughter whose parents are trying to get her to eat ‘her broccoli.’ You might well think that the fact that you don’t *want* to eat your broccoli is a reason why you should be allowed not to do so (thank you, very much!), and, indeed, that it is a reason why you should *allow yourself* not to do so” (2003, p. 442).

¹³ On the relationship between self-interested reasons and well-being, I am assuming that an agent has a self-interested reason to φ only if φ -ing is conducive to her well-being. The converse claim is trickier, since there are conditions which enhance our well-being which we have little or no control over (our being loved, for example). But we can say that, if

Richard Kraut's account (1994). According to Kraut, an agent's life will be good insofar as she pursues – or is otherwise appropriately related to – something which she loves, which is worthy of that love. Susan Wolf's proposal in her recent Tanner lectures has a similar shape, but she envisages the two conditions as underwriting a meaningful life, specifically. “[M]eaning in life arises when subjective attraction meets objective attractiveness, and one is able to do something about it or with it,” she writes (2010, p. 26).

I think that these accounts are certainly along the right lines.¹⁴ But the accusation hybrid theorists routinely face is that they are simply hedging their bets. The worry will remain pressing until the need for both the objective and subjective conditions, as opposed to the single condition its opponents favor, has been firmly established – along with the relationship between the two. Translating hybrid accounts of well-being into claims about practical reasons tends to be rather sticky too. Suppose, on the one hand, that we take the liberal line here – i.e., we allow that *either* a suitable value or a suitable desire would provide the agent with a reason to bring about the relevant state of affairs (with some balance of both types of reasons figuring essentially in a good life, perhaps). In that case, as Ruth Chang points out, the proponent of the hybrid view could only avoid by fiat effectively ‘double-counting’ the agent’s reasons to do independently valuable things that she wants to do on this basis (2004, p. 62). On the other hand, suppose we say that the agent’s desires have to be held for the right reasons. The trouble is that, if a desire is both based on and underwrites the reason itself, the proposal smacks of circularity. Joseph Raz (another prominent objectivist) deems the idea that desires strengthen the very reasons on which they are based positively paradoxical; “it requires too much by way of

something is conducive to our well-being and we *can* act so as to help bring it about or maintain it, then we have a self-interested reason to do so. Thus, there is a close connection between the metaethical debate over reasons internalism and the debate about the substantive conditions that make a person’s life go well.

¹⁴ Ironically, Parfit himself briefly expressed sympathy for a hybrid position in the closing lines of his appendix to *Reasons and Persons* (1984). However, he has evidently since changed his mind, perhaps due to his subsequently coming to draw a sharp distinction between *liking* and *desiring*. In any event, he firmly believes nowadays that desires play no essential or important role in the best account of reasons, which figure in turn in his view as the sole fundamental normative entity. Kraut too later abandoned his hybrid picture for an objectivist account, although on very different grounds. His *What is Good and Why: The Ethics of Well-being* (2007) is a book-length argument for Aristotelian perfectionism.

mental gymnastics,” he complains (1999, p. 61).¹⁵ Finally, if *both* the value and the desire are held to be jointly necessary to ground a practical reason, then how could it be that these very different sorts of conditions both make essential but indistinguishable (or at least overlapping) normative contributions? So the hybrid theorist will be driven to cast about for an appropriate metaphysical picture according to which values and desires play distinct but complementary – and ineliminable – normative roles. The veto power view provides just such a picture, in saying that objective values play the role of first laying out the smörgasbord of goods, as it were, whereas desires play the role of then picking and choosing amongst them.

Admittedly, one could go the other way here and say that our desires generate the reasons, and objective goods instead play the role of the *sine qua non* conditions. Or, one could go for a ‘no priority’ view, as has become popular in the literature on causation.¹⁶ But since the veto power view accords a particularly plausible role to the *will* here (our desires having, on my view, the role of picking and choosing from the value-laden smörgasbord), there are already grounds to prefer it. The veto power view also makes better sense than either of its siblings of an important fact about deliberative priority, which is especially evident from the first-person standpoint. We do not deliberate from what we *want*, after all; we deliberate by looking out into the world at what would be valuable for us. This is often billed as the fatal flaw of internalism (see, e.g., Raz, 1999, and Dancy, 2000), but it does not afflict my account. On the veto power view, values do impinge on our deliberations from the outside: but, in order to give us reasons, these values must then engage with our internal motivational mechanisms: i.e., our desires. And, as we have been seeing, this is a substantive further

¹⁵ There is also the weaker idea that the agent’s desires just have to be explicable in terms of the presumed underlying *value* of their objects. While this is a tempting suggestion, I think it fails too. What needs explaining here is not just that we want X, but also that we want X *rather than* Y. But we may want X rather than Y not because we take it to be more valuable for us, but merely because of admitted happenstance. For instance, many happily partnered but unsentimental people will admit that we might, in some nearby possible world, have also been happy with unknown persons who are now, effectively, *personae non gratae*. But having had the good fortune of meeting and falling in love with our *actual* spouse, these alternative merely possible spouses cease to attract us in the least. At least, so it might be hoped (by our spouses, for example). Accordingly, the undoubted value of the love object falls short of explaining what needs to be explained here – namely, the *choices* which structure our lives.

¹⁶ That is, some theorists argue that the attempt to differentiate causes and *sine qua non* conditions is misguided, and that this is largely a context-sensitive matter.

constraint. For we can recognize that something would be valuable for us without automatically coming to want it – due, for example, to our recognition that we have an incompatible project or commitment, and only one life to lead.¹⁷ So I think that my view is considerably more promising than certain alternatives in the same family.

I have argued that the metaphysical details (and detail) of my proposal defuses characteristic worries about the messiness and/or vagueness of hybrid theories of well-being: a significant theoretical boon. But a theoretical boon is also an explanatory burden. For, you might well ask, how could values and desires be related to each other in this push-pull way? In fact, the distinction within normative metaphysics which I am drawing, i.e., between *reasons* (facts which provide normative force), and *necessary preconditions* (facts whose absence would neutralize the potential reason-providing force of such facts), has a long and respectable history. It is positively crucial in epistemology, for starters. Consider, as a representative example, the fact that a reliable barometer indicates that it is going to rain this afternoon would be an epistemic reason for me to believe that it will rain, but for the fact that I don't believe the barometer *is* reliable. Perhaps I have misleading evidence that it is broken, or simply lack any evidence about its accuracy. Alternatively, I may not know how to interpret it. And without this internal constraint on what externally valid (e.g., reliable) sources of information can do for me epistemically, there would be nothing to stop me from reasoning as follows about a reliable barometer *of* whose reliability I am completely unaware: "This barometer says it is going to rain and (so) it is going to rain. So I now have some evidence that the barometer is reliable." This kind of inductive procedure is known as bootstrapping (see, e.g., Vogel, 2000; Cohen, 2002). And despite the reliability of the

¹⁷ Hence, the veto power view is well-placed to account for the importance of personal projects, an important theme in Williams' philosophy – for which his central example was Van Gogh's friend, Paul Gauguin (who Williams has us imagine abandoning his family, in order to dedicate himself to painting: "Moral Luck," in 1981, pp. 22-23). I'll leave the original moral issues to one side here. But it does seem very plausible to think that, if someone finds himself totally dedicated to a certain pursuit, then he may find himself completely *indifferent* to the allure of competing alternatives. Judging by his letter, Van Gogh was a case in point. The veto power view says that such agents may have unopposed self-interested reasons to pursue their project – as long as what they do want would be independently good for them to get.

reasoning involved, it seems clearly circular; we are tempted to write it off immediately. Indeed, bootstrapping is an epistemological disaster for any theory of justification that sanctions it (or so say nine-out-of-ten epistemologists, by my off-the-cuff estimations). However, we would hardly say in the alternative that my trusting the reliable barometer is itself a reason to believe that it is going to rain this afternoon. After all, my trust will persist during fine weather too.

The point is very general; on all but the most mad-dog externalist epistemologies, a potential source of information can only be a source of information *for me* if I take it to be a trustworthy, reliable, and informative source. Admittedly, spelling out just what this trust amounts to, and thus when exactly its absence comprises a defeating condition (to now express the point negatively), is by no means easy. And whether there are exceptions to the rule for basic sources of evidence (e.g., perception) is highly controversial. But for my purposes, it does not much matter. Considering the disparities between the constitutive aims governing theoretical and practical reasoning, we would not expect the loose analogy I'm drawing between *wanting* and *trusting* to be perfect in any case. And we should not forget the suggestive historical fact that epistemologists originally borrowed *their* notion of a defeating condition from a distinctly practical normative domain – that of the law.¹⁸ For, as H.L.A. Hart emphasized (1961), the fact that (e.g.) Smith signed a contract requiring him to pay Jones a certain sum of money would be a legal reason for him to do so, but for the fact that Smith lacks legal standing to enter into this contract. For example, he is a minor, or he is mentally incompetent. The neologism 'defeater' was thus originally due to Hart, who bemoaned the lack of an English translation of the Latin *conditio sine qua non*.

¹⁸ They also altered it a bit. The usual distinction between *rebutting* defeaters (countervailing evidence) and *undercutting* defeaters (evidence, or the absence thereof, which undermines the force of certain putative evidence) is arguably misleading; the former aren't aptly construed as defeating or *sine qua non* conditions at all, they are just further pieces of evidence. In any case, the veto power view is roughly analogous to placing a *no undercutting defeaters* constraint on justification or warrant as a necessary condition thereof. Tellingly, a general 'no defeaters' constraint of this kind is sometimes taken to be the main criterion of internalism in epistemology (see Bergmann, 1997).

Despite the general inattention in ethics to this distinctive kind of normative force (with the notable exception of Jonathan Dancy's work), it would in fact be rather surprising if there were no such thing as defeaters for practical reasons.¹⁹ After all, defeaters are ubiquitous in systems of rules, and rules are often thought to give rise to practical reasons for agents who are playing the game (so to speak). So desires may, similarly, play this reason-sanctioning role. I will go on to argue in section 7 that it would explain a great deal if they did.

But first, we must clarify what the defeating condition for reasons is supposed to amount to on my view, exactly: in other words, what it means *not to want something* – as distinct from *wanting the opposite*, which is stronger. In the process, I will say something about the composite states of *indifference* as well as *ambivalence*, which will be important in what follows too.

5: THE OPPOSITE OF DESIRE

Many philosophers (including, at times, both Williams and Parfit) represent the opposite of a desire for a state of affairs to occur as a desire for it not to happen. While roughly right, this is an unfortunate mode of expression. For one thing, it is easily confused with merely *not wanting* the event to happen, which (as we will see) would be a disastrous, if common, conflation. And even if we are careful to mind this distinction, the expression is quite hard to parse. Due to the promiscuity of the negation operator (i.e., the often long, sometimes open-ended list of ways in which events can fail to occur) the intended meaning will frequently get

¹⁹ Dancy explicitly argues for the applicability of the concepts of enabling/disabling to practical reasons. His *Unprincipled Ethics* (2004a) contains an edifying chapter on the subject (Chapter 3), in which I encountered the line from Plato with which I opened. See also Raz on qualifying/disqualifying reasons (1999). But, as far as I am aware, no one has proposed applying these defeater-type notions to desires. Indeed, it is interesting that Dancy and Raz both take considerable pains to deny that desires generate reasons, even though this does not exhaust the ways in which desires could have pervasive normative force, by the lights of their own distinctions. On the other side of the theoretical divide, Mark Schroeder (2007) has defended a version of the Humean view according to which desires always generate reasons, but the reason is *constituted* not by the desire itself, but rather by its object (i.e., by *what* we desire). But although this metaphysical distinction is a relative of mine, Schroeder invokes it in order to defend one of the competing views which I am trying to broker a détente between.

lost. And it just seems inapt to construe someone's being against an occurrence as their being *for* or in favor of its *non*-occurrence.²⁰ Compare the suggestion that we construe my disliking skiing as my liking not-skiing. Unless there is some clear, contextually delineated alternative to skiing, e.g., lounging in the chalet drinking cocoa, this statement means next to nothing. And even if the alternative is plain, it would be a very backwards way of expressing oneself – which is why it sounds close to a joke.

Perhaps in light of this difficulty, Parfit typically eschews the unpromising “wanting *not* that state of affairs” locution in favor of “wanting to *avoid* that state of affairs.” This is fine as shorthand, and I will often make use of the expression myself. But just as we should be cautious about assuming that desires are always future-oriented (I can desire that you agree with me now²¹), there are analogous grounds to take wanting to *put an end* to something as a motivational state in the same family as wanting to avoid or prevent it.²² So it will be convenient to have a general label for the motivating state that is the opposite of desire: with apologies to Elie Wiesel, I will label it *aversion*. But since aversions can be cashed out in terms of wanting, they are not a new ontological outlay in the economy of mental life.²³ By the same logic, the usual distinctions which apply to desires will apply, *mutatis mutandis*, to aversions. For example, I will assume that, as with desires, we can be *intrinsically* or *extrinsically* (e.g., instrumentally) averse to a state of affairs, i.e., for its own sake, or because of its

²⁰ The obvious exception is for states of affairs where the contrary and contradictory concepts are one and the same (e.g., for potentially living things, being dead is the same thing as not being alive).

²¹ For an edifying discussion of desires which appear to be present-oriented, see Lloyd Humberstone on the example of wanting to stay in the bath (1990, pp. 108-109).

²² I think we should also be cautious about conflating *expectations of liking* with desire in the motivational sense in which I use the term throughout this thesis (following both Williams and Parfit, p. 43). Take Parfit's own example (which appears in a different context) of wanting life to go well for a hitherto stranger one met briefly on a train (1984, p. 494). Assuming that one knows one cannot do anything to bring this about, this will not typically be a *motivating* desire, I suggest. Yearnings in this category may more aptly (and variously) be called hopes, wishes, fantasies, and daydreams. Parfit does not make this distinction, however.

²³ Note that I will now cease to use the terms ‘desiring’ and ‘wanting’ interchangeably, as I've been doing so far, since the expression ‘desiring to avoid’ would only serve to muddy the distinctions I've just put in place.

relation to other states of affairs.²⁴ And I take it that – as with reasons – desires and aversions can be either *self-interested* or *altruistic* in nature. They can also be both: desires for the success, or aversions to the failure, of a team or collective of which you are a part are a classic case in point. Nor should intrinsic desires be assumed to be invariably self-interested. We may desire that a friend cease to suffer for its own sake, and for *her* sake, not for ours.

As well as the individual motivating states of desire and aversion, there is also the composite state of *indifference* to consider. For in light of the preceding distinctions, it is obvious that there are countless events which I am *neither* for nor against (no apologies to G.W. Bush). If we understand desire and aversion as the *presence* of a positive and negative motivational state (respectively), then indifference is best thought of as the *absence* of either state. We often express indifference by saying that we don't mind, or don't care, one way or the other. We also typically lack an aversion to events which we desire – call this *pure desire* – and, similarly, lack a desire for events which we are averse to – call this *pure aversion*. However, the composite state of *ambivalence*, i.e., desiring and being averse to the same state of affairs, is not uncommon either. I will say a bit more about ambivalence in a moment.

²⁴ The same distinctions apply to reasons. E.g., we can have an instrumental reason to do something, or a reason to act which is self-standing (i.e., intrinsic).

But first, to lay things out perspicuously here, let S name some state of affairs. The four possible individual motivational states (described using both modes of expression) are as follows:

- D**... desire for S ... wanting to bring about/maintain S
- ¬A**... lack of aversion to S ... not wanting to prevent/end S
- ¬D**... lack of desire for S ... not wanting to bring about/maintain S
- A**... aversion to S ... wanting to prevent/end S

These possibilities may combine to yield four possible overall motivational configurations, like so:

- D & ¬A**... pure desire for S ... desire for and lack of aversion to S
- ¬D & ¬A**... indifference to S ... lack of desire for and lack of aversion to S
- D & A**... ambivalence to S ... desire for and aversion to S
- ¬D & A**... pure aversion to S ... aversion to and lack of desire for S

I think that it is crucial, in discussions about the normative significance of desires, to mark these distinctions clearly.²⁵ But then why is there a tendency, even within philosophy, to slip and slide between them? I suggest it is because these distinctions tend to be highly sensitive to – and hammered out in – context. Suppose someone asks you, “Do you want to go skiing?” If this question is understood as a hopeful and welcome invitation, then a flatly negative response will ordinarily count as a refusal, and carries a fairly strong implicature that you are actually averse to the idea.²⁶ On the other hand, if skiing is merely mooted as one possibility among others, then the previous implicature of a negative reply, i.e., that you are actually averse to going skiing, is weakened considerably. And if the question is just a disinterested inquiry into your preferences (imagine that the inquirer can’t go themselves, and is just helping you think through your options) then it is

²⁵ In particular, if we only distinguish between ‘wanting S ,’ ‘not wanting S ,’ and ‘wanting *not* S ,’ as is customary, then we cannot express *lack* of aversion – nor (in consequence) indifference. We could add ‘not wanting *not* S ’ to the list, but the expression is unwieldy, since it could be read as a double negative. But if double negation elimination is performed, then indifference – ‘not wanting S and not wanting *not* S ’ – will be tantamount to ‘not wanting S and wanting S ’: an explicit contradiction. So it is much clearer to trade in ‘wanting *not* S ’ for ‘wanting to prevent (or end) S .’

²⁶ This is a phenomenon common among litotes generally, of so-called ‘negative raising’ – e.g., in many contexts, saying “I don’t like skiing” has the implicature “I *dislike* skiing.”

virtually cancelled. It wouldn't be strange at all to shrug in response and say: "No, I don't want to go skiing... I mean, I'm not *averse* to the idea, but there are certainly things I'd *rather* do." Admittedly, the verbal form of aversion can sound a bit archaic on its own. However, when considerations of politeness are not paramount, there are many sure-fire ways of conveying aversions via a combination of gestures and hyperbole – by responding to questions like the above by rolling one's eyes, miming a shudder, and saying things like "There's nothing I want to do *less*" and "God no, I'd rather die" (remember being a teenager?).

A few words about ambivalence are in order before we continue. It is plausible to suppose in the first instance that ambivalence is always soluble in principle, because the conflicting motivations in play must owe to separate and separable features of some composite event (e.g., in the case of Van Gogh's ear, drama without the pain might have served just as well). But besides ambivalence proper, there are also mere *conflicts* – where we want several incompatible events – and *dilemmas* – where we are averse to several events, one or more of which is inevitable. And on one way of thinking about things, conflicts and dilemmas are virtually the norm. Suppose I have two potentially good choices, ranked in order of preference. For example, I would prefer the strawberry shortcake to the apple pie for dessert; I would enjoy both, but I want only one. It is still natural (although not compulsory) to say that I have a weaker motivating desire for the apple pie, which will kick in if the kitchen runs short on shortcakes. This would be a conflict. It could also be that, even though I don't want both desserts (having just enough room for one, say), I *would* want both *ceteris paribus*, and so face a contingent dilemma – between getting too full, and missing out on something delicious (these both being outcomes I'm averse to, we're here imagining). But as we will soon see, some putative conflicts and dilemmas are not contingent in this way: they appear to be insoluble even in principle.

6: THE VETO POWER VIEW, TAKE TWO

With the foregoing distinctions in place, I am now in a position to issue the canonical statement of the veto power view about the normative impact of desires on self-interested reasons:

An agent's *lacking an intrinsic, self-interested desire* for an outcome *defeats* the intrinsic, self-interested reason for her to bring it about or maintain it; and, likewise:

An agent's *lacking an intrinsic, self-interested aversion* to an outcome *defeats* the intrinsic, self-interested reason for her to prevent or to end it.

This is obviously not a comprehensive view about self-interested reasons. The crucial unanswered question – what is objectively valuable, and hence reason-providing? – is left hanging quite deliberately though, since the idea is that the view be hitched to one's preferred account of objective value. Following Parfit, though, I think we can assume that pleasures and pains have a strong claim to inclusion on the list of objective goods and ills (respectively), especially if we are careful not to construe these mental states in an unduly narrow or overly sensationalist way, so to speak. So I will focus on these clauses as applied to hedonic-cum-affective states hereafter.²⁷ These clauses then jointly ensure that an agent can have no reason either to bring about, or to prevent, episodes of (say) joy or suffering about which he is genuinely *indifferent*. Hence, the veto power view will back Williams' denial that Van Gogh had self-interested reasons to pursue those would-be goods, or to avoid those would-be ills, about which he simply didn't care. But notice that these clauses allow that he *could* still have had these reasons if he had instead been *ambivalent* – as he might well have been in the case of

²⁷ It admittedly remains to be seen whether lack of desire for such putative goods as knowledge, the admiration of beauty, and love could plausibly be held to defeat the self-interested reasons to pursue them. I tend to think so, but I am also not convinced that we will feel the *need* to make additions to the list of objective goods once we have the power to make the appropriate subtractions that the veto power view furnishes us with (as we'll see in the next section). But I won't try to defend these claims here. It is quite ambitious enough for now to argue that the above clauses hold for hedonic-cum-affective goods and ills. I am also open to the possibility that the preceding clauses may need embroidering. In particular, if we allow that reasons come in different strengths, then we may want to allow that a disproportionately weak desire can *diminish* a reason's force, as well as for its total absence to *defeat* it. But I'll leave such possibilities aside here.

cutting off his own ear. That is, suppose we assume (as seems very natural) that he had an aversion to the pain involved, in addition to the rogue desire to do it. In that case, the veto power view leaves his self-interested reason *not* to cut off his own ear perfectly intact. The same line will apply to a large portion of the cases where people knowingly invite some painful ordeal. After all, political protesters and cosmetic surgery patients alike typically want to minimize or anaesthetize themselves to the pain to whatever extent they can, given the aim of the exercise.²⁸

Why are the preceding clauses restricted to intrinsic, self-interested desires and reasons? When there is no risk of confusion, I will often drop the caveats to save words. But the following (admittedly toy) example will serve to illustrate why it can matter, as well as to further elucidate the contours of the veto power view. Suppose soft-hearted Sarah has no intrinsic, self-interested desire to give petulant Peter her cupcake. She might still want to give it to him on grounds that are solely *instrumental* (e.g., she wants to win him over) or *altruistic* (e.g., she wants him to enjoy it). In that case, the satisfaction of her instrumental or altruistic desires should not be held to contribute to *her* well-being, even by the lights of purely internalist accounts. Sarah will

²⁸ In still other cases, e.g., some physical challenges, pain is part of the *point* of the whole exercise. But I take it to be very intuitive to hold that (say) the marathon-runner has no self-interested reason to avoid the pain that she willingly and unambivalently embraces, out of dedication to the sport – even though she *would* otherwise have self-interested reasons to avoid the associated aching lungs and burning muscles. There is also an important class of cases where the person is not really *feeling* much of anything, in which case it doesn't count as pain at all in the phenomenologically loaded sense of the term I intend here (following Parfit, p. 51). I'm thinking primarily of the 'anhedonic' tendencies exhibited by some depressives and dissociatives – along with practitioners of certain advanced Buddhist meditative techniques, and people who suffer from a rare congenital insensitivity to pain. In these cases, Parfit and I will be in agreement that the person may have no self-interested reason to avoid the phenomenologically neutered form of 'pain,' assuming no lasting bodily damage that would eventually catch up to them, making their lives go worse. On the other hand, people who do not seem to care about their health in the rather more jubilant sense of being *bon vivants* who smoke, eat with abandon, and never see the inside of a gym do not strike me as being aptly criticizable. As long as they are not indulging in fantasies of receiving their centennial telegram from the Queen, then I am inclined to take a pretty permissive view here. "It's up to them," we'll find ourselves saying, so long as they are not harming anybody *else*. Thus, I take my account of self-interested reasons to deliver (or at least to be consistent with) our intuitive verdicts in an impressive range of real-life cases. As for highly idealized cases, like Parfit's example of a man who is inexplicably indifferent to agony on future Tuesdays, I have serious doubts that such cases are actually intelligible (see *n.* 7). For an interesting point of comparison, see Street (2009), in which she argues that the weight of intuitions are still on the internalist's side, once even this case is properly filled in.

be no better off, having handed it over, if Peter continues to dislike her – or the cupcake. On the contrary, she will then have given up a cupcake for nothing.

The restrictions in these clauses ensure that the veto power view will respect the same distinction. For suppose that, on top of this, Sarah would actually rather enjoy Peter's likely reaction upon her magnanimously presenting him with the gift of the cupcake: i.e., visible embarrassment at his having previously been such a grump. (As the saying goes, we never forgive those who we've wronged.) Despite knowing herself full well, however, suppose Sarah wants to be a bigger person than this (giggling inwardly at Peter's discomfort wouldn't do anyone any harm, exactly, but nor would it be particularly classy). Consequently, Sarah may not want this ignoble thrill of moralistic triumph or *schadenfreude*. The restrictions in the above clauses ensure that Sarah's instrumental or altruistic desires would not interfere with the fact that her lacking the intrinsic, self-interested desire for this ignoble pleasure would on my view act as a defeater for her reason to indulge it. Moreover, the ignoble pleasure she indeed felt 'despite herself' (as we often say) upon watching Peter squirm would not then contribute on my view to her well-being – it being the kind of pleasure that she does not care to have.

Suppose too that, upon feeling this spiteful or aggrieved satisfaction, Sarah was to feel a slight pang of shame. She nevertheless might not want to avoid the associated discomfort – whose undoubted unpleasantness would not then, on my view, detract from her well-being. In other words, the veto power view denies that feeling bad (e.g., guilty) is inevitably bad *for* us – when we see it as deserved, for example, and so wouldn't want to just shrug it off. And although this might be a bit of an overreaction in the cupcake case, there are plenty of other cases of guilt, shame, or remorse where we would take a pretty dim view of someone who would choose to simply rid herself of this negative emotion, given the opportunity (e.g., if offered a special guilt-blocking pill). Tellingly, we might then query it *as* a genuine instance of guilt, shame, or remorse.

7: INTERNALISM, REGAINED

I will now shift gears from elucidation back to argumentation, i.e., to the remaining task of urging the veto power view over its purely objectivist rivals. Recall that the view I propose differs from pure objectivism in saying that unwanted ‘goods’ are not genuinely good for us (i.e., beneficial), and that ‘ills’ we are not averse to are not actually bad for us (i.e., harmful). We have already seen several examples which suggest that this proposal resonates with our sense of what counts as a good life, from the perspective of the person living it – and the many other examples to be considered in this section will serve to further drive home the point. But those who follow Parfit in espousing pure objectivism could try to block my proposal by saying one of three things. They could say, one, that – contrary to the hypothesis which Parfit and I share – agents inevitably *do* want every potential pleasure and to avoid all instances of pain; two, that even if they don’t want this pleasure (or mind this pain), these hedonic states *are* intuitively beneficial (harmful) regardless; or, three, that certain pleasures (pains) to which we will often be indifferent are *never* beneficial (harmful), even when they are the object of a desire (aversion). All three of these lines are attempts to short-circuit my claim that not wanting a certain pleasure, or not minding a certain episode of suffering, prevents it from benefitting or harming you (respectively). I will now consider these three preemptory responses in turn, and show by means of counter-examples that none of them can be made to work. After that, I will be back on the offensive on behalf of my position.

Firstly, the pure objectivist could try to insist that there is no such thing as an undesired pleasure, or suffering we don’t mind. (This would be to deny that the gap the veto power view posits between pleasures and desires on the one hand, and suffering and aversion on the other, so much as exists.) Taking this line would involve saying, in cases like Sarah’s, that she actually *does* want the feeling of moralistic triumph, *ceteris paribus*; she just wants to be a good person even more. This is rather strained, given that this is hardly a contingent conflict: so the cash value of the *ceteris paribus* clause is minimal at best. Moreover, there are some kinds of pleasure where

our being reluctant to, i.e., averse to, experience it is constitutive of its being the kind of pleasure it *is* – indeed, it is a familiar fact of life that sometimes the greater our opposition, the more the pleasure grows. This notion of an illicit thrill is implicated in a perhaps worrying large family of concepts, which includes stickybeaking, rubbernecking, rebelliousness, titillation, exhibitionism, voyeurism, and arguably sadism. But despite the pleasure to be found in thereby subverting or perverting one’s own self-critical aversions, I think it would be mistaken to be too pessimistic about human nature in consequence. For it seems to me clear that sometimes a person’s deep-seated or thoroughgoing aversion towards such a taboo pleasure (or to deriving it from this source) leads him to have *no* desire to indulge in or acquire a taste for it, or even to the waning or cessation of his previous desire to indulge. Conversely, suppose a theorist were to deny that we are so much as capable of being turned wholly *off* (motivationally speaking) as well as *on* (hedonically speaking) by the fact that the fruit is forbidden (i.e., by our being averse to this pleasure). They would thereby incur a serious theoretical cost: namely, a commitment to positing insoluble ambivalence on the part of everyone about every illicit thrill in the offing. But does it seem plausible to suggest that someone who remains capable of getting a genuine thrill from (say) voyeuristic pursuits can never stop wanting to look, even on account of their moral scruples? This cynicism seems unwarranted, and threatens to collapse important distinctions between grades of peeping Toms that moral psychology should respect.

The pure objectivist could also try to insist, secondly, that even genuinely undesired pleasures are always intuitively good for us, and even pain we don’t mind is always intuitively bad. (The idea here would be that the veto power view makes a substantive claim, but one that is not even *prima facie* plausible.) But as we’ve been seeing, this line seems false to the phenomenology. This is brought out even more clearly by imagining experiencing pleasure against one’s will. These admittedly unusual occurrences need not owe to external

coercion, which would tend to muddy the waters.²⁹ For just as our wholehearted enjoyment of some activity we began only reluctantly can reveal to us that we really *did* want to engage in it (perhaps we were just being lazy), we sometimes fail to grasp that we *don't* want to do something until it's already too late. A particularly vivid example of this kind is to be found in Milan Kundera's novel *The Unbearable Lightness of Being*. In the following scene, we watch Tereza enact her plan to betray her philandering husband by seducing a stranger ('the engineer'), with psychologically disastrous results:

Peering into the engineer's face, she realized that she would never allow her body... to take pleasure in the embrace of someone she neither knew nor wished to know. She was filled with an intoxicating hatred. She collected a gob of saliva to spit in the stranger's face. He was observing her with as much eagerness as she him, and noting her rage, he quickened the pace of his movements on her body. Tereza could feel orgasm advancing from afar, and shouted "No, no, no!" to resist it, but resisted, constrained, and deprived of an outlet, the ecstasy lingered all the longer in her body, flowing through her veins like a shot of morphine. She thrashed in his arms, swung her fists in the air, and spat in his face. (Harper Collins, 1999, pp. 155-156)

Apart perhaps from the weirdly joyless mirth of tickling, a better example of pleasure (purely) opposed yet augmented precisely because it *is* opposed could evidently hardly be wished for. And the veto power view not only vindicates our sense that such pleasures are not worth having, but even makes room for the possibility that they come to us sometimes as an unadulterated harm.

The internalist addendum I've proposed is thus recommended partly to inoculate objectivist theories of well-being against a problem with hedonism that would otherwise infect them, and which is both old and deep. Consider Plato's contented oyster, Mill's happy swine, and Moore's example of a man perpetually indulging in bestiality. (See the *Philebus*, 21c, *Utilitarianism*, Chapter 2, and *Principia Ethica*, Chapter 3.³⁰) One point these

²⁹ Although, it is worth asking why we would care so much about consent if pleasure and pain were what really mattered, and a person's consent was just a good *indication* of what she would enjoy or find unpleasant. We can imagine various replies on behalf of the pure objectivist, but I am not sure how convincing they would be.

³⁰ Mill's example here is originally due to Thomas Carlyle, who famously held that Jeremy Bentham's doctrine of hedonism was 'pig philosophy,' "a doctrine worthy only of swine." I should also own that, given that the term 'bestiality' was sometimes used in a more general sense (to denote crass behavior) until quite recently, Moore may not have envisaged the man's pursuits in quite such vivid detail as a modern ear may incline us to.

examples might be taken to make is that not *only* pleasure is good. Perhaps Moore's character should go to visit his grandmother, or get to the opera, once in a while too – although our suggestion to the oyster to follow suit may fall rather flat by comparison. We may thus be moved to follow Mill here in distinguishing between higher and lower pleasures, or to go the whole hog in adding other objective goods and ills to the hedonist's hitherto brief list.

But these additions would do nothing to address the arguably deeper lesson to be learned here: namely, that not *all* pleasures are good for us, nor are all pains bad, even in the intended *pro tanto* sense. We will feel the need to make *subtractions from* as well as *additions to* the list, at least when the list is meant to catalog the interests of fellow human beings. For merely sentient creatures, pain *is* surely always to some extent bad for them, and pleasure most likely always to some extent good. But a more complex story needs to be told about sapient animals like us, whose very reasons-responsiveness (*inter alia*, perhaps) issues in our capacity to take complex critical and stoical attitudes to our own pleasures and pains. Moreover, and as we have already seen, the veto power view allows us to vindicate this deep thought on the cheap. We do not thereby lose our claim on the underlying objectivity of what is valuable, and hence is in people's interests. Nor do we have to imperiously deny that different people's interests lie at different points on a continuum with that of animals, whose interests may well be specifiable in hedonic terms alone.

This brings us neatly to the third and final tack the pure objectivist could take in fending off the internalist supplement to objectivism which the veto power view recommends. That is, they could diverge from Parfit by saying that at least some of the hedonic considerations which I have suggested are *sometimes* silenced were mute from the beginning. In other words, at least some of the pleasures and pains to which people are often indifferent are *never* good or bad for us, respectively, even when we *do* want to experience or eschew them. This might be dubbed a 'manual deletion' option, insofar as we remove states like *immoral pleasures* from being on anyone's objective list of benefits, or *altruistic suffering* from being even a potential objective ill. But I think

that this way of deflecting the veto power view on the part of the pure objectivist would actually be the worst way yet. The subsequent implication that wicked and sadistic people can never genuinely advance their own interests by exploiting others, and that decent people can never be made genuinely worse off by the altruistic self-sacrifices they make only reluctantly, is not only insulting but incredible.³¹ In a godless world, why should this be so? The principle that man cannot profit by his own wrong may be a principle of justice, but the world is not a fair place. And if this is not leveled as a metaphysical claim, then it is hard to see how it could be a serious one – as opposed to a piece of wishful thinking, or an imperative about how to use our words.

I am similarly skeptical of the related claim that immoral pleasures aren't *real* pleasures, and noble suffering isn't *genuine* suffering. The implication would be that people whose motivations are powerfully affected by the richest emotional capacities like love and loyalty will actually cease to *like* the 'pleasure' from which they may therefore choose to abstain, out of decency – or to *dislike* the 'suffering' which they may therefore resolutely embrace, out of grit. But this characterization seems high-minded and unrealistic, at least for we mere mortals. I suppose true saints might be like this, although the claim that they become effectively anhedonic – affectively impervious to the relevant bodily and psychological sensations – would make their very saintliness a bit of a puzzle. We might then prefer to call them extraordinarily self-disciplined, like a yogi, or admirably scrupulous, like the "deadly insensible" curmudgeon of the Kantian imagination (1996, §4:398). Whatever the case, decency and grit are not tantamount to sainthood; they are the sorts of traits which ordinary people can strive for, and become better people for having tried. And given that at least some of our dispositions towards pleasure and pain are more or less monolithic facts, rooted in our core biology, the move in question is not only independently implausible, but endangers decency and grit as fundamental but perfectly ordinary character ideals.

³¹ Virtue ethicists who identify flourishing as the *telos* of virtue need not, and for the most part do not, make the incredible claim. They need only claim that there is a *general* self-interested reason to live up to a certain moral ideal, since it represents an attractive, distinctive kind of human excellence – just as there might be a general moral reason to obey the law, without its being implied that individual moral and legal reasons can never come into conflict.

As an illustration to tie things together, picture a loving husband, caring attentively for his ailing wife. We can imagine him feeling great sadness at her suffering, and fearing her further deterioration. He might also experience pangs of unreasonable anger or even physical disgust – for, in illness, the vulnerabilities and ignominies of the body are made all too plain. He is also likely to be exhausted. Nevertheless, we can imagine him having no desire to avoid these emotional and physical discomforts by distancing himself from his wife during her most vulnerable hour: *for she is his wife*, whom he loves (echoing here a famous phrase of Williams³²: see his “Persons, Character, and Morality,” in 1981, p. 18). Were he an articulate man, he might say by way of further explanation – not a justification: none is required – that she is the person who has his deepest love and loyalty, and that providing such unflinching care is partly constitutive of commitments of this kind.

Given the good husband’s way of seeing his situation, I imagine nobody will feel quite right in saying that he still has self-interested reasons to distance himself from his wife so as to reduce his empathetic distress and physical discomfort – reasons which might be *outweighed* but could never actually be *silenced* by the stronger countervailing moral reasons to stay loyally by her side. But, as we have seen, this is what pure objectivists like Parfit will be inexorably pushed to say.³²

There is one remaining road for the pure objectivist to go down, and I believe (from conversation) that it is the route that Parfit himself would be inclined to take. They can bite the bullet and say that, although self-interested reasons cannot be silenced by our lack of desire for certain sorts of potential goods, or lack of aversion to certain sorts of potential ills, this isn’t as problematic as I have been hitherto suggesting. The idea would be that there are (invariably or at least typically) strong *countervailing* reasons in cases like these. Hence,

³² Nor could they cite the husband’s *sensitivity* to these countervailing moral reasons as the explanation for his not having these self-interested reasons to withdraw. If this sensitivity was construed on conative (i.e., desire-like) lines, then this would be capitulating to the internalist agenda. But, if his sensitivity is instead construed along cognitive (i.e., belief-like) lines, then the pure objectivist would run afoul of the stipulation that we are dealing here in reasons which hold (recall from section 2) regardless of what the agent does or does not believe. And, as I have said, the idea that the husband is just *impervious* to his standing reason to distance himself from his wife – with the implication being that he is therefore less than fully cognizant of his true interests here – strikes me as being irredeemably off-key.

we mistake a reason's falling silent for it being effectively drowned out. I'll now consider this defensive maneuver, and argue that – as well as being rather weak – it rests on a patent falsehood. So the bullet turns out to be fatal to pure objectivism, I'll eventually suggest.

For, as the example of the caring husband can be readily filled in to illustrate, the objectivist cannot bank here on there *being* stronger countervailing reasons in cases of this kind: indeed, there might well be none such. For example, this man's wife might be too ill to even notice the difference if he detached himself a little in consoling and caring for her. If so, most of will feel uneasy (at a minimum) in insisting that he nevertheless has strong moral reasons not to do so – it's not doing her any good, seemingly, and it might not even matter *to* her that he stay so close at this time. Hence, if the husband had a slightly different conception of his role, borne out by different desires, I think he might well have an *unopposed* reason to conserve his energy whenever he could, when to do so would be no betrayal. But if he genuinely doesn't *want* to minimize his own distress by distancing himself from the situation, then we are loath to advise him to do so. And it's not just that we will respect his decision in the nominal sense of not intervening; we will take it to be *up to him* in a much more thoroughgoing sense. For another example in this vein, Van Gogh's profound disinterest in bourgeois comforts may have been personal and particular rather than political and general in its basis. *Not for me*, we can imagine him thinking about worldly success, yet never so much as supposing there to be anything the matter with fame, or any good reason not to go after it. Thus, must our genuine indifference to some potential good or ill inevitably reflect atomistic countervailing *reasons*, as opposed to more holistic ideals of ours, or even visions *of* our unfolding lives? Insisting as much would be unduly dogmatic, I suggest, in view of cases like the preceding. Yet our supposed self-interested reasons to pursue these undesired would-be goods,

and to avoid these embraced would-be ills, seem to fall intuitively flat even so – i.e., even if there aren't any reasons running in the opposite direction.³³

The point is bolstered by considering *ignoble pleasures*, e.g., guilty pleasures, hollow pleasures, intoxicated pleasures, as well as the simulated pleasures of the sort epitomized by Robert Nozick's experience machine (1974, pp. 42-45). Such enjoyments may be certified morally benign, and are typically but by no means invariably imprudent to indulge in. But even if there is no prospect of a hangover, an addiction, or any other negative long-term consequences, I might well be indifferent or even purely averse to having such experiences. This will be, in all likelihood, because the pleasure strikes me as somehow sordid or vaguely distasteful, and I don't want to indulge myself thus. (At least, not at the moment; whereas, at other times, a guilty pleasure is *exactly* what I want. In which case, I would be reluctant, as well as hypocritical, to deny the self-interested reason to watch a trashy TV show.) So again, the pure objectivist's proposed explanation here – that we confuse a reason's being silenced with its instead being *outweighed* – fails to accommodate an important class of cases, in which there are no apparent counterweights of the required kind. But the implication that I could be flouting the dictates of practical reason in (say) declining a harmless but unwanted hit of morphine surely strains credibility: indeed, it borders on absurd. “*I just don't want to!*” I want to say, and that should be the end of it.

We sometimes decline to engage in perfectly innocent enjoyments without even these more holistic or broadly aesthetic bases for electing to decline. I suspect that we do need to be able to say *something* to render

³³ An exception may be ‘adaptive preferences,’ a concept we owe to Amartya Sen (see, e.g., 1999) and Martha Nussbaum (see, e.g., 2001), wherein someone comes to have fewer or more modest desires due to her oppressive circumstances. But, on closer inspection, I think that the veto power view says the right thing about these preferences too, allowing us to explain the distinctiveness and insidiousness of the oppressive social forces which they are adaptations *to*. For we must not underestimate the damage that such forces can wreak on the oppressed, thus sometimes (temporarily or permanently) *altering* their interests, as well as merely deceiving them as to where their true interests lie. Picture Mill's hothouse flower (1869), now withering in the ground: it may no longer be capable of thriving in the very soil in which its seedling not only would have blossomed, but would have put down deeper roots. If so, then we must minister to its present needs, even as we try to restore function. But we can still reprimand the gardener for perverting or stunting its growth.

our refusal intelligible: otherwise we may be guilty of *akrasia*, or just not thinking things through. But that something need not be either general or strong. The activity may just seem somehow boring or silly or lacking in point – or merely be such that, on occasion, we can't be bothered engaging in it (“for no *particular* reason,” as we might then be wont to explain). Perhaps the phrase “Not tonight, honey” will put the reader in mind of one set of examples, but there are plenty more of its kind. Just think of some pastime you've tried and quite enjoyed – e.g., chess, knitting, rock-climbing – that some well-meaning friend has urged you to take up as a hobby, but which fails to pique your ongoing interest. Unless you're at a loose end and fully expect you would find the endeavor positively *blissful*, a diagnosis of *akrasia* would often seem off-base, or at least overblown. The phrases “Only if you want to” and “Not unless you want to” capture this common sense idea: viz., that one's marked lack of interest in an activity has decisive normative authority – when it comes to our *own* interests, at any rate.

Interestingly, Dancy takes these sorts of cases as grounds for rejecting the *reasons-ought platitude* – I ought to (or, more weakly, *should*) do what I have most reason to do – for so-called “enticing” reasons (2004b). But, on my view, we need not make this unprincipled exception to accommodate unopposed reasons to do what would be fun, exciting, appealing, etc. I think these enticements *can* take us to a perfectly valid ‘should’ or ‘ought’ claim – it is just that they are particularly liable to vetoing, since our desires for diversions wax and wane. Similarly, Scanlon admits that some reasons in this family – such as the reason to listen to a piece of music you'd enjoy – do appear to be “optional,” depending on whether or not we “decide” to act on them (‘decide’ being perilously close to ‘want,’ given Scanlon's denial of internalism) (2009, Lecture 1, pp. 4-5). The pure objectivist might try to explain away this appearance. But, on the veto power view, the explanation for this genuine phenomenon is simple and ready to hand: namely, sometimes we want to do these things, and at other times we don't. Note too that Scanlon and Raz both feel obliged to allow that desires can act as occasional tie-breakers, when the agent is faced with two equally good options (see 1998, p. 48, and 1999, p. 62, respectively). This is yet another rather *ad hoc* exception which the veto power theorist can avoid. On my

view, what is doing the normative work in these cases is not that the desired alternative is ruled *in*, but rather that the undesired alternative is ruled *out*.

So, where does all this leave us? Recall from my opening example of Van Gogh that reasons are supposed to be intimately connected with *advice* – the idea being that reasons underwrite ‘ought’ claims, and ‘ought’ claims capture in turn what an omniscient, benevolent advisor would recommend you to do, were she to have your ear.³⁴ But, over the course of this chapter, we have seen that hedonic considerations often fail to fit the bill, intuitively speaking. That is, they are *not* always the sorts of considerations that your ideal advisor would seem wise to automatically take into account. They are silenced, I have suggested, precisely when you genuinely don’t *want* that kind of pleasure, or honestly don’t *mind* that kind of pain. But the attractive resulting idea – i.e., that hedonic considerations are capable of being genuinely *silenced* – is unavailable to the pure objectivist, or at least not without great cost. I argued that they would be better off in biting the bullet, and saying that these reasons *cannot* in fact be silenced and are only ever outweighed. But this is a hard bullet to bite, since in many of the cases in which these reasons fell intuitively flat, countervailing reasons were pretty thin on the ground. So, any way they turn, the pure objectivist seems to need a helping hand from his internalist would-be foe. I therefore invite the conclusion that the proposed internalist addendum is well worth the objectivist’s adoption. In other words, the veto power view carries the day over a purely objectivist position.

In this chapter, I have argued that we do not have to choose between internalism and objectivism: we can endorse them both. Moreover, when it comes to self-interested reasons, I argued that we should do so. In the next chapter, I argue that the same goes for moral reasons, appearances notwithstanding.

³⁴ The ‘advice’ model is not compulsory, but there is rather a paucity of alternatives for reason centrists like Parfit and Scanlon, who want to reduce all practical normativity to practical reasons, and bill it as a substantive endeavor – which it would not be if practical reasons were conceived just as generic normative widgets. I come back to this issue in the next chapter.

Chapter Two

REGIMENTING REASONS INTERNALISM: MORAL REASONS AND ITERATED 'OUGHTS'

OR: You Shouldn't Have To

"It is undoubtedly in this way, again, that we are to understand the passages from scripture in which we are commanded to love our neighbor, even our enemy. For, love as an inclination cannot be commanded, but beneficence from duty – even though no inclination impels us to it and, indeed, natural and unconquerable aversion opposes it – is practical and not pathological love, which lies in the will and not in the propensity of feeling, in principles of action and not in melting sympathy; and it alone can be commanded."

~Kant, *The Groundwork of the Metaphysics of Morals*, 4:399

Bernard Williams famously propounded reasons internalism, claiming that you don't have a reason to seek things that you really don't want, or to pursue ends about which you just don't care (1981). So, don't care about fame or fortune? It can't matter *for* you unless it matters *to* you, Williams plausibly suggests – hence, someone so unworldly has no reason to pursue these worldly goods.³⁵ Indeed, the notion of desire-independent duties to the self being notoriously dubious, internalism about self-interested reasons *in toto* is plausible; I defended the position in Chapter One. (A brief recap: I argued that, although desires do not provide self-interested reasons, they nevertheless sanction or enable them – thus vindicating Williams' claim that indifference to a certain would-be good robs it of its goodness for you.)

In this chapter I want to focus though on moral reasons, which are a different kettle of fish, intuitively speaking. For consider (to adapt Williams' example, 1995): the fact that an abusive husband simply doesn't care about his wife's suffering hardly renders it acceptable that he mistreats her (we should all agree). It certainly doesn't ameliorate the harms done *to* her – quite the contrary, one might think. Nor does it inoculate this brute, seemingly, against our condemnation of him (as opposed to our collaborative advice or admonitions – which may well be withdrawn or hurled at him merely instrumentally, a distinction Williams

³⁵ 'Reason' is intended here in the strictly normative sense, i.e., as the kind of consideration that bears directly on what you ought to do – and where what you ought to do is understood, in turn, as what you have most or decisive reason to do. The word 'reason' is sometimes used instead in the motivating sense, to mean considerations which help move us to action. Obviously if Williams is right, then there is an intimate connection between normative and so-called motivating reasons. But to avoid confusing the issue, or begging any questions, I'll continue to eschew this usage here.

emphasized; see 1995). But I think that we can hold on to these undoubted moral truths without compromising, indeed while clarifying, Williams' driving insight about reasons – at least if we acknowledge that he did not always put his point as clearly as he might have done.

This chapter, accordingly, has two ultimate aims: firstly (and primarily), to offer a new way of understanding Williams' internalist position, as applied to moral matters; secondly (and secondarily), to make a provisional plea for the view, so understood. I'll approach the issue though from an admittedly unusual starting point: namely, certain puzzles and distinctions in deontic logic. But the expressive-cum-conceptual resources to be garnered in this context seem to me to shed considerable light on both the entrenched debate about the veracity of Williams' position, and also the question of why it has become so entrenched, over the past thirty-odd years. So I'll be taking the back roads towards my destination, over the first part of this two-part chapter.

Here's a brief preview, and a few spoilers: in part one, I will be developing a kind of error theory about deontic modals. I argue, drawing on various highlights in the literature, that sentences like 'John ought to X' admit of two distinct, genuinely normative readings. Firstly, they can mean (roughly) that John is *defective* in some respect insofar as he doesn't X; call this the *evaluative* reading. Secondly, they can mean (again, roughly) that John is actually *called upon* to X; call this the *prescriptive* reading. Now, these propositions usually stand and fall together, but by no means always. For example, if John is called upon to do something only insofar as he, being defective, has gotten himself into a bad situation, then the two can come apart. More importantly, for our purposes, if John is *not* called upon to do something insofar as he, again being defective, wouldn't be so much as capable of pulling it off, then the two propositions will again diverge in their respective truth-conditions. So it is important not to conflate the two propositions in examples of this kind.

Moreover, these two senses of 'ought' can interact and iterate in non-trivial ways: as we see, for example, in the counterfactual claim that 'If John *weren't* defective, he *would* be called upon to X.' Or, to do some backwards translation: 'John ought to ought to X' – which thus has a true reading, and a meaning all of its

own. But this is, of course, a horrid little sentence, which rubs up aggressively against proper English syntax.³⁶ Little wonder then that, as philosophers accustomed to paraphrasing all practical norms in terms of ‘ought’ claims, we have by-and-large been overlooking the possibility and significance of propositions of this kind. But such propositions are crucial to have in one’s repertoire, I’ll argue. For, sometimes, when philosophers insist that a proposition of the form ‘John ought to X’ is true, they are mistaken about what they really mean to claim. They should be invited to substitute the iterated claim ‘If John *weren’t* defective, he *would* be called upon to X.’ (Or, if you absolutely insist: ‘John ought to ought to X.’) Since this captures their intended meaning, the appeal of the original, controversial sentence (i.e., ‘John ought to X’) will tend to dissipate, along with much of the surrounding debate. Indeed, I’ll go so far as to suggest in section 1.4 that Chisholm’s puzzle (the so-called paradox around which deontic logic is oriented) can be elegantly handled by means of this maneuver.

I’ll go on to argue, in part two of the chapter, that the debate about Williams’ reasons internalism is closely analogous to the quandary over Chisholm’s puzzle, and responds to the same style of treatment. That is, the Williams-style internalist should be understood as proposing that, sometimes, sentences of the form ‘John has a moral reason to X’ will fail, despite it being true that ‘If John *weren’t* defective, he *would* have a moral reason to X.’ The interlocutor’s insistence that ‘John has a moral reason to X’ *must* be true – lest we be left stranded in the wasteland of vulgar moral relativism – should be dispelled by the internalist’s freshly minted resources with which to condemn John. (Indeed, his not having a moral reason to X is partly *constitutive* of his defectiveness, the internalist may say.) But the internalist still has a substantive and controversial claim about how the latent moral prescription expressed by the failed reasons claim can be short-circuited – namely, by John’s defectively not caring about the locus of the moral concern. So the real philosophical issue turns out to

³⁶ But so too does the sentence: ‘John *will* ought to X.’ Future obligations are hardly a suspect class, though, so I don’t think this point should worry us unduly. It is just that English grammar makes some normative statements difficult to express using only ‘ought,’ the normative term so favored by philosophers.

be an oldie but a goodie, with unsurprisingly Humean origins, rooted in sentimentalism. It is time to get back to our roots, I'll eventually suggest. But we'll start with more formal considerations, and work our way to the moral issues from there.

PART ONE: ON DEONTIC MODALS

1.1: CHISHOLM'S PUZZLE

The 1960s was a good decade for deontic logic as well as free love. Following on from G. H. Von Wright's pioneering work in establishing deontic logic as a species of modal logic, Roderick Chisholm's three-page paper "Contrary-to-Duty Imperatives and Deontic Logic" (1963) was largely responsible for establishing it as a distinctive modal logic, in identifying the following puzzle. Consider the following quartet of propositions, which Chisholm expressed thus:

- (1) It ought to be that a certain man go to the assistance of his neighbours;
- (2) It ought to be that if he does go he tell them he is coming; but
- (3) If he does not go then he ought not to tell them he is coming; and
- (4) He does not go.

These propositions are logically independent of each other *and* perfectly consistent, intuitively speaking – they describe a readily comprehensible state of affairs, since lousiness and laziness are sadly hardly uncommon. But, in standard deontic logic (SDL, hereafter), the logical independence of these propositions can be retained only at the expense of their consistency. Take the natural first-pass formalization, where O denotes the necessity-type operator 'ought to:'

- (1) Ogo
- (2) $O(go \rightarrow tell)$
- (3) $\neg go \rightarrow O\neg tell$
- (4) $\neg go$

Now, in SDL (the normal modal logic **KD**), we can derive:

- | | | |
|------|-------------------------|--|
| (c1) | $O\neg tell$ | from (3) and (4), by <i>modus ponens</i> |
| (c2) | $Ogo \rightarrow Otell$ | from (2) and the axiom K : $O(p \rightarrow q) \rightarrow (Op \rightarrow Oq)$, by <i>m.p</i> |
| (c3) | $Otell$ | from (1) and (c2), by <i>m.p</i> . |
| (c4) | $\neg O\neg tell$ | from (c3) and the axiom D : $(Op \rightarrow \neg O\neg p)$, by <i>m.p</i> . ³⁷ |

So contradiction follows; a paradox is born.

1.2: ITERATED DEONTIC MODALITIES

How should we react to Chisholm's puzzle? A seemingly unrelated three-page paper published shortly afterward holds, I think, the key. In "Iterated Deontic Modalities" (1966), Ruth Barcan Marcus was, to the best of my knowledge, the first (and one of the few subsequent) logicians to question the widespread assumption that deontic operators cannot iterate, or nest, without thereby collapsing. (Hence the frequent addition to SDL of **A4**: $O(Op \rightarrow p)$, from which $OOp \rightarrow Op$ then follows via **K**.) Barcan Marcus builds her case by considering the sentence:

(PARK) Parking on highways ought to be forbidden.

As Barcan Marcus notices, sentences like this make perfectly good sense, even though iterations of deontic operators often *do* look like redundant 'stuttering' (e.g., 'You ought to ought to X'). So, we will be moved to ask: whither and why the difference?

³⁷ And even if we drop **D**, because we want to be able to accommodate conflicts of duty, they evidently shouldn't be demanded by the logic itself (that is, on pain of inconsistency). Moreover, the situation envisaged is *not* naturally construed as involving a conflict of duty at all. So, as far as solutions go, this is a blind alley.

Another blind alley involves subjecting (2) and (3) to a uniform semantical treatment. But if we formalize (3) on the model of (2), then (3) follows from (1); if we formalize (2) on the model of (3), then (2) follows from (4). But, intuitively, these propositions are logically independent of one another. It is thus widely agreed that these two propositions must have a different logical form. See McNamara, 2008.

Barcan Marcus' characteristically neat suggestion was that sentences like **PARK** trade on and evince a distinction between two different senses of 'ought' – the *EVALUATIVE* sense and the *PRESCRIPTIVE* sense. Now, taken by itself, the suggestion that there are two sorts of 'oughts' – the *evaluative* sense expressing something in the vicinity of *how things ought to be*, and the *prescriptive* sense expressing, roughly, *what ought to be done* – was not unique to Barcan Marcus. In fact, it has been championed and deployed by several philosophers since (as we'll see in the next section, where I turn to clarifying the distinction).

What was ingenious in Barcan Marcus' paper was the idea that, having distinguished the two senses, non-redundant *iterations* of these distinct deontic operators could be accommodated and expressed to salutary effect. Let's see how this works using the example sentence **PARK**. It says, broadly speaking, that a certain novel prohibition ought to be put into effect. Now, while Barcan Marcus thinks that prohibitions are naturally understood as inverted prescriptions, and are thus issued in the same mode, she suggests that it would be misleading to read the outer 'ought' similarly – or **PARK** as thus prescribing a prohibition. For as Barcan Marcus points out, **PARK** would then be equivalent (*mutatis mutandis*) to:

(**PARK***) It is forbidden that parking is permitted on highways.

But, writes Barcan Marcus, sentences like **PARK*** are “either peculiar or false. How can it be forbidden that a state of affairs prevail unless it is part of a blueprint for creation?” she asks (p. 580). In other words, how can one now forbid a (permission-conferring) event that – while its *effects* persist – has itself already happened? A mandate to alter the status quo is one thing, retrospectively forbidding its genesis quite another; the idea makes little sense, for reasons soon to emerge.³⁸

³⁸ And if we read **PARK** as prohibiting *continuing* to permit parking on highways, then a number of difficulties arise. For it is perfectly possible to advance **PARK** without knowing or even presupposing that some person or body *could* legitimately revoke the permission, given the relevant social and legal mores. One might even know that the repeal would *not* be legitimate, but level one's remark in a purely jurisprudential vein – i.e., as a remark about what would make for a better world, or a more just society, *ceteris paribus*. Cf. “It would be better if owning a handgun was forbidden in the US.

I do not want to suggest that the situation is totally hopeless for someone who takes **PARK** and **PARK*** to be deceptive equivalents. But it is hard not to agree with Barcan Marcus that it is considerably more natural to instead read **PARK** as a *prescriptive claim embedded inside an evaluative claim*, such that the sentence in its entirety says, roughly, that the world would be a better place were a certain prohibition to be put into effect. There are other sentences that call even more forcefully for this kind of reading: the remark “You shouldn’t have to (but you do)” is hard to parse in any other way. Such sentences are hardly semantic anomalies, either. Indeed, it is tempting to think that the general-purpose jurisprudential question (“Which laws ought there to be?”) might best be understood along these lines, and likewise for rule reform generally speaking. But we can still accept that unmixed iterations of evaluative ‘oughts’ – as well as duplicate prescriptive ‘oughts’ (i.e., indexed to the same agent) – are redundant, as seems likely.³⁹

In what follows, I plan to take seriously Barcan Marcus’ closing suggestion that “some of the controversy about what are or are not intuitively acceptable deontic truths would be resolved if the two uses of ‘ought,’ although closely connected, were not confounded” (1966, p. 581). And despite the lack of uptake of her subsequent suggestion that the two distinct ‘oughts’ are capable of iterating without collapsing, I believe that it sheds light on puzzles which have dogged both deontic logic and metaethics over the past several decades. So I propose in this chapter to revive and fully exploit her insights.

1.3: A TALE OF TWO ‘OUGHTS’

Earlier I stated my view that, in “Iterated Deontic Modalities,” Barcan Marcus provides the key to the puzzle which Chisholm identified just prior. I’ll spell out that thought in the next section, but we still lack the

But such a law would be unconstitutional, so workarounds are needed in lieu of a constitutional amendment.” This is, in fact, a reasonable statement of the dominant view among American liberals. Thus, there are arguably better examples than **PARK** which Barcan Marcus could have used to make her point instead.

³⁹ The question of whether an evaluative claim can be intelligibly nested inside a *prescriptive* claim is trickier, and I’ll set it aside here. Duties to improve oneself, or imperfect duties to make the world a better place, would be natural first-pass examples though.

requisite grease, which we can garner by undertaking a preliminary clarification of the distinction between prescriptive and evaluative ‘oughts.’

We might initially suppose that the two ‘oughts’ differ – at least partly – in that to which they *apply*. Barcan Marcus herself (perhaps inadvertently) suggests as much, in the following passage:

When we assert that it ought to be the kind of world where a person’s actions always flow from his obligations, this expresses our belief about *a state of affairs which, if it obtained, would make for a better world*. Such beliefs may provide reasons for justifying a prescription, but they are themselves not prescriptions. (1966, p. 580, my italics)

So there is a temptation to think that the ‘ought’ of evaluation applies exclusively to states of affairs in the *world*, whereas the prescriptive ‘ought’ applies exclusively to *persons* – insofar as rational actors alone may be intelligibly permitted, enjoined, or forbidden to act in certain ways. Such a picture is elsewhere made explicit. Lloyd Humberstone – in his illuminating paper “Two Sorts of Oughts” (1971) – paraphrases the distinction in play as the *situational* versus the *agentive* ‘ought’ (in discussing the lament “It ought not to have happened,” said of some natural disaster).⁴⁰

However, I think that this way of glossing the distinction between the evaluative and the prescriptive ‘ought’ is actually misleading. For, as well as speaking about what ought to be the *case*, we can speak meaningfully, and seemingly non-metaphorically, about what *people* ought to be like over and above what they ought to do.⁴¹ In other words, we regularly identify ways in which people are to various degrees *defective*. For example, we criticize each other’s moral characters all the time.⁴² And it is surely very implausible to think that (e.g.)

⁴⁰ Notice here that there might well be an *implicature* that the relevant bodies (or deities) should have prevented or ameliorated the terrible situation. But this is not an entailment, since it is easily cancelled by the further declaration: “There’s nothing anyone could have done though” or “I don’t believe in God.”

⁴¹ On the other hand, the idea of instructing the world to lift its game *is* deluded or comical – no doubt about that; hence the joke about the man who shakes his fist at a cloud.

⁴² Epistemology and aesthetics are further robustly normative domains which seem to involve evaluating some aspect of the *person* without instructing them to change their beliefs or tastes respectively – the former being a nonsensical

wickedness can be understood just as a propensity to fail to do what one ought to do, in severe and repeated ways. It is rather a property of whole persons, surely, encompassing the attitudes and mental habits that comprise their entire moral outlook.⁴³

Consider the things we say about people we deem truly immoral (through-and-through, or just in some respect). For, as J. L. Austin remarked, ordinary language is by no means the last word philosophically, but it is at least the first. So it is suggestive that, sometimes, only an evaluative claim conveys what we wish to say about such a person – expressing a thought that an alternative sentence employing the prescriptive ‘ought’ seems not to adequately capture. On the contrary, the alternative prescriptive claim may seem not only insipid but inapt. Gilbert Harman, independently drawing the same distinction as Barcan Marcus and Humberstone – i.e., between what *ought to be* and what *ought to be done* – puts the point particularly sharply. He writes:

...it sounds odd to say that Hitler should not have ordered the extermination of the Jews, that it was wrong of him to have done so. That sounds somehow “too weak” a thing to say. Instead we want to say that Hitler was an evil man. Yet we can properly say, “Hitler ought not to have ordered the extermination of the Jews,” if what we mean is that it ought never to have happened; and we can say without oddity that what Hitler did was wrong. Oddity attends only the inner judgment that Hitler was wrong to have acted in that way. That is what sounds “too weak.” (1975, p. 7)

In this passage, Harman amply demonstrates that evaluative ‘oughts’ can apply to people as well as situations, insofar as we say (in his example) that *Hitler was an evil man*. Now, as Harman readily admits, negative evaluative judgments about people *often* carry the implicature that they should buck up and lift their game

instruction, by the lights of most epistemologists, as doxastic involuntarists. The same plausibly goes for the latter instruction, i.e., to alter one’s aesthetic sensibility. One can’t shrug off philistine tendencies by sheer force of will.

⁴³ This claim is very much in line with Williams’ overall ethical outlook. Williams is deeply critical of the ‘reductivist enterprise’ of trying to assimilate all ethical notions to moral obligations, duties, or reasons (terms he uses more or less interchangeably). In “Morality, the Peculiar Institution,” he writes of Ross’ ‘duty of gratitude’ that “what Ross is trying to force into the mold of obligation is surely a different ethical idea, that it is a sign of a good character to want to return benefits. This characteristic is not the same thing as a disposition to do what one is morally obliged to do.” He goes on to decry Ross’ ‘duty of justice’ as an ‘extraordinary’ distortion of the notion in question. “The requirements of justice concern, in the first place, *what ought to happen*.” In contrast, “the way in which a given requirement of justice relates to *what a given person has reason to do, or more specifically is under an obligation to do*, is a matter of how that person stands to that requirement” (my italics). Williams concludes: “It is a mistake of morality to try to make everything into obligations” (1985, pp. 179-180).

(even if the mandate tends to be rather unspecific). But not always – or at least the implicature may be blocked. Indeed, sometimes it seems *constitutive* of the wickedness of the person with whom we are dealing that the prescription would inevitably fail to move them (as we might well imagine with Hitler). Knowing this, we find the prescription simply dying on our lips. “Too weak a thing to say” is *itself* too weak a thing to say, I would in fact suggest.⁴⁴ (Indeed, this is one way of putting Williams’ driving insight; but I anticipate.)

The suggested upshot is that evaluating a person negatively does not mean or entail that we thereby instruct them to *change*, or even to change their behavior. (Nor need it imply that, advanced at some suitable place and time, such an imperative *would* then be apt.) For the subject of the negative evaluation may be quite *incapable* of responding by changing in the relevant mental respect, or even improving their conduct.⁴⁵ And the prescription and the corresponding prescriptive claim – which articulates, I have suggested, the latent aptness of that prescription – plausibly take for granted the capacities that would enable it to be *followed*. If that is right, then there is a sense in which the ‘ought’ of prescription may be taken to imply ‘can.’ Whereas the ‘ought’ of evaluation carries no such entailment, for all that has been said so far. As Humberstone suggests:

The sincerity and felicity conditions for the issuing of imperatives involve the presumption that the addressee is able, should he so decide, to obey the imperative. Consequently, when ‘oughts’ entail imperatives, they may be expected to imply ‘cans.’ It follows, by a simple transition, that when ‘oughts’ do not entail ‘cans’, they do not entail imperatives. (1971, pp. 10-11)⁴⁶

⁴⁴ Harman goes on to say as much himself, in his (I think) somewhat misleadingly entitled, but otherwise enlightening paper, “Moral Relativism Defended” (1975). That is, Harman argues that such ‘inner judgments’ are not only “too weak” but actually *false* for agents who lack suitable motivations. This position is thus a clear predecessor, and close relative, of Williams’.

⁴⁵ It may be all downhill from here: all roads leading to Gomorrah, so to speak. Moreover, unlike prescriptions, evaluative notions seem to come in degrees (Humberstone, 1971). So we may use them to mark out an *ideal* which the agent is in some sense not obligated to actually achieve. The envisaged actions are then supererogatory, not obligatory. As Williams elsewhere remarks, during the course of criticizing the idea that all ethical considerations can be assimilated to the class of moral obligations, “ethically outstanding or possibly heroic actions... in being more than obligations, are not obligatory, and we cannot usually be asked to do them or be blamed for not doing them” (1985, p. 188).

⁴⁶ It is worth noting that Robert Stern (2004) argues that Kant himself only makes claims in the vicinity of ‘ought implies can’ when discussing commands specifically. Cf. the famous passage from Kant with which I opened.

I think Humberstone is onto something important here, and I'll return to his point in section 2.3 – at which time I'll be ready to take these seeds of an argument on Williams' behalf and coax them into fruition.

In the meantime, we are now equipped to deal with Chisholm's puzzle, before we move on to consider certain parallels and implications for the fate of Williams' position.

1.4: A SOLUTION TO CHISHOLM'S PUZZLE

Putting Barcan Marcus' suggested lessons together with Harman's, we now have it that (a) there are two sorts of 'oughts,' evaluative and prescriptive, (b) that they can (when combined) iterate without collapsing, and (c) both sorts of 'oughts' can apply properly to persons, to mean two quite distinct things. So, in our deontic logic, we may – having distinguished the two 'oughts' using suitable deontic operators, O_E and O_P – now trade in \mathbf{K} : $O(p \rightarrow q) \rightarrow (O_p \rightarrow O_q)$ for a rather more nuanced alternative.

Now, although I cannot hope to make an adequate argument to this effect here, Chisholm's original formulation of his four propositions is already quite suggestive. It suggests that the 'ought' of evaluation will be more appropriate than the alternative prescriptive 'ought' to taking conditional propositions in their entirety in its scope. Whereas, the 'ought' of prescription will apply felicitously only to suitable objects of individual actions. We might naturally take these – on a first pass – to be simple (i.e., atomic or negated atomic) propositions, such as the consequents of such conditionals.⁴⁷

⁴⁷ So here is a hypothesis, already implicit in (or at least extractable from) Chisholm's paper: wide-scopers (e.g., John Broome) are going for the 'ought' of evaluation, whereas narrow-scopers (e.g., Niko Kolodny) are going for the 'ought' of prescription instead, in their debate over whether we ought to satisfy the requirements of rationality. If so, they may be talking past each other.

It would then be natural to replace **K** with:

$$\mathbf{K}^*: O_E(p \rightarrow q) \rightarrow (O_E p \rightarrow O_E O_P q) \quad (\text{with } q \text{ simple, hereafter})$$

Or, to express the idea more intuitively, in the manner of a natural deduction-style rule:

$$\mathbf{K}^*\text{-rule} \quad \frac{O_E(p \rightarrow q) \quad O_E p}{O_E O_P q}$$

K* says (roughly) that, if it ought to be that an agent Xs, and that $X \rightarrow Y$ ought to be made the case by her too, then it also ought to be that the prescription ‘Y!’ applies to the agent, by virtue of the fact that she Xs.

A natural complement to **K*** would be:

$$\mathbf{InvK}^*\text{-rule} \quad \frac{O_P q \rightarrow p \quad O_E \neg p}{O_E \neg O_P q}$$

In axiomatic terms, **InvK*** would be $(O_P q \rightarrow p) \rightarrow (O_E \neg p \rightarrow O_E \neg O_P q)$.⁴⁸ **Inv-K*** addresses the inverse possibility, saying (roughly) that, if it ought to be that an agent does *not* X, and that she only ought to Y if she *does* X, then it ought to be that the prescription ‘Y!’ does not apply to her, seeing as she duly doesn’t X.

⁴⁸ $O_E O_E p \rightarrow O_E p$ is a very plausible additional axiom, as is $O_{P_i} O_{P_i} p \rightarrow O_{P_i} p$ (where i is the index for a specific agent). We will probably also want the relevant analogues of **D**. Note that the evaluative version, i.e., $O_E p \rightarrow \neg O_E \neg p$, guarantees that the two introduction conditions for **K*** and **InvK*** cannot hold simultaneously. The natural analogue of **A4**, $O_E(O_E p \rightarrow O_P p)$ is plausible as well, but I won’t consider further candidate rules which mix the two deontic operators here. At that point, a nice model theory would be required. And as far as the model theory goes, I would be inclined to base it on logics for belief, i.e., with non-factive necessity-type operators (indexed to each agent) to capture the prescriptive ‘ought’. The evaluative ‘ought’ might be handled using a single operator, assuming we impose an ordering on worlds with respect to their optimality in the various relevant respects. The resulting logic may be a bit more complex than we’d like, but there are no evident technical obstacles to constructing it in this fashion. And, without such finessing, deontic logic may remain “a misleading technical exercise,” as Barcan Marcus fears (1966, p. 582).

In order to get a better fix on these rules, and to see why they are instructive in this connection, let's apply them to Chisholm's quartet, which we can now re-express as follows:

- (1') $O_E go$
- (2') $O_E(go \rightarrow tell)$
- (3') $\neg go \rightarrow O_P \neg tell$
- (4') $\neg go$

Now we can only derive:

- (c1') $O_P \neg tell$ from (3) and (4), by *modus ponens*
- (c2') $O_E go \rightarrow O_E O_P tell$ from (2) and **K***: $O_E(p \rightarrow q) \rightarrow (O_E p \rightarrow O_E O_P q)$, by *m.p.*
- (c3') $O_E O_P tell$ from (1) and (c2'), by *m.p.*

So, much as before, *modus ponens* will give us $O_P \neg tell$. But **K*** plus *modus ponens* now gives us $O_E go \rightarrow O_E O_P tell$, which, via *modus ponens* again, only yields $O_E O_P tell$. Hence there is no direct conflict here: our agent *ought to be such* as to render the prescription to tell his neighbors he's coming *apt*, by intending to go to their aid. Or, to trade idioms: *a decent person in this situation would have a decisive reason to call ahead*. But, as it is, this man *ought not* to tell his neighbors he is coming, seeing as he has no *intention* of going to their assistance, and it would be wrong to lead them on. A contradiction is thus avoided, and the former temptation to assert the jettisoned $O_P tell$ is seen to rest on a certain subtle confusion, between simple and complex (i.e., iterated) 'ought' claims.⁴⁹ But, the confusion being subtle, it is not at all puzzling why the puzzle should have arisen in the first place – a further constraint on any adequate treatment of a logical-cum-philosophical puzzle.

⁴⁹ So we see that, using some standard terminology from the literature, we have *factual detachment* but not *deontic detachment*. But the **K*** and **InvK*** rules above together comprise an excellent *substitute* for deontic detachment, capturing much of what theorists in this camp want to say, I believe. Cf. Paul McNamara's useful *Stanford Encyclopedia of Philosophy* entry (2008), for its overview of the literature on this and other points. Another helpful survey is Carmo and Jones' entry in the *Handbook of Philosophical Logic*, Vol. 8. As Carmo and Jones explain, the intuition behind deontic detachment is largely that the original obligation has not been *defeated* or *overridden*: rather, it has been *violated* and *usurped* (2002, p. 282). Jones and Porn write, along similar lines, that there is a prevalent feeling that "in some sense or other, the agent concerned *does* have an obligation to tell his neighbors that he is coming" (1985, p. 277, my emphasis). The error theory I am proposing aims to respect these intuitions, by offering a new way of cashing them out.

Moreover, if we also endorse **InvK***, we can say still more. Given strong substitutability (as in SDL), $O_E go$ and the rules for classical negation give us $O_E \neg go$. We can also assume the plausible pair of propositions $go \rightarrow O_P tell$ (given that $O_E(go \rightarrow tell)$) and $O_P tell \rightarrow \neg O_P \neg tell$ (the prescriptive analogue of **D**). Thus, with the help of transitivity, **InvK*** yields $O_E \neg O_P \neg tell$. In other words, our agent *ought to be such* as to render the prescription to *not* tell his neighbors he is coming *inapt*, by not *failing* to go their aid. Or again, and rather more clearly: *only a louse would have a decisive reason not to call ahead in this situation*. (Of course, the man may not be a louse in general; he may be behaving uncharacteristically. In which case, we might say, suggestively: “You’re acting *like* a louse,” or “You’re *being* a prick,” etc.) But, as things stand, this man ought *not* to tell his neighbors he is coming, seeing as he *is* lousily neglecting to go.

The strategy evidently generalizes.⁵⁰ And it represents, I believe, an elegant, satisfactory – and novel – solution to the puzzle Chisholm identified.⁵¹ I hasten to say, however, that it is my own proposal only in the sense that

⁵⁰ For example, it also applies to James Forrester’s well-known related paradox (1984): “If you kill your Mother, you ought to kill her gently.” Given the expressive resources now at our disposal, there are two ways of reading this sentence: (1) If you kill your Mother, it would be *better* (evaluatively speaking) if you killed her gently, or (2) If you kill your Mother, then you ought (in the prescriptive sense) to kill her gently. (1) might be termed a ‘best of a bad lot’ conditional, and it shows that evaluative notions can pertain to ‘suberogatory’ as well as supererogatory actions – which fall below the threshold of decent conduct enshrined by the bare evaluative ‘ought,’ but nevertheless are better than certain going alternatives. In any case, neither (1) nor (2) implies that you ought (in the evaluative sense) to kill your Mother – gently or otherwise – let alone that she be killed (cf. the Good Samaritan paradox, which can be dispensed with in a similar fashion). Indeed, given that the *opposite* is true (i.e., $O_E \neg kill$) we can say that $O_E \neg O_P kill$ gently, as in Chisholm’s puzzle. There is also more to be said about the *pragmatics* of the contrary-to-duty imperative (i.e., $O_P kill$ gently), since we might be loath to get ourselves embroiled in this sordid business by *issuing* this prescription even so. But I’ll confine myself to the semantical issues here.

It is worth noting that distinct evaluative claims can also iterate, given sentences like (1). To wit: ‘*It ought not to be that it would be better if you were to kill your Mother gently*’ (given that are about to kill her, one way or another). The same strategy can be employed in examples like Prakko and Sergen’s case (1996), where: (1) It ought to be the case that there are no dogs, (2) It ought to be the case that if there are no dogs, then there are no warning signs, (3) If there are dogs, then it ought to be the case that there are warning signs, (4) There are dogs. Here, I think we should say that *there ought to be warning signs*, but that *it ought not to be the case that this itself ought to be* (on account of there being dogs). At any rate, I hope to have made a preliminary case to the effect that the current proposal gives us rich resources for dealing with these core puzzles in contemporary deontic logic.

⁵¹ I can’t hope to defend this proposed solution to these puzzles against – or even to distinguish it from – its various rivals in the literature. Very briefly though: it does have some forerunners, in the form of accounts which distinguish between what ‘ideally ought’ to be done and what ‘actually ought’ to be done. (See, e.g., Jones and Porn, 1985.) But it is considerably more general, and hence doesn’t come unstuck on *atemporal* variants of the puzzle (see *n.* above).

I deserve the blame if it turns out to be wrong. If it is along the right lines, then all the ingredients were evidently already there in the literature – all that remained to be done was to put them together appropriately.

What does all this have to do with moral philosophy though? It is time to get back to our original subject, viz., the entrenched debate over Williams' reasons internalism – which I think succumbs to essentially the same treatment as Chisholm's puzzle case.

PART TWO: ON MORAL REASONS

2.1: REASONS AND RELATIVISM

Williams' internalism about reasons – i.e., the claim that practical reasons require the backing of suitable motivations – is a bit of an odd duck, as far as philosophical claims go. The idea seems on the surface to be perfectly straightforward. But a great deal of ink has been spilt over it, with no sign of the flow abating. Nor is the issue easily dismissed as resting on a clash of intuitions, a misguided question, or an illicit equivocation. So it stubbornly remains, as a kind of divisive stain on metaethics. Some prominent theorists find Williams' thesis intuitively compelling, if not on Williams' grounds. At least an equal number, however, think it deeply mistaken. Some philosophers take it to be an important truth; many more think that it is an obvious falsehood. The literature burgeons, but the debate seems to be at a standstill.

What is the explanation for this puzzling state of affairs? I think that part of the explanation lies in the widespread tacit assumption that internalism could only be true if desires *provide* reasons. But this is an illicit assumption, since the internalist can say alternatively that desires *enable* reasons – that is, desires allow objective (i.e., desire-independent) would-be reasons to exert their potential normative force.⁵² (I defended

⁵² An interesting possibility which arises in this connection is that the distinction between would-be and actual reasons may be marked in natural language by the phrases 'there is a reason' versus 'he has a reason.' (Although note that some authors use these two phrases to mark a different distinction, between *objective* and *subjective* reasons, the latter being 'reasons' only by the lights of the agent's possibly erroneous *beliefs*.) Williams once suggested as much himself: "When we say that someone ought to have acted in some required or desirable way in which he has not acted, we sometimes say

this position in Chapter One.) If so, then the deadlock may arise partly because friends of internalism are focused on cases where a lack of desire seems to *defeat* an objective would-be reason; whereas, its foes are concentrating on cases where a (foolish) desire seems *not to give* someone a genuine reason to act. But, on my view about desires, both parties to the debate could well be right – except insofar as they continue to argue at cross-purposes.

This strikes me as an adequate explanation of the stalemate when it comes to self-interested reasons. But it can't be the whole story about reasons *writ large*, i.e., including moral reasons. Now, the line I'm pushing happily does allow friends of internalism to deny that someone's wicked desires *provide* reasons for her to do wicked things. But what about someone who (on the flipside) *just doesn't care* about some important moral matter? Whether or not they still have a *reason* to lift their game remains, for all I've said, a genuine bone of contention between internalism's friends and foes.

I think that the remainder of the explanation goes something like this: most of us, hopefully, share an aversion to vulgar moral relativism (let's call it *VULGARISM*, for short) of the kind that would prevent us from saying, for instance, many of the things I said in opening about the abusive husband's defects. But there is an evident potential worry that, having espoused *internalism*, *vulgarism* would be only an inference away.

that *there was a reason* for him to act in that way – he had promised, for instance, or what he actually did violated someone's rights. Although we can say this, it does not seem to be connected in any secure way with the idea that *he had a reason* to act in that way. Perhaps he had no reason at all. In breaking the obligation, he was not necessarily behaving irrationally or unreasonably, but badly. We cannot take for granted that he had a reason to behave well, as opposed to our having various reasons for wishing that he would behave well" (1985, p. 192). I am indebted to Judith Jarvis Thomson for discussion on this and other points. My claim about the connection between 'defectiveness' and evaluatives, and the distinction between evaluative and prescriptive notions, was also greatly inspired by her recent book *Normativity* (2008). However, I should make it clear that I understand and flesh out these notions in very different ways. In particular, she does not approve of my subsequent suggestion that 'ought' is ambiguous, or my endorsement of reasons internalism.

(Whereas the self-interested analogue of relativism – i.e., allowing that what someone does with her own life is in some sense up to her – is neither vulgar nor counter-intuitive: hence the disanalogy, I believe.⁵³)

I suspect that there is one camp of morally serious, anti-relativist philosophers who think or assume that the inference from internalism to vulgarism is compulsory or compelling – and thus move to dismiss internalism in virtually the same breath. (Indeed, and to anticipate, some philosophers in camp one have staked their research programs on an assumption which would elide the very transition.) Whereas those in the second camp are more suspicious of (or possibly blind to) this tricky putative inference; they thus aren't under pressure from that direction to deny internalism off the bat.

Williams himself obviously belonged squarely in camp two. Not only did he explicitly renounce vulgar moral relativism, but he explicitly denied the *inference* to it from internalism, in the following well-known passage:

There are many things we can say to people who lack appropriate items in their *S* [their motivational set]. Suppose, for instance, I think someone (I use 'ought' in an unspecific way here) ought to be nicer to his wife. I say, 'You have a reason to be nicer to her.' He says, 'What reason?' I say, 'Because she is your wife.' He says – and he is very hard case – 'I don't care. Don't you understand? I really do not care.' I try various things on him, and try to involve him in this business; and I find that he really is a hard case: there is *nothing* in his motivational set that gives him a reason to be nicer to his wife as things are. There are many things I can say about or to this man: that he is ungrateful, inconsiderate, hard, sexist, nasty, selfish, brutal, and many other disadvantageous things. I shall presumably say, whatever else I say, that it would be better if he were nicer to her. There is one specific thing the external reasons theorist wants me to say, that the man has a reason to be nicer. (1995, p. 39)

He went on, of course, to decline to say so once more, just as he had some fifteen years prior. This is the passage in which Williams most clearly articulates the intended scope of reasons internalism – a task by then long overdue, but better late than never.

⁵³ To spell out this asymmetry between moral and self-interested reasons more precisely, I take it to be unquestionable that we ought to care about the suffering of others. But I believe that there is nothing that we *ought* to care about, simply for our own sake – and nothing to say to the person who lacks certain self-interested aims, other than that they're missing out on certain objective goods. I argue elsewhere that this kind of freedom is at least partly constitutive of our autonomy (the sovereignty of our will, literally speaking). I also believe that we *have* to care about ourselves to some minimal extent, in order to maintain a grip on our identities; but that is another story.

Consider, as a counterpoint, a passage by Derek Parfit – who has his flag planted firmly in camp one, and is thus a staunch opponent of internalism (he calls theories in this family ‘subjectivist’). He writes:

When cruel people make others suffer, we can call these people vicious, odious, and callous. But on subjective desire-based theories, some of these people have no reason *not* to make others suffer. These other criticisms become much weaker if we must admit that, on our view, these people have no reason to act differently. We may even have to admit that, in harming themselves or others, these people are doing what we believe that they have *most reason* to do, and what they *ought* in the reason-implying sense to do. We may also have to admit that, on our view, these people have no reason to care about the fact that they are being imprudent, or that they are vicious and callous. Why should these people care about these criticisms if, as we believe, they have no reason to care? (Forthcoming, p. 716)

What is interesting about this passage is the difficulty Parfit finds in securing any content for Williams’ assurances that someone could be cruel without thereby having a *reason* to stop behaving cruelly. Hence Parfit’s worry that, having admitted that such people may have no such reasons, “we may also have to admit that... these people have no reason to care about the fact that they are... vicious and callous. Why should these people care about these criticisms if, as we [i.e., Williams et al.] believe, they have no reason to care?” (p. 716). The question hangs unanswered, suggesting that the inference from internalism to vulgarism is virtually irresistible. Parfit is thus baffled by Williams’ internalist position; he (ruefully) denounces Williams as a moral nihilist, for whom “no one has any reason to care about anything” (p. 712).⁵⁴

We can all agree with Parfit that, if these vulgar consequences *were* the result of adopting Williams’ internalism, then the position would be hopeless. Whatever we say about reasons, we must be able to say the appropriate things morally about the abusive husband: no question, to my mind – or to Williams’, I suggest.

⁵⁴ Similarly, Parfit writes: “Since Williams assumes that *ought* implies *has a reason*, but also believes that all reasons are internal, Williams concludes that the *ought* of moral obligation, when applied to [some] people, is an illusion. This view has unwelcome implications. On Williams’s view, we cannot claim that there are some things that it would be wrong for anyone to do. We cannot even claim that it would be wrong for anyone to torture other people for his own amusement. Given some sadist’s motivations, this man may have no reason to act differently. On Williams’s assumptions, this man’s sadistic acts would not then be wrong” (forthcoming, p. 708). Parfit does not seem to apprehend the possibility of a notion of wrongness which is *evaluative* rather than *reason-involving* and hence *prescriptive*. Cf. the passage I quoted from Harman, in section 1.3.

2.2: REGIMENTING REASONS INTERNALISM

But the foregoing remarks about Chisholm's puzzle show us that Parfit's reasoning here (and that of camp one generally) is actually highly suspect. For, as we have seen, it simply does not follow from the fact that

- (P1) one *ought* to X, and
- (P2) one ought, *if* one Xs, to Y,
- that
- (C) one *ought* to Y, even if one does *not* X.

In fact, that's precisely the inference that got us into trouble in Chisholm's example – which cleverly arranges it such that if one does not X, one actually ought *not* to Y. (But the point is perfectly general, since these inferences call for a uniform semantical treatment.)

Filling in the blanks here, we can say that it does not follow from the fact that

- (p1) one *ought* to care about the suffering of others, and
- (p2) one ought, *if* one cares, to stop making them suffer,
- that
- (c) one *ought* to stop making others suffer, even if one does *not* care.

So it does not follow that Williams, in espousing the premises, is bound to endorse the conclusion – or to deny (p1), having denied the conclusion and endorsed (p2). Thus, we see that Parfit's argument that Williams is a moral nihilist, for whom “no one has any reason to care about anything” (p. 712) rests on an invalid schema which he foists onto Williams. Of course, to say that an argument is an instance of an invalid schema is not yet to say that it does not go through in this particular case. (Every argument is an instance of the invalid schema: P, hence Q.) So, at the moment, the possibility that Parfit's inference on Williams' behalf may fail is offered merely as a coherent metaethical possibility. However, there is a prospective analogy with

Chisholm's puzzle (our flagship counter-example to the inference) to be pressed into service here, and I'll do so in the next section.

But first, something going on in the background here is that Parfit is what I call a *reasons centrist*. That is, he holds that reasons are the chief, indeed the sole, fundamental normative entity. Parfit is, accordingly, a *buck-passer* about goodness and value, wherein evaluatives are analyzed and understood solely in terms of reasons. For example, buck-passers hold (very roughly) that what it *is* for something to be bad (say) is for there to be a reason to stop, prevent, or destroy it, as appropriate to its nature. And similarly for all other evaluative notions, including those that apply to persons, which explains Parfit's blank insistence that "we cannot criticize or blame people for failing to do what we believe that they have no reason to do" (forthcoming, p. 708). In other words, the idea we encounter in Williams – that someone could be labeled *vicious* without it being implied that they have a *reason* to lift their game – fails to be intelligible for Parfit: it rests on a distinction which he professes not to be able to hear.

Now, practical reasons are generally assumed to be prescriptive entities, which bear on *what ought to be done* – recalling that what *ought to be done* is widely understood, in turn, as what we have *most* or *decisive* reason to do. (I discuss the possibility of dropping this assumption in the next section.) So I suggest that the foregoing discussion of deontic modals has rendered the suppressed distinction audible, intuitively speaking. Indeed, the distinction between evaluative and prescriptive claims seemed highly useful – perhaps even vital – in that context, Chisholm's paradox having stubbornly resisted previous attempts at an elegant resolution. (Deontic logicians hence now tend to employ dyadic non-compositional deontic operators in their semantics, a bit of a last resort.⁵⁵) So, notwithstanding Parfit's reasons centric program, the evaluative/prescriptive distinction

⁵⁵ McNamara (2008) summarizes, regarding Chisholm's puzzle: "So it seems like we are left with a dilemma: either (1) you allow factual detachment and get the consequences earlier noted to the effect that simply because someone will act like a louse, he is obligated to do slightly mitigating louse-like things, or (2) instead you claim that 'if *p*, then ought *q*' is really an idiom, and the meaning of the whole is not a function of the meaning of its conditional and deontic parts. Each seems to be a conclusion one would otherwise prefer to avoid."

surely counts as independently motivated at this juncture – our having acquired excellent evidence that it is a distinction which makes a difference.

And with this distinction (and the license to iterate) now in hand, we can reassure ourselves that Williams' claims about the abusive husband have content. To wit:

“Suppose, for instance, I think someone (I use ‘ought’ in an unspecific way here) ought to be nicer to his wife”

≈ $O_E be's\ nicer$

“I shall presumably say, whatever else I say, that it would be better if he were nicer to her.”

≈ $O_E she's\ not\ mistreated$

“There are many things I can say about or to this man...”

Inter alia: $O_E be\ cares\ about\ her\ qua\ person,\ qua\ woman,\ qua\ his\ wife$

“There is one specific thing the external reasons theorist wants me to say, that the man has a reason to be nicer...”

Instead, we can offer: $O_E O_P be\ nicer!$

$\neg O_E \neg O_P be\ nicer!$

Let's flesh this out a bit in plain English. And while we're at it, let's ditch Williams' word 'nice' here. It's altogether too anodyne to apply to the situation we might envisage.

First off, there is no doubt that the abusive husband *ought* to care about his wife's suffering (this is leveled as an evaluative claim).⁵⁶ Furthermore, and as should go without saying, it *ought not be the case* that the wife is treated and viewed in this way (another evaluative claim). It might even be incumbent on *us* to help her to escape (this would of course be a prescriptive claim). For, at this time, her husband is a thoroughly hard-hearted brute (evaluative again). More precisely, he is *defective* insofar as he does not care about her *qua* person (his being nasty and selfish), *qua* woman (his being sexist), or even *qua* his wife (his being ungrateful and

⁵⁶ Indeed, I'll assume henceforth that statements about *what we ought to care about* always involve evaluative rather than prescriptive 'oughts.' (Cf. the passage from Kant which I quoted in opening.) For we typically take the structure of our will not to be subject to direct manipulations of the self-same will, on pain of an infinite regress, or even conceptual incoherence. This also helps explain why Parfit's usual phrase 'reason to care' is quite unnatural, it being a veiled prescription; 'ought to care' is better, albeit subject to the unfortunate ambiguity just noted.

inconsiderate). However, we can say (as in Chisholm's puzzle) that any decent husband *would* care about his wife's suffering – in which case, if Williams is right, we *could* intelligibly instruct or advise him to start to treat her properly. We can also say, inversely, that only a brute would *not* be instructable in the proper behavior towards his wife. We might even add that the husband's indifference towards her is precisely what *renders* him – as well as shows him to be – a bad and dangerous person.

So there are indeed many things for Williams to say here, none of which collapses as a logical matter into what Parfit thinks he has to say. The Williams-style internalist just denies that we have the conceptual right to remonstrate with or *advise* the unresponsive brute against treating his wife badly – as opposed to managing, threatening, locking him up, or submitting him for treatment. As a matter of logic, we *can* deny this and only this, just as Williams has it – as we have to in Chisholm's puzzle, as we saw in section 1.4.

But *should* we deny this (and this alone), following Williams' lead? That is, does the invalid inference Parfit touts fail in this case too? In other words, does the analogy with Chisholm's case stretch as far as Williams needs here? I'll briefly sketch an argument to an affirmative answer in what remains.

2.3: BEYOND THE REACH OF REASONS

I will not try to mount a full-throated defense of Williams' internalist thesis here – at the end of the chapter, there will be premises left outstanding. And in a way, it would be a slightly half-hearted defense anyway, since it depends on an admitted stipulation about the meaning of the word 'reason.' Although this stipulation follows one of two conventional philosophical usages, I see little point in insisting that the English word must be used in one way rather than another: it is not the kind of word that has deep resonances, in which case it *can* be important to defend one's preferred usage. Still, on the most prevalent conception of a reason – which Williams espouses and Parfit appears to need, given other of his ambitions – internalism is very plausible, I'll now argue.

On this conception of practical reasons, reasons are intimately connected with *advice*. In other words, reasons are typically understood as pros and cons to be tallied in a benevolent, omniscient adviser's deliberations on your behalf about *what you ought to do*. Williams himself often employs this notion (and the advice-like notion of the outcome of shared *deliberations*) in characterizing reasons – it is not supposed to be an analysis. Parfit employs the concept of 'advice' from time to time as well, defending claims about the reasons someone has by pointing to the advice that we would give them. (Nor does Parfit ever explicitly distance himself from the advice-based conception of a reason. He would make life difficult for himself if he did, I'll argue in closing.) In a similar vein, Williams writes that it is crucial that claims about a person's reasons can be advanced in the "If I were you..." mode (1995).

Recall now the earlier conceptual point that offering someone advice that they could not *follow* would be at heart a nonsensical exercise. Or, to put the point as Humberstone does: a prescription or piece of advice (which are concepts in the same family, I assume) is a kind of *speech act* – which requires, for its felicity conditions, its recipient be capable of following it. And if these felicity conditions do not hold, then the corresponding prescriptive claim presumably fails in its own way. In other words, it is false, or at least not true – it being afflicted by something like presupposition failure. Tellingly, Williams concurs with this constraint on practical deliberation. He writes:

The fact that moral obligation is a kind of practical conclusion explains several of its features. An obligation applies to someone with respect to an action – it is an obligation to do something – and the action must be in the agent's power. '*Ought implies can*' is a formula famous in this connection. As a general statement about *ought* it is untrue, but it must be correct if it is taken as a condition on what can be a particular obligation, where that is practically concluded. If my deliberation issues in something I cannot do, then I must deliberate again. (1985, p. 175)

Even more tellingly:

...we must consider the question of *ought* as it occurs in the deliberative question 'What ought I to do?' There do seem to be considerations which provide a good case for distinguishing such a 'practical' *ought* from the general propositional *ought*. One reason that has been given for making such a distinction is that the practical *ought* is heavily governed by actual reality, whereas the general propositional *ought* is

permitted to be adapted to speaker's whim. So 'A ought to do X' (practical) does imply that it is possible that A do X; in general, however, 'O(A does X)' does not – if it is not possible that A do X, then all the speaker has to concede is that it ought to be the case that it be possible that A do X.

(“Ought and Moral Obligation,” in 1981, p. 118)

Finally:

Now 'ought to' in the modality of advice implies 'can,' because advice aims to offer something as a candidate for a deliberative conclusion. If X-ing is not available to the agent, 'You ought to X' cannot function as a piece of advice about what he should now do; when it is a matter of what I am to do, manifestly 'I cannot' acts as a stopper. (1995, p. 40)⁵⁷

Now, thinking back to Chisholm's example, it is easy to see why the advice to go to his neighbor's aid (and hence call ahead) cannot now be followed by the unneighborly man in question. For, the issue of whether or not to go has *already* been settled in the example. This is what renders it a *contrary-to-duty* situation; note too Chisholm's word choice here – “he *does* not go,” as opposed to “he *will* not go” (1963, p. 35). The issue is done and dusted, so our advice to him is constrained by and structured around this fact. We can say that he ought, in the prescriptive sense, to *have gone*, but the present-tense prescription is short-circuited. (Imagine a snippet of conversation, between this man and his mother: “You ought to go help him!” says his mother; “I'm not going,” says he, “And, anyway, it's too late now,” he adds. She realizes this to be true. “Still, you ought to be there,” she might say, by way of a retreat. “If you were a good fellow, you'd have done your duty,” his father could interject, iterating 'oughts' on cue.)

But why think that the case of the cruel husband conforms to the proposed analogy? What is the 'stopper' supposed to be here, exactly? For although this man currently *does not care* about his wife's suffering, there is no supposition in this case that the time to care has passed. On the contrary, he might perhaps come to his senses some day, and we would be very pleased. So the parallel with Chisholm's example appears to be in

⁵⁷ This facet of Williams' thinking might also explain his insistence in “Internal and External Reasons” that a person's reasons must be capable of *explaining* her actions. For it seems to follow from the preceding idea that your reasons be capable of being *acted on by you*, that you might *act on this basis*, and that your action might *be thus explained*.

danger. But the parallel will be restored by the – I think plausible – idea that there are people who presently *can't* respond to moral advice as such. (This 'can't' is intended to mark their present psychological incapacity, not in a temporal or predictive vein. Some day they might wake up. But we can only remonstrate with the agent who is currently standing before us.) For consider: in order to *follow* an instruction, or *take* a piece of advice, its recipient must possess not only the relevant behavioral and logistical capacities to act it *out*, but also the relevant cognitive and conceptual capacities – to act *out* of it, so to speak. Otherwise, how could we explain the fact that non-human animals can be commanded but not advised – and that they can obey but not defer?

The suggestion is that, in order for you to be receptive to a piece of moral advice, you have to be capable of understanding it *as* moral advice. And for Williams, this receptivity necessarily involves *caring* (at least deep down) about the locus of the moral concern – and thus being capable of being moved by it, literally and figuratively speaking.⁵⁸ Otherwise, you could certainly *behave* in the way I want you to behave, but not for the proffered reason. You would instead be so acting because I said so, or on a whim, or because the law demands it (cf. the choice phrase 'internalizing negative externalities'). You wouldn't be acting on the moral reason as such, which is just what is involved in *taking* my advice, or following my instruction. You would instead be *acting on command*, in the derogatory sense of the phrase.⁵⁹

⁵⁸ On the subsequent connection between *reasons internalism* and so-called *motivational internalism*, Williams writes: "Does believing that a particular consideration is a reason to act in a particular way provide, or indeed constitute, a motivation to act? If it does not, then we are no further along. Let us grant that it does – this claim indeed seems plausible, so long at least as the connexion between such beliefs and the disposition to act is not tightened to that unnecessary degree which excludes *akrasia*. The claim is in fact *so* plausible, that this agent, with this belief, appears to be one about which whom, now, an *internal* reason statement could truly be made: he is one with an appropriate motivation in his *S*. A man who does believe that considerations of family honour constitute reasons for action is a man with a certain disposition to action, and also dispositions of approval, sentiment, emotional reaction, and so forth." (1981, p. 107)

⁵⁹ By *command* I mean reliably eliciting certain behaviors, as we do from animals, by giving it certain cues. As in dressage, these cues or 'aids' may be highly complex and subtle – even though horses are manifestly not very bright. On the other hand, with our fellows, there is a distinctive mode of interaction, which issues in the person acting *on the basis of* the reasons which are offered in and as advice. Williams thus distinguishes between "...two possibilities in people's relations. One is that of shared deliberative practices, where to a considerable extent people have the same dispositions and are

Moreover, it is plausible to suppose that some people *do* lack the requisite psychological capacities to apprehend and subsequently act on certain pieces of distinctively moral advice. They lack, say, an ability to see certain of their fellows *as* fellow human beings, as evinced and partly constituted by their not caring about their suffering. The hard case husband appears to be an excellent case in point, vis-à-vis his wife. It seems fitting to say that he is deaf, in something close to a literal sense, to our moral advice about how to treat her properly. It would not then merely be browbeating, or a boring and pointless exercise, to keep on in urging him to start to treat her better. We might say of this man that there is “no *sense* in reasoning with him” (the ‘with’ bears emphasis too): the suggestion being that this phrase expresses something close to a literal truth, as opposed to merely registering our unwillingness to engage him. Telling this brute to treat his wife with the proper respect would be a bit like telling a stone not to fall on her head, or berating it for having done so: ridiculous, in other words. But the ridiculousness will also have a disturbing or even tragic edge to it when the impervious object comes in the shape of a person – i.e., the kind of being who *ought to be* (and will often be mistaken as being) responsive to our remonstrations. Moreover, there is a standing possibility that his profound moral defects are at least partly his own fault. The stone, in contrast, was not morally slack in allowing itself to harden and fall. Unlike the man, it *never* had self-control, or at least the capacity to develop it.⁶⁰

helping each other to arrive at practical conclusions. The other is that in which one group applies force or threats to constrain the other. The fiction underlying the blame system helps at its best to make a bridge between these possibilities, by a process of continuous recruitment into a deliberative community,” he writes (1985, p. 193). It would be instructive to compare Strawson’s famous distinction (1962) between the *interpersonal* and the *objective* stance, and their respective views about *blame* and *resentment* – but that would take me too far afield for present purposes.

⁶⁰ So I make no claim to the effect that we can’t *blame* the man whole-heartedly, to the extent that he may be responsible for his defective moral sensibility, if not for his current behavior. In this regard, I may diverge from both Williams and Strawson, at least according to some interpretations of their respective views about blame.

Time to gather together the various threads. Here then, at long last, is an argument for internalism about moral reasons:

1. Reasons contribute to prescriptive 'ought' claims.
2. Prescriptive 'oughts' presuppose that their recipient can follow the relevant prescription (e.g., instruction/piece of advice).
3. Following a prescription requires not only the logistical wherewithal – to act it out – but also the psychological capacities to act out of it.
4. Some people lack the psychological capacities to act out of certain moral instructions/pieces of advice.
5. Having these psychological capacities necessarily involves caring, at least deep down, about the locus of the moral concern.

Conclusion (by successive applications of *modus tollens*): people like the husband – who doesn't care about his wife's suffering, even deep down – don't have moral reasons to behave better. (Although they ought to be such as to have these reasons.)⁶¹

A word about premise (1) – as I've indicated, it is ultimately just a stipulation, but with an important caveat. If someone wants to eschew the notion of a practical reason as a consideration which informs apt advice, then they are welcome to do so in my book, and I do not claim they will still be pushed to be reasons internalists. But other difficulties may emerge, depending on the theoretical role which they want reasons to play. In particular, they had better not mean by 'practical reason' something like 'generic normative widget' (i.e., the minimal unit of practical normativity) whilst simultaneously claiming that the program of reducing all practical normativity to practical reasons is a sweeping theoretical reform. The idea that a whole is composed of its constitutive parts is not an exciting claim. So Parfit would appear to be on the horns of a dilemma here. On the one horn, internalism looms; on the other horn, his reductionist program looks to be essentially trivial.

None of the other premises are obvious truths, but premises (4) and (5) are the least secure at this juncture.

In fact, the anti-rationalist or broadly sentimentalist idea I find buried here in Williams about the nature of

⁶¹ How does my argument relate to Williams? I am not sure, since I am not sure that I understand Williams'. It is pretty cursory. See *n.* 57 for one suggested connection.

immorality deserves a thesis of its own. It is partly an empirical claim, and partly a claim about how best to think about our moral capacities (noting that the notion of a capacity has a built-in flipside: namely, that of incapacity). For, unlike the brute propensity to judge correctly – which would tend to merely be better or worse – we would naturally expect moral conceptual-cum-perceptual capacities to be entirely lacking in some people. To give some sense of the kind of person I have in mind here, I close with the following first-hand account of an abused wife, taken from Catharine MacKinnon.

Not a day went by I didn't think was my last as he totally lost control. He slept with a gun beside him every night as he promised he would kill me and then shoot himself if I didn't submit to his obsession of slavery and bondage and beatings during sex. I was raped 11 times between March '84 and November '86. I had four broken hands during my marriage, caused by my husband. I was put into the hospital in traction for two weeks due to a beating by him. I walked with a walker several months after that. When I was raped by Jerry, I was always tied to my bed. Tied where my legs were spread apart. He tied me with nylon cords and extension cords. I even got tied up while I was sleeping at times. He would then penetrate me with objects such as his rifle... (quoted in 2006, pp. 19-20)

In light of such testimony, the suggestion that there are brutes who are beyond the reach of reasons may strike the reader as not without intuitive force. Whether the internalist line can be ultimately maintained admittedly remains to be seen. But it is not to be summarily dismissed or disparaged, as some critics of Williams have done. That it is a viable and morally serious position is the intended upshot of this chapter.

Little now remains of the original debate about reasons – which is all to the good, it seems to me, since it was beginning to seem interminable. The residual philosophical action is elsewhere, I suggest, recalling the central issue which divides moral rationalists from sentimentalists. Namely, is moral advice of the kind that every rational being can respond to if they are just thinking straight, i.e., properly exercising general rational capacities? Or does it instead require distinctively moral capacities that are tied up with what we care about? My money's on the latter, although I don't think the standard sentimentalist construal of such capacities gets it right either. In the next chapter, I turn my hand to developing the beginnings of an alternative story about what such capacities might consist in – a story more conceptual than classically sentimental. The ultimate aim would be to replace the notion of 'responsiveness to reasons' with a notion of basic social awareness.

Chapter Three

RECONCEIVING ETHICS: DOING WITHOUT REASONS IN PRACTICE

OR: It's Just What We Do

*"You don't understand. I don't shoot at people. You see, it's against my upbringing. Call it etiquette, whatever."
~Ned Nederlander, The Three Amigos*

Doing something because it's 'the done thing' is not exactly a popular way of justifying an action these days. It might not even be thought to provide much of an explanation. Prevailing notions of social convention tend to make the modifier 'mere' seem virtually redundant. And codes of conduct are generally thought to be the province of professional ethics, rather than ethics proper. At best, such rules and regulations are thought to be useful rules of thumb, not sources of genuine ethical requirements in themselves.

I think that this is a mistake, and not a small one either. So, in this chapter, I want to make a preliminary case for the idea that some norm-governed social practices have pervasive and irreducible normative force for the people engaged in them. For reasons I'll shortly explain, I'll call this 'the idea of having an ethos' (and 'the ethos view' for short). The idea, in slogan form, is that practices can generate practical requirements for the participating agents. Or, to put it another way, we can be obliged to do something because – in the final equation – it's simply the done thing for someone in our social position. Accordingly, someone's having an awareness of their relationships with other people may be a source of genuine ethical insights. I actually think it may be a crucial source of insights, but I'm already out on a limb as it is.

I take it that the idea of having an ethos is pretty unpopular in metaethics and normative theory today. (The tenor of much of the literature on moral dumbfounding is another discouraging sign for someone who holds my views.) But it's a bit hard to say, really; the question seldom comes up. Elsewhere in philosophy, the social realm is getting increasing attention. First-order ethics and political philosophy is obviously chock-full of concerns along the lines of: 'What are we doing, what have we done, and what are we going to do about it?'

Theorists with a more metaphysical bent are asking questions like ‘What can we do together?’, ‘What are we doing together?’ and ‘Can we try this thing together?’ (Also: ‘Who, on earth, are ‘we?’) Even epistemologists have been getting in on the act lately. But moral philosophy of a theoretical stripe seems not so much to be anti-social as socially oblivious on the whole. Not many theorists argue that social practices have solely derivative moral importance; but this it is because it is tacitly assumed to be the case, or sometimes quite explicitly.⁶²

So it may be helpful for me to declare who I take my friends to be here right at the outset, by way of locating my own project. By the end of this paper, I will have incurred special debts to ideas to be found in three contemporary essays: John Rawls’ “Two Concepts of Rules” (1955), Cora Diamond’s “Eating Meat and Eating People” (1978), and Sally Haslanger’s “But Mom, Crop Tops are Cute” (2007).⁶³ Admittedly, none of these three papers would be most naturally filed under the heading of metaethics. Which goes toward my point. Anyway, I’ll flag the various connections along the way, as I shamelessly play bowerbird.

But why even bother to build a case for this allegedly unpopular view regarding this none-too-hot topic? (Why dig my own grave while playing in my own sandpit?) Because – to put it bluntly – it seems to me that we speak as if it’s true all the time, and that we’re often speaking truly when we do so. It’s easy enough to

⁶² Discussing the possibility that conventional moral rules (unlike rules of etiquette) give rise to reasons directly, Parfit writes: “On some views, it is we who create these moral requirements. That is true, I believe, only in limited and superficial ways. What we can create are only the particular forms that, in different communities, more fundamental, universal, and uncreated requirements take” (2011, p. 145). No real argument to this effect ensues. However, Parfit does grant that our relationship to people with whom we have ‘close ties’ (much like our resemblance to our future selves, he says) “can give us special reasons to be specially concerned with their well-being” (2011, p. 136). One of my concerns is that there seems little room to actually fit this idea into his overall picture. This is especially so since it is hard to see how ‘close ties’ could be anything other than manmade. But there is also the issue that, as a description of the reasons that spring from love, it seems to me to be seriously misleading. I discuss these issues in section 4.

⁶³ The same might have been said of Raz’s “Liberating Duties” (1989), had I not discovered it too late. I also found myself in sympathy with parts of Philippa Foot’s “Morality as a System of Hypothetical Imperatives” (1972). Bernard Williams will testify for the defense at various points as well. There are also people whose work has been invaluable inspiration in this endeavor, although the theoretical connections are less immediate here than they will be when it comes to the subsequent parts of the project which have more to do with moral perception than metaethics *per se*. See *n.* 95.

overlook the ubiquity of such talk. Given the current linguistic preference for ‘ought’ and its weaker cognates (i.e., ‘should’ and ‘have a reason’), we are liable to forget that not all normative claims come to the party outfitted in ‘ought’ garb. Even setting aside evaluative language and commands expressed in the imperative or prohibitive mood, there are many normative claims that aren’t typically or even naturally expressed in terms of what we ought to do. For example, we often speak instead of what we *have* or *have got* to do, *must* do, or *can’t* do. (I’ll come back to these locutions later on.) Another common – and even more neglected – class of statements comprises sentences like the following, where we speak of what is or is not *done* (or will or won’t be done) by people of a certain kind, or who occupy a certain social position. For example:

Boys don’t cry;
No son of mine does ballet – it’s for sissies.
Friends don’t let friends drive drunk;
We look after our own;
In this house, we do our chores without being told twice;
Everybody’s doing it.⁶⁴

These statements are an interesting hybrid: apparently descriptive propositions which typically function (i.e., are leveled and heard) as normative claims. More precisely, they *prescribe* or *proscribe* by means of an ostensibly factual claim about the behavior of people who fit some description.⁶⁵ At a first pass, the normative force of these statements is a sort of inescapability claim. The first invites the inference: if you’re a boy, then you don’t cry. The only way for a crying boy to get around this truth would seem to be to dry his tears. Otherwise, he is not a boy, or is at least a defective boy – i.e., a boy who isn’t ‘real’ or ‘true.’ (Compare the taunts: “What are you, a man or a mouse?” and “Real men don’t eat quiche.”) This being held to be a disastrous consequence,

⁶⁴ Compare Sally Haslanger’s examples of generics such as “Women are submissive,” “Blacks are violent,” and “Cows are food” (2010) which make a claim about the structure or character of the social world. (See also Leslie, 2008.) There is also the eponymous example: “But Mom, crop tops are cute” (2007). Haslanger points out that there is a temptation to deny all of these claims, holding them to be somehow defective. In section 6, I’ll turn to issues of these kinds.

⁶⁵ There are also second and third-person variants where the target audience is implicit, e.g., “You don’t hit a woman – you just don’t,” “You do not talk about Fight Club,” and “One does not simply walk into Mordor.” And there are minor variants of such sentences containing ‘always’ and ‘(would) never.’

the threat of *modus tollens* seems to work the prescriptive magic. That we hear it as a straightforward prescription shows how convincingly the trick comes off. But how does it work, exactly?

Or rather, how does it pretend to work? For obviously the claim that ‘Boys don’t cry’ is not true – that is, it expresses an illicit prescription. But other claims of this kind do seem to be true – that is, they express plausible normative claims. “Friends don’t let friends drive drunk” is one such example. Handing over the car keys to one’s inebriated pal is an action unbefitting of a true friend, generally speaking. So the slogan appears to express a valid prescription by dint of a facially descriptive claim.

What’s behind this trick? And why doesn’t it work with similar generic statements? For consider the sentences: “Men of that generation don’t cry,” and “American parents don’t serve their adolescent children wine with dinner.” These statements would naturally be heard, in most contexts, as stating straightforwardly descriptive facts. (They could be leveled in a rueful tone or an approving one, or in a perfectly matter-of-fact way.) On the other hand, notice that the normative reading begins to sound natural again given certain small adjustments – e.g., “Men of *my* generation don’t cry,” and “We don’t serve *our* adolescent children wine with dinner.” I will go on to argue that a neat explanation for these linguistic phenomena can be adduced if we allow that we commonly presuppose that the social norms of practices we participate in have genuine normative significance. Is this presupposition at all plausible? I will argue that it can be. Hence the seriousness of the problem when it turns out to be false – as in “Boys don’t cry,” for example.

So we see that, in one way, the unpopularity of the idea of conforming one’s actions to established social custom is entirely understandable. To again be blunt, established social custom can be *awful*. It will not help to restrict ourselves to practices where people *want* to play the game either. (Yeats: “The best lack all conviction, while the worst/ Are full of passionate intensity.”) And the fact that someone was ‘just following orders’ can hardly be held to render what they did morally unproblematic. We certainly can’t have it that what was *done* is therefore an event of neutral value for its casualties. On the contrary, the fact that grave ills can be wrought

by perfectly ordinary people, whose actions are readily comprehensible although not defensible as instances of rule-following, is partly why we condemn both the orders and the regime. People are naturally inclined to be much too obedient. (This is one reason why the Milgram experiments are so chilling, I suspect. Would we have kept upping the voltage too, as per the experimenter's instructions? The odds that the screams in the next room would have stayed our hand are probably not as good as we'd like.) But there is no problem in granting right at the outset that there are bad practices, such as gendered childrearing, as well as good practices, such as caring and protective friendships. There is also plenty in between, or having elements of both. Plausibly only good practices have a *prima facie* claim to justifying doing what's required without further ado. (I think it will end up being slightly more complicated than that, but not in especially surprising ways. For example, being caught up in an indifferent activity might warrant finishing what you've started, if other people are counting on it.) The primary challenge is thus to sort the wheat from the chaff and – ideally speaking – to provide an explanation of the criteria for doing so. I'll turn to these issues in sections 5 and 6, when I'll be in a position to make the problem maximally hard for myself, and then show how I might go about getting myself out of trouble.

The resistance to the idea of playing by the rules has theoretical origins too though. I suspect that my opponents are thick in the grass, but one camp is particularly visible at the moment. I have in mind reasons centrists, such as T.M. Scanlon and Derek Parfit, who hold that the practical reason is a fitting master concept in ethics. To recall the previous two chapters, the idea is that practical normativity ultimately comes down somehow to what we have reasons to do. Reasons are in turn supposed to be natural facts which count in favor, in an irreducibly normative sense, of the agent performing some action, or bringing about some state of affairs. This is not supposed to be an analysis or even an informative characterization: Scanlon and Parfit hold that no such account of reasons is in the offing.

This makes things tricky, dialectically. For, as we saw at the end of Chapter Two, theorists may use the term ‘reason’ in two very different ways (or slide back and forth between these different usages). On usage (1), reasons are not supposed to be theoretically interesting entities. They are just the minimal basic unit of practical normativity, as a matter of stipulation or by conceptual *fiat*. So to say that a person has (for example) a moral reason to do such-and-such is just to say that there would be a moral problem of some kind with his behavior if he didn’t. On usage (2), reasons mean something more than that. Reasons are the sorts of things which a person may be intelligibly reasoned *with*, say. Now there is conceptual room at the outset for moral problems with a person’s not doing such-and-such, even though he admittedly had no moral *reason* to do it. (And if there ends up being no room, the gap won’t be plugged by the concept of a reason *per se*.) So there is a real danger of equivocation on the term ‘reason’ here. And this danger is evidently inherited by the term ‘ought,’ given the widely accepted platitude that one ought to do what one has most reason to do.

So what’s to be done about all this?⁶⁶ I think it’s best for our purposes in this chapter to try to skirt these messy terminological issues. To do so, we need to designate some deliberately bland terms for basic concepts in practical normativity. We may borrow them from epistemology, especially given that the analogy between practical and theoretical reason is one of the main motivations behind reasons centrism (as discussed in the introduction). Then everyone will have the opportunity to voice their normative claims in a theoretically neutral way, whenever that is called for. To this end, we can speak about what would (partly or wholly) *justify* an action, and (for variety) what would *warrant*, *favor* or *rationalize* doing such-and-such. We will also say that we are *allowed* or *permitted* to do something insofar as there is sufficient justification for doing it, and that we

⁶⁶ One option here is to mark the distinction between (1) and (2) by using the locution ‘there is a reason’ for (1) and reserving ‘he has a reason’ for (2), as Bernard Williams once briefly suggested. (See *n.* 52.) But this can get confusing and unwieldy, and these two locutions have recently been put to other work. Anyway, I don’t think it much matters whether one goes for a theoretically neutral or a substantive notion of a reason, as long as one is clear. The other important point is that, if reasons demarcate the boundaries of normative terrain, as in (1), then it’s hard to see how they’ll be of much use in explaining its contours, as we might hope given (2). In other words, if reading (1) is in play, then reasons are effectively the *explanandum* rather than the *explanans* in metaethics. I suspect that Parfit in particular may be trying to have it both ways here, as I argued at the end of the previous chapter.

are *obliged* or *required* to do it insofar as one is not permitted not to. When an agent acts with some putative justification consciously in mind (whatever that means), and this suffices to explain what she did (ditto), then we can say that this is her *motive* or *rationale* for so acting. If not, then not. When her motive does not align with any genuine justification, then although she may *think* her action was justified, it really wasn't – even if all of her evidence pointed to the fact that it was, and so her behavior was perfectly understandable. Still, she wasn't actually allowed to do that, at least if some kind of justification is called for in order to be allowed to act in this way.

Given this advent of a universal language, the idea of having an ethos may be understood as a challenge to the conception of practical normativity which Scanlon and Parfit espouse. On their view, all justifications for action consist in there being some natural fact which favors the agent performing some action or bringing about some state of affairs. In my view, separating out the natural and the normative (the *favorer* and *favoring*, or the justifying *fact* and the justifying *relation*) will not always be conceptually feasible. For I will argue that there are important justifications for action which take a different form. Namely, the favorer is a moral-cum-social fact partly *constituted* by my being required to perform it according to some social practice in which I'm a participant. So ethical chromatography turns out to be an ill-fated idea, based on a false assumption about a certain contrast. In my view, sociality and practical normativity cannot be kept apart from each other, in practice or even in theory.⁶⁷

In the next section, I'll turn to clarifying this more holistic and socialized normative conception, together with the idea of having an ethos more broadly. I'll turn in section 2 to showing how the requirements of a practice might become practical requirements proper, on an analogy with promising. Following on from there, I'll work my way up to suggesting that this possibility is not merely theoretical, and that the idea of having an

⁶⁷ In this respect, ethics may be quite unlike epistemology. Or not, depending on how sympathetic you are to something along the lines of a naturalized epistemology, with a hefty side of holism. You have to be Quine, in other words.

ethos is worth buying into. This is partly because it enables us to neatly explain certain moral phenomena that would otherwise continue to be elusive and downright puzzling. To that end, I'll spend section 3 discussing three well-known puzzle cases: Jonathan Bennett's case of Huck Finn (1974), Michael Stocker's hospital visitor case (1976), and Bernard Williams' example of a husband who had 'one thought too many' (1981). I will argue that the idea of an agent effectively playing their part will help us to better understand what is going on in these cases. Not only does it help us to *explain* these agents' actions, but it also helps to explain why we have certain intuitive reactions not so much to what they did but to the *manner* in which they did it. The ethos view is also a theoretical boon, I'll argue in section 4, where I'll explore some connections between the idea of having an ethos and ethical intuitions, practical necessity, and so-called partial reasons. In any case, the proof of the pudding here is meant to lie entirely in the eating – the ethos view is not supposed to be a conceptual truth or anything like that. Even though I'll be making conceptual claims along the way, these are adduced in the hopes of eventually developing a view which you might find appealing and satisfying. I'll also spend some time, in sections 5 and 6, trying to show to a first approximation that things are unlikely to take a sour turn here.

But before we continue, it may be helpful to get the relevant theoretical contrast firmly in view here (and lest it be thought that I'm drawing a spurious line in the sand). So it is worth quoting from Parfit at length, who writes:

...it will help to distinguish between two conceptions of normativity. On the reason-involving conception, normativity involves reasons or apparent reasons. On the rule-involving conception, normativity involves requirements, or rules, that distinguish between what is correct and incorrect, or what is allowed and disallowed. Certain acts are required, for example, by the law, or by the code of honor, or by etiquette, or by certain linguistic rules. It is illegal not to pay our taxes, dishonorable not to pay our gambling debts, and incorrect to eat peas with a spoon, to spell 'committee' with only one 't,' and use 'refute' to mean 'deny.' Such requirements or rules are sometimes called 'norms.'

These conceptions of normativity are very different...[and] may conflict. When there are such rules or requirements, we may have reasons to follow them. But these reasons are mostly provided, not by the mere existence or acceptance of these rules, but by certain other facts, most of which depend on some people's acceptance of these rules. If we drive on the correct side of the road, we shall be less likely to crash. If we use words with their correct spelling and meaning, that may make us seem better educated,

and help us to be understood. When there are no such reason-giving facts, we may have no reason to follow some rule or requirement. We may have no reason, for example, to follow some fashion, or to refrain from violating some taboo. When I was told, as a child, that I shouldn't act in certain ways, and I asked why, it was infuriating to be told that such things are not done. That gave me no reason to do these things. (2011, pp. 144-145)

Actually, I think they might have done. This chapter is intended to explain why I think so, and to invite you to draw the same conclusions. (If I'm going to out myself as one of those shady characters who think that it is partly "*we* who make things good" then I might as well go for broke.) But failing that, I'd like to at least show that this position is more plausible, and has more conceptual resources at its disposal, than is suggested by Parfit above. I'll lean particularly on an analogy with games to begin with, and laws later on.

1: OH, THE GAMES WE PLAY

One of the difficulties with the idea of a social practice governed by certain rules is that it isn't quite the right idea. Not only is 'rule-governed social practice' pretty clunky, it is both over and under-inclusive in certain key respects. The examples to which I want to accord pride of place include *marriage*, *friendship*, and familial and social relations of other particular kinds. Our social practices with regard to animals provide further important cases, which we'll turn to towards the end. I am also of the view that the social roles that are associated with – one might even say constitutive of – gender and race belong in this category too (see Haslanger, e.g., 2000). Defending this claim in full generality would require more space than I can give it here. But we can at least include particular social practices such as gendered childrearing, patriarchal marriage, and slavery on the list. Somewhat less ethically interesting, but conceptually illuminating nevertheless, are the practices associated with being a member in clubs and societies, and with participating in games or sports governed by a set of rules and standards.

But the idea of a rule-governed social practice is less than ideal partly because it naturally encompasses social phenomena that are quite a bit more diffuse than what I have in mind. The practice of riding on the T in

Boston is a social practice governed by certain rules (e.g., no smoking allowed), but it's not a natural inclusion on the list – or at least, I don't want it to be on my list. This honor is meant to be reserved for cases where the rules seem to be at least partly constitutive of the practice. This doesn't mean that the rules can't be changed. It is rather, as Rawls puts it, that engaging in such practices *means* following the rules, whatever they happen to be at the time. For example, if one neglects to follow the rules in baseball, then one would be in imminent danger of no longer playing the game at all (as we might say – “he seems to be off playing with the fairies”). The distinction between not playing the game properly, and not playing it whatsoever, is thus a difference in degree, not in kind. Rawls writes:

To engage in a practice, to perform those actions specified by a practice, means to follow the appropriate rules. If one wants to do an action which a certain practice specifies then there is no way to do it except to follow the rules which define it... This point is illustrated by the behavior expected of a player in games. If one wants to play a game, one doesn't treat the rules of the game as guides as to what is best in particular cases. In a game of baseball if a batter were to ask ‘Can I have four strikes?’ it would be assumed that he was asking what the rule was; and if, when told what the rule was, he were to say that he meant that on this occasion he thought it would be best on the whole for him to have four strikes rather than three, this would be most kindly taken as a joke. (1955, p. 164)

Now, as Rawls effectively went on to say, it is certainly a concern if some of the rules are unduly onerous or too easy on some of the players, or have been shown to detract from the very point of the exercise. One might recommend that the practice be reformed in certain ways on such bases. For example, in baseball, the infield fly rule was adopted to prevent ‘ungentlemanly’ conduct on the part of the fielders (which, for complicated reasons, made it too difficult for the runners to strategize). Similarly, it was decided once upon a time that a batter hit by the pitch would get an automatic walk to first base, rather than it counting as a ‘ball.’ This was due to certain perverse incentives which would otherwise persist. So a practice-based conception of rules need not be indifferent to their consequences. Practices can and should evolve; the rules are not set in stone.⁶⁸ We may even propose scrapping a practice entirely, if we think that this game should not be played

⁶⁸ It may even be possible to bend the rules a little if little is at stake (e.g., in a Little League game). One might say: “It doesn't matter. We aren't really trying to play a proper game here.” Compare the idea of ‘pub rules.’ I owe this point to Julia Markovits, and am grateful for the many extremely helpful discussions we've had about these issues.

anymore (e.g., football is simply getting too dangerous; the equestrian three day event has become cruel). But all of this is not in spite of but precisely because engaging in the practice and being subject to its rules are not conceptually distinct affairs.

Another thing to be said about a set of rules which constitutes (or, I will also say, governs or regulates) a practice is that they can have a recursive structure. That is, sets of rules can gain additional expressive power by employing rules that refer (implicitly or explicitly) back to the set of rules itself. For some examples, it might be a rule of the game that, in order to play it properly, you have to play to win. Obviously there are rules to the effect that you can't throw the game, but this is often more for the bookies' sake. But there is also a rule in NFL football to the effect that the team as a whole is required to make a reasonable effort to win, even if losing would be a strategic advantage on account of the schedule – which fans and bookies alike could be expected to factor in. There are also games in which, in order to be a good sport about it, you have to *want* to abide by the rules, even when you could get away with breaking them. And there are rules which are tantamount to exclusivity or non-compete clauses in contracts: i.e., you shall not pick up the bat for another team. (The original version of which read: "I am the Lord thy God, and thou shalt have no other Gods before me.") And of course a set of rules might specify that, if you break the rules, then certain consequences will ensue, or you will be disqualified from the game. Finally, a set of rules might specify that surrendering is not an option. In other words, there are games where one isn't *allowed* to just throw in the towel, or stalk off the field in protest. To withdraw from the game is effectively to lose.

Since this issue will crop up frequently in what follows, it is worth noting that games can be inescapable for their players in one of two general ways: (1) the player cannot *help* but play the game somehow (or at least to want to win), and (2) the player is not allowed to leave the field in any case – the rules specify that he has to keep going if he possibly can. Suppose the player moves to skulk off the field regardless, and we want to get him back out there. We probably can't drag him back, even setting aside the little matter of his rights. For if

this is one of those many games which has to be played willingly, then forcing him to play would be pointless, even borderline incoherent. We might be able to cajole or threaten him into playing; we could promise him a bigger (or smaller) salary. But what *guarantees* that we can get him back into the action? Seemingly nothing short of brute force. And it might be essential to the game that he be playing under his own steam. In which case the answer to our question will become: nothing, or so it seems.

But we still want to be able to *say* something to him or about him. If (1), then we are pretty much home free. He is going against his own nature, or letting himself down badly, or there is no getting off the island, or he is bound to get himself tied up in knots – something along those lines. But in general, it's a bit hard to see why (1) would actually be true. Maybe if he'd have to swim through shark-infested waters to exit the stadium, or walk over hot coals, or he'd been in training for this game his whole life, then we'd think him rash or even crazy not to heed our warning. There is also the possibility that he has long fantasized vocally about being the next David Ortiz (a real 'team player' as well as a fine one). So he is in real danger of lapsing into incoherence here: he can't have it both ways, he is going to have to make his choice. Those are some of the reasons why (1) may sometimes hold. But ordinarily, if he walks away without a backwards glance, then he seems to have given us pretty good evidence that (1) is in fact false.

In these instances, we may have to retreat to inescapability mark (2). Here the situation is dialectically more complicated. We could point out that he loses either way, so he might as well get out there and give it his best shot. And other people are counting on him, so he *cannot* let them down. These are certainly things we can say, and they may sometimes even be effective. But neither strategy is anything like foolproof. He might be loath to make the slightest effort, and no longer care a whit about winning or letting his teammates down (if this is impossible, return to (1), above). "You're taking the easy way out!" we might say. "Yes." he replies agreeably. Or: "I got my reasons," said enigmatically, with no subsequent elaboration. (Compare: "It is not irrational to prefer the destruction of the whole world...") Still, perhaps he *should* care somehow about the

outcome – perhaps we can see, as he can't, that this is really the match of his life. And, perhaps more to the point, he is letting other people down. Is he making a mistake then, in simply walking away? Not necessarily, or not exactly, and at least not by his own lights. And, for reasons we saw above, this might be dialectically important. It is more that he himself is out of line, out of order, or off with the fairies. He may be a quitter and a spoilsport; he is certainly acting like one. So, if what we had hoped here was to be able to yell at him for being hopeless or irresponsible, or shake our heads ruefully to ourselves, then we could walk away feeling furious but satisfied, having said our piece. But insofar as we had hoped to show him the error of his ways, or put it to him in a way he could understand, we may be out of luck. He does not even seem willing to listen. Maybe we'll try again tomorrow, hoping that he is merely having a "Manny moment."⁶⁹

This scenario should give us a preliminary sense of the different kinds of inescapability that might be at issue here. Although it's admittedly a tendentious version of events: for it's also meant to show how someone like me, who holds an internalist position about reasons, may still have a notion of inescapable normative requirements in her swagbag. (Although notice that we were able to paraphrase reasons claims away here entirely.) It is just that, as internalists, we think that these requirements may be entirely beyond a person's grasp (see Chapter Two). The claim is essentially that some obligations do not move some people, and that this means that certain sorts of normative criticisms no longer make sense when leveled against them. They would at least have to be able to register. This sort of thing is not unprecedented: as we'll effectively see in the next section, there are some communicative acts which need to make an impression on their audience in order to stand a chance of the proposition expressed coming out as true. (As in performative utterances – i.e., the things we do with words.)

Anyway, in what follows, I want to continue to draw on Rawls' analogy between normative requirements and the rules of a game. This might sound dubious or off-putting, since games are often envisioned as basically

⁶⁹ <http://bleacherreport.com/articles/37109-the-view-from-landsdowne-street-my-favorite-manny-moments>

trivial. But I suggest it is not the game itself that is trivial; it is rather that the stakes are low. Still, we do need to be more specific here. We play all sorts of little games – e.g., Wittgenstein’s game of ring-a-ring-a-roses. The social practices I want to focus on are analogous to games that (a) are multi-player, and (b) have rules which are rich in complexity and nuance. (See Haslanger on ‘thick social practices,’ 2003.) To see what I mean by this, consider my first example of marriage. I don’t mean the legal institution of marriage, or at least not in the first instance. The marital contract is in some sense the state’s attempt to recognize and regulate the relationships which I’m more interested in – namely, common law marriages, de facto partnerships, and intimate domestic relationships in general. There is a social notion of marriage which is partly independent of the legal notion, although the two are intimately entwined, and would likely march in step ideally. This is why it can make sense for two newlywed men to say: “Finally, the state has recognized our marriage!” Or for a straight person to say: “I refuse to get married until all couples who are *essentially* married enjoy the same right.” Or for an anthropologist to say: “Marriages based on love are a fairly recent social phenomenon.”⁷⁰

Relationships of this kind are generally governed by a complicated nexus of expectations, rituals, intimacies, and standards of behavior on the part of the participants. These norms (which I’ll use as a covering term) tend to have a firm footing in communal expectations about what this kind of relationship does and should involve (hence the notion of a ‘1950s marriage,’ which it is still possible to enter into today). (Raz, 1989, is instructive on this point.) Norms are the sort of thing about which the partners might agree – “yes, that’s a part of our marriage.” (One partner’s norm may be the other’s mere regularity; this can lead to squabbles and misunderstandings.) These norms generally have a history, both personal and social. But the norms need not be rules *per se*, or at least not in the usual sense. For one thing, love as a feeling cannot be commanded, as Kant famously pointed out. But it might still be expected. On a more humdrum note, it is not exactly a *rule* of Mr. X’s marriage that he takes care of the grocery shopping while Dr. X handles routine car maintenance; in a

⁷⁰ Compare religious marriage, i.e., being married in the ‘eyes of God.’

sense, it's just what they do. Even norms with more explicit normative content – it's his *job* to do the dishes – may not be aptly called rules, partly because they apply to a body politic of two who double as the ruling bodies. The notion of 'house rules' thus tends to be pretty jokey.

It will be useful to have a covering term for the set of norms which governs a social practice of the sort I want to consider as possible sources of obligations. Rather than speaking of charters or constitutions, I suggest we turn to the somewhat old-fashioned idea of an 'ethos.' This English term, a borrowing from Greek, is meant to reflect the fact that the relationship and social roles I want to focus on have a certain *character*. (It is also the ancient root of the word *ethics*, which seems like a good omen.⁷¹) Such practices are not just governed by a code of conduct; their tenets, principles, or creed is generally informed by a certain *spirit*, which may also be an integral part of what makes the practice what it is. Families and relationships generally have an ethos, but so do many corporations. An ethos may be more or less formal (although informality shouldn't be equated with lack of strictness). An ethos may be largely or wholly implicit. For example, company ethos might dictate that you go out to after-work drinks on Fridays, even though nobody would say a word to the new girl as she ducked out week after week. *They* might not even be able to pinpoint what the problem was, exactly. Still, her colleagues' lips would get more and more tightly pressed together as the weeks went by. The point is that an ethos must be abided by to the extent that the person subject to it is living up to the ethos, or playing their part properly – or, eventually, playing it at all. And the consequences of not playing the game can be pretty severe. This might be partly what makes it *true* that Friday drinks are part of the company ethos, even though it's hardly stipulated in the handbook. Perhaps the new girl couldn't even have been expected to read between the lines before becoming the office pariah, or being summarily fired for failing to 'fit in.' (She just wasn't working out, they all felt. And this business is like a family. Apart from the fact that its members are chosen and disposed of with great comparative ease.)

⁷¹ The original meaning is apparently 'customary place' or, charmingly, 'the habitat of horses.' The Latin equivalent is *mores*.

Other plausible examples of social entities with an overarching ethos include such behemoths as organizations, movements, and (sub)cultures, be they political, ideological, or geographical in flavor. These social entities comprise collectives or communities rather than anything personal and particular enough to be happily called a relationship. I'll be focused on relationships proper for the most part, but there is no harm in allowing the idea of an ethos to apply pretty expansively. And because we have fairly robust intuitions about what counts as a (more or less robust) ethos, I think there's little harm in leaving this notion loosely characterized rather than looking to tighten the account at this juncture. The same can be said about the issue as to how ethos-governed practices are to be individuated: we have at least a rough idea of where one ethos ends and another begins, or where there are overlapping ethes in play. Drawing boundaries is less important than it might be partly because I have no interest in claiming that *all* practices governed by an ethos have genuine normative force for its participants. (To anticipate, I will argue in section 5 that this can only be the case if the practice has some kind of independent *value* for its participants, at least under decent social conditions. At a minimum, it must not be positively bad for them, assuming that they're not being a bad sport or playing other dangerous games. Still, for reasons I'll explain, it is not a reductionist account.) Equally importantly, as an internalist, I believe that an ethos has normative force for a participant only if she *cares* about playing the game, or is called upon to care for one reason or another. (Note that this doesn't mean that any old game she wants to play will do: the internalist claim is a necessity claim only, as discussed in Chapter 1.) The upshot is that we do well to make a beginning here by focusing on the ties that bind.

But what about this idea of an ethos being, at some level, just what one *does* as a participant in a certain social practice? This stands in need of further comment straightaway. After all, we were taught at our mother's knee that what people happen to *do* is one thing; what people *ought* to do quite another. But this distinction, crucial as it is, becomes trickier and more delicate when it comes to practices which one can become *committed* to continuing to participate in. Suppose I've led other participants to rely on the fact that I'll continue to do the 'done thing,' in keeping with the ethos. Or I'll continue to perform my assigned role, rather than going against

the grain. Then – all things being equal – the thought might naturally occur to me: *well, I've made my bed, now I have to lie in it*. Or at least I have to negotiate and coordinate with my bedfellows before I make a move. If this thought turned out to be accurate, then this plausibly might – absent defeaters – turn the ethos governing the social practice into genuinely ethical claims on me, as a participating agent. It remains to be seen if and when it is in fact accurate. But, in the next section, I examine how this might work in theory.

2: PRACTICAL MAGIC

I suggested just now that the practical magic that might breathe normativity into social norms (thus raising them beyond the level of mere conventions) may have to do with the fact that other people are counting on you to abide by a certain ethos, in virtue of your apparently *taking yourself* to be bound by it. In other words, the normativity of social practices might stem from something that resembles, in generic form, a promise. This issue is doubly worth taking some time to investigate, given that the normativity of promising has itself been a source of controversy lately. On one prevalent view (most recently due to Rawls but traceable to Hume's writings on 'artificial virtues'), the obligation to keep one's promises depends essentially on there being a social practice of promising. In a chapter of *What We Owe to Each Other*, Scanlon argues against this way of thinking about things. He writes:

I do not doubt that there is such a thing as a social practice of promising, which consists in the fact that people accept certain norms, which they generally follow and expect others to follow. The question is what role this practice plays in generating obligations to keep one's promises. According to the standard institutional analyses, these obligations arise from a general duty to comply with just and useful social practices. I will argue, however, that the wrong of breaking a promise and the wrong of making a lying promise are instances of a more general family of moral wrongs which are concerned not with social practices but rather with what we owe to other people when we have led them to form expectations about our future conduct. Social practices of agreement-making, when they exist, may provide the means for creating such expectations, and hence for committing such wrongs. But I will argue that these practices play no essential role in explaining why these actions are wrongs. (1998, p. 296)

Scanlon builds his case by telling a just-so story about a stranger in a strange land who inadvertently throws his good hunting spear over a raging river. An aboriginal appears on the opposite bank, having haplessly

thrown *his* boomerang across the turbulent billabong as well. So my weapon is at his feet, and his weapon is at mine. Scanlon points out that if I manage to induce him (through gestures and grunts or what have you) to believe that I will return his boomerang if he returns my spear, then this is much like a promise. That is, failing to hold up my end of the bargain seems just as wrong intuitively as if I'd promised to return it.

It is much like a promise – too much like one, in fact. That is, on the most natural way of mentally filling in the case, it *is* a promise in effect. Admittedly, some people insist on reserving the term 'promise' for cases in which the promiser has uttered the local equivalent of the words "I promise" to the promisee. But almost everyone would agree, I take it, that giving one's word is about as serious a commitment as promising *per se* – and breaking it is wrong for much the same reasons. Giving one's word is no more than a metaphor either: not sharing a language is only a contingent barrier to making a commitment to another person. It doesn't matter if one says it or whistles it or pours it in somebody's ear. If I induce someone to count on my doing something in virtue of the fact that *I now consider myself bound to do it*, then it is tantamount to a promise – a swamp promise, if you like. (Why aren't contracts promises? Because no such assumption applies. The default assumption in law is that you *will* break the contract if it is in your best economic interests to do so; hence the notion of efficient breach. And the contract may also serve as a *post hoc* guide as to how that situation will be adjudicated and remedied.⁷²)

It is crucial to promising that I represent myself as feeling bound to keep my word to you – at least indirectly, via my own sense of honor (as in "I'm a gentleman; my word is my bond"). So, if lying promises are promises at all, it is because this is how I've succeeded in representing myself, even though I've misrepresented my true intentions. Nevertheless, breaking a promise and deception of this kind are different kinds of wrongdoing. Although they may be co-present in instances of lying promises, these wrongs can easily come apart. Consider

⁷² Another view might be that a contract *is* a promise, but only a contingent one. As in: I promise to do this thing so long as the benefit I expect to get from you outweighs the cost of my performance.

Scanlon's swamp case again. If I merely try to get the native to *expect* that I'll do such-and-such, then this might be manipulative and deceptive, but it is not the moral equivalent of a promise. What makes it natural to think of the situation as involving an agreement between two people is imagining it as having an open social character – i.e., we meet face to face, and form a spontaneous pact under conditions of relative transparency.

But imagine instead that the situation was such that the native couldn't see me initially, my having concealed myself behind a handy gumtree. There I have the time and wherewithal to rig up a cunning disguise. Suppose I subsequently get him to believe that I'll throw the boomerang back if he throws back the spear by pretending to be a mischievous minor deity. Like many poltergeists, I have no sense of obligation: I insinuate that this is just how I happen to roll, or how the spirits tend to move me. In this case, nothing like a promise would have been made (not even a lying promise). Rather, I would have given the man something closer to a mischievous threat, or perhaps even a threat if I seem particularly menacing. And a threat is not a promise.

Moreover, creating a sense of expectation is not necessary to promising any more than it is sufficient. I can promise to do such-and-such even though the promisee never takes it seriously, so bad is my track record in these matters. It seems that I just have to have made some honest attempt to get her to take my word for it, via our mutual recognition that I have in fact given her my word. Thus, I cannot make a promise to someone which I deliberately keep a secret from her.⁷³ Even if the illocutionary force is lacking for contingent reasons (e.g., she doesn't hear what I said), then I am called upon to repeat myself. In other words, uptake seems required. Suppose now that she does hear and understands my utterance, but regards it with extreme skepticism (i.e., my utterance has the illocutionary but not the perlocutionary force of a promise). I take it that I've succeeded in making a promise nevertheless. All things being equal, we ordinarily think that I now am under a *pro tanto* obligation to keep my word – however surprising it would be to the promisee if I actually

⁷³ What about promises to oneself though? They're a spandrel at best I suspect, and possibly just a misnomer for what would be better termed a *resolution*. See Holton, 2009.

followed through with it. (The promisee might arguably not even be *around* to care about my keeping my word any longer, having subsequently breathed her last.) Still, a promise is a promise and must be kept, or so we're taught as children. In other words, one of the norms governing promises as we know them is that one is not allowed to break them.⁷⁴ (This is also the basis of a familiar joke. Although rules can certainly be bent, the idea of them being *made* to be broken is a mildly amusing idea the first half dozen times you hear it. Compare the prescriptive liar: "The only rule is: there are no rules!")

So there seem to be three elements in my making you a promise to do such-and-such. (1) I have to convey to you that I *will* do such-and-such, because (2) you can *count* on me to do it, because (3) I *told* you I would, didn't I? (see 1). (Cf. "I may have to hold you to that.") Note that your expectation in (2) is normative, not predictive; you have an automatic complaint against me if I violate these expectations, *ceteris paribus*. (Compare the 'ought' of expectation, as in the neutral speculation: "He ought to have left by now – I guess he won't be there if I call him at home." Such usages might plausibly be a spandrel of genuinely normative expectations.) And the expectation here seems to be the flipside of the coin which makes the promise an obligation on my part.

I'm sure this characterization of promising needs plenty of adjustment and finessing. But in lieu of a starting point along these lines, I think someone would have to admit: "I don't really believe in promises. I know we *talk* about such things. But I'm just not buying the whole idea."

Unsurprisingly, Scanlon recognizes all of these problems from the get-go. Regarding the boomerang story, he writes:

⁷⁴ This is assuming that she doesn't go so far as to 'officially' discharge me of my obligation – which is different from her simply not *caring* about my making good on the commitment. The latter might make a moral difference insofar as I wouldn't have to worry about upsetting her unnecessarily. But it would make no moral difference vis-à-vis the promise itself, unless I can presume that she'd release me under the circumstances (on analogy with presumed consent).

Here, it might be suggested, a practice of agreement-making is covertly presupposed. For in the absence of such a practice, what reason would the stranger think me to have to return his boomerang once I had recovered my spear? And how could he have a reason of the right kind? If the reason he attributes to me has nothing to do with the thought that I will be moved by a 'sense of obligation,' it may seem that what I have done could not be the same kind of wrong as that involved in breaking a promise. I believe that it is the same kind of wrong. (1998, p. 297)

He then effectively goes on to say so: i.e., that deceptive manipulation and promise-breaking *do* involve the very same wrong. But this is surely very implausible. They might be as bad as one another, but they look to be horses of a different color – i.e., wrong for different reasons, not to mention wrong at different times. If I had every intention of returning the boomerang as per our agreement, but wig out at the last moment, I would not have done anything wrong up until that point. Not so with pretending to be a demigod. But Scanlon initially fudges by only considering promise-breaking where one's intentions were cynical from the beginning. Promises can be broken in other styles though. (In then essentially going on to recognize this, he moves for an elaborate reconstruction which again – it seems to me – brings in the implicitly normative notion of a promise.⁷⁵) So Scanlon's account of the normativity of promising is itself not very promising, even though he is ostensibly not in the business of offering a revisionary account. (He introduces the discussion by writing that Rawls' practice-based account "no longer seems to him to provide the best account of the matter.") This leads one to suspect that there must be something more at stake here, theoretically.

And indeed there is. The idea that the normativity of promising owes partly to irreducibly social facts – i.e., the fact that X gave Y her word to do such-and-such – is the thin edge of the wedge for a reasons-based

⁷⁵ Namely, Principle F, which reads: "If (1) A voluntarily and intentionally leads B to expect that A will do X (unless B consents to A's not doing so); (2) A knows that B wants to be assured of this ; (3) A acts with the aim of providing this assurance, and has good reason to believe that he or she has done so; (4) B knows that A has the beliefs and intentions just described; (5) A intends for B to know this, and knows that B does know it; and (6) B knows that A has this knowledge and intent; then, in the absence of special justification, A must do X unless B consents to X's not being done." (1998, p. 304) Scanlon then goes on to say that "Principle F is not just the social institution of promising under another name." But, insofar as I understand it, it seems to me that it *presupposes* this notion. The key question is: what exactly is meant here by the term 'assurance'?

conception of practical normativity.⁷⁶ If what I've argued is along the right lines, then one has to give one's word – i.e., to represent to my promisee that I consider myself bound by the promise – in order to make a genuine promise – i.e., to be bound by it in fact. Thus, in saying “I promise,” we can effectively make it so. Or, in rougher but simpler form: my obligation comes down to the fact that I promised. None of this will be terribly surprising in the wake of J. L. Austin's brilliant account of these sorts of things. But it is important to remember Austin's point that one of the things that can apparently be done with words, given the appropriate social backdrop, is to change our normative situation. (Promises being one of his earliest examples.) We needn't follow John Searle (1964) in saying that an 'ought' can therefore be derived from an 'is.' The moral as I see it is that some 'is's must already contain 'oughts.'

Redirecting the lesson back to the question which originally motivated this side investigation into promising, we see that promising might provide a good first-pass model as to how the norms which comprise an ethos might come to have genuine normative force. Namely, they might be constitutive of a practice in which one is bound to hold up one's end of the bargain once things are underway. In the last section, we examined in schematic form some of the reasons why this might in turn be the case (e.g., one is not allowed to simply walk away from games of this kind). To be more specific now, in cases where a promise is given (freely, I mean), we are plausibly bound to do what we said we would because the promisee can expect us to keep our word, not as a behavioral prediction, *contra* Scanlon, but as a normative matter. Obviously that still doesn't get us very far though. What we will end up saying, I suggested, was that this normativity will ultimately come down to the moral-cum-social fact that we promised, and that promises are made to be kept.

Now, as we also saw, another element of promising is to have made an explicit, public commitment to the promisee to that effect. So it is crucial for the analogy to work that we are playing a game that involves other

⁷⁶ There is thus a tension between Scanlon's reasons-centrism and his contractualism. Alternatively, we might say that his view is not really contractualist in the classical sense; he himself expresses misgivings in the introduction to *What We Owe to Each Other* about sticking with the name he originally chose for the view.

people, not just following any old code of conduct that we made up off the top of our heads. (Compare the notorious idea of a private language.) The norms governing fundamentally public practices like games, laws, promising and possibly language thus seem interestingly disanalogous in one way from the norms governing epistemology and aesthetics, for example. Epistemologists and aestheticians are not the thought police; one's beliefs and tastes are in some sense one's own affair – although they may be as wrong as can be, and right only insofar as other people are around to help us see the light. Ultimately though, the responsibility falls on (in) the individual. Whereas practical philosophy is in some sense oriented around the fact that we're living in a society, and we have to find some way of all rubbing along together. This is why we generally approach the subject of ethics with a mixture of hope and fear, rather than sheer intellectual curiosity.

Anyway, a natural variant on the theme of promising would thus be cases in which the commitment is likewise public, but isn't made to any particular person, or has a more open-ended content. This is where the idea of an ethos naturally fits back in. There may also be cases where we find ourselves with a commitment to a particular person which was acquired in a more haphazard way. Friendship seems like a plausible case in point.

To see this, let's return to the linguistic puzzle I started with, which will also help to summarize and apply the key points in the story so far. When the ad campaign "Friends don't let friends drive drunk" came out, it wouldn't even have occurred to most audience members to take its central slogan as an empirical or descriptive claim. What explains this fact? I think it is this: what it is to be someone's friend is at least partly a matter of the norms which one becomes subject to in actually becoming their friend. This is rarely something that we consciously decide upon – in fact, this might not be strictly speaking possible (as opposed to consciously deciding to try to *become* or *make* a friend). But even if it develops gradually, a friendship is a relationship which one enters into willingly, by very definition. And friends are obliged to help one another, console one another, take an interest in each other's affairs, and refrain from undermining each other's

interests, etc. (If you reject these ideas, then you don't really believe in friendship in the customary sense.) For *what it is to be* a friend involves, as a constitutive matter, helping one another, consoling one another, taking an interest in each other's affairs, and refraining from undermining each other's interests, etc. In other words, acquiring a friend and the attendant responsibilities are not even notionally separate events.⁷⁷

What might make them genuine as opposed to merely supposed responsibilities? What, in other words, might make the concept of friendship credible? I think that, if anything, it would have to be the fact that a friend is also required to stay true to her friends, and to keep up her end of the filial pact by doing what a friend does. This is not supposed to make the practice self-justifying; in other words, we have yet to see what *makes* it a valid pact (we'll get there). But many of us have strong intuitions that it is in fact valid. I won't try to drum these intuitions into anyone in this chapter (or anywhere else for that matter). In the next section, I'm just hoping to show how a friendship's characteristic moral claims might be vindicated by those of us who are already inclined to view them as licit. Call it a transcendental argument. My intended conclusion will be that, if we owe our friends anything in particular, then the inbuilt norm of loyalty is very likely the reason why.

There are a plethora of expressions which testify to this implicitly normative notion of friendship being very much the ordinary one. "It's just what friends do," we might say, in order to explain our insistence on going to great lengths on a friend's behalf – or, even more suggestively, by asking the rhetorical question: "What are friends for?" Other expressions capture the fact that we typically feel bound to stick by our friends, even when doing what that requires comes at a significant personal cost. For example: "A friend in need's a friend indeed," and "Fair weather friends are no friends at all." The slogan "Friends don't let friends drive drunk" deftly calls attention to both a specific provision of the pact of modern friendship and the final catch-all

⁷⁷ There are other things we call friendships which are considerably less demanding. But they're plausibly parasitic on the real deal; e.g., a 'Facebook friend.'

clause: you don't abandon your friends.⁷⁸ So how is it that these facially descriptive statements can function as normative claims? I have suggested it is because concepts like friendship, being normative, import normative considerations into the conversation as a matter of course. So, if I am right, these claims typically not only function as normative claims, but actually *are* normative claims, semantically speaking – at least relative to the customary normative meaning of the concept of a friend. When one isn't part of the practice, these statements may rather be taken to be merely anthropological reports about how *they* feel obliged to do things, rightly or wrongly.⁷⁹ But most of us have friends, and so feel the force of these claims ourselves. If true, they would apply to us, for we are part of some relevant 'we.'

It remains to be seen whether these claims are ever in fact true. That is, is it ever the case that the norms comprising an ethos can have genuine normative purchase, in virtue of our being bound to play our part in the resulting social practice? We've seen how this might work in theory, by examining the nature of promises; but does it work in practice? I'll move for a positive answer by considering three flashpoints in contemporary ethics which I believe reveal a gap in the dominant normative conception – a gap which I'll suggest the idea of having an ethos is well-positioned to fill.

3: IN THE THICK OF THE ACTION

I'll start with the case of Huckleberry Finn, first discussed by Jonathan Bennett (1974). Huck believes, wrongly, that he ought to snitch on Jim, a runaway slave. The two have become companionable, floating down the river together in their flimsy raft (a thinly veiled moral metaphor: i.e., 'the same boat'). Still, Huck is increasingly plagued by his conscience – his sense of having been complicit in the great crime of stealing Jim from his "rightful owner" (2010, p. 103). From Huck's perspective, his eventual failure to turn Jim in is just

⁷⁸ Hence, upon being let down one time too many by a supposed friend, someone might say: "I don't feel like we're even friends anymore. In fact, maybe we never were."

⁷⁹ Cf. Williams on thick moral concepts which one does not share (1985, p. 144). He effectively builds in the distance by picturing the linguistic community in question as a cohesive and remote tribe, so one can talk about 'What *they* call F.'

that: a failure. But he ends up doing just what the situation calls for, demonstrating great cunning in protecting Jim, not to mention considerable courage. So he ends up doing the right thing, and not for the wrong reasons, exactly. A more accurate way of putting it would be that he lacks any real grip on what he did and why. He infers that he shielded Jim from the slave-hunters only out of weakness, rather than anything that could properly be called a moral motive. And indeed there was nothing explicitly moral in his thinking at the time – not even any obvious signs of moral recalcitrance or ambivalence. Still, Huck is not suffering from a case of straightforward *akrasia* in which the will happens to be at odds with the dictates of morality (such that the weakness ends up pointing the agent in the right direction, as luck would have it). While it has been dubbed a case of *inverse akrasia*, Huck's case is psychologically and morally more complicated than a case in which I erratically fail to follow through with some bone-headed or dastardly plan I've devised.

So it's all a bit puzzling. One thing to say at the outset is that, insofar as we are loath to condemn Huck and might even want to praise him, it has to do with what we might loosely call his instinctual grasp of what to do. Admittedly, Huck behaves well in spite himself – but it's not by accident, either, since his spur-of-the-moment *volte-face* seems to stem from his inchoate but nagging sense of the true moral contours of the situation. (Or at least, we can fill in the case this way imaginatively, although I think it also happens to be the right reading of the novel.)

Setting aside for the moment the question of the degree to which we are warranted in praising Huck's motives, there is a prior question about how to even characterize or explain what he did. The literature is interesting insofar as it tends to miss the mark on this score. Bennett himself, for example, has it that Huck ignores his conscience, and instead acts on his sympathies. But it is striking just how *unsympathetic* Huck actually is to Jim at this point in the novel. As the time to tattle draws near, never does a tender moment compete with such bitter, self-righteous ruminations as:

Jim talked out loud all the time while I was talking to myself. He was saying how the first thing he would do when he got to a free State he would go to saving up money and never spend a single cent, and when he got enough he would buy his wife, which was owned on a farm close to where Miss Watson lived; and then they would both work to buy the two children, and if their master wouldn't sell them, they'd get an Ab'litionist to go and steal them. It most froze me to hear such talk. He wouldn't ever dared to talk such talk in his life before. Just see what a difference it made in him the minute he judged he was about free. It was according to the old saying, 'Give a nigger an inch and he'll take an ell.' Thinks I, this is what comes of my not thinking. Here was this nigger which I had as good as helped to run away, coming right out flat-footed and saying he would steal his children – children that belonged to a man I didn't even know; a man that hadn't done me no harm. I was sorry to hear Jim say that, it was such a lowering of him. (2010, pp. 99-100)

And Huck is deeply satisfied with, even smug about, his plan to hand Jim over.

My conscience got to stirring me up hotter than ever, until at last I says to it, 'Let up on me – it ain't too late, yet – I'll paddle ashore at the first light, and tell.' I felt easy, and happy, and light as a feather, right off. All my troubles was gone. (p. 100)

Julia Markovits suggests that Huck's subsequent turnaround can be explained by his recognition of Jim's value as a fellow human being (2010, p. 208). Nomy Arpaly offers a similar explanation (2003, p. 77). This is closer to the mark, I think, but it still doesn't strike me as quite right – it sounds a bit too high-minded if intended as a take on Huck's sophomoric sensibility. He is a very ordinary boy. (He was also the original unreliable narrator of American literature, which is partly why the book is a classic.) Furthermore, Huck's fixation on Jim's utility as a *slave* (that is, as a piece of property) seems pretty hard to reconcile with recognizing his value as a fellow human being. Admittedly, we live with minor cognitive tensions on a fairly routine basis, probably by hiving them off from each other somehow. But Jim's status as a slave remains unquestioned and is (at this juncture) very much at the forefront of Huck's uncritical mind. So what happened?

It's important to notice the moment in the novel when everything changes for Huck. What happens is that Jim comes out with this:

Pooty soon I'll be a shout'n for joy, en I'll say, it's all on account o' Huck; I's a free man, en I couldn't ever ben free ef it hadn't been for Huck; Huck done it. Jim won't ever forgit you, Huck; you's de bes' fren' Jim's ever had; en you's de only fren' ole Jim's got now. (2010, p. 100)

Twain makes it quite clear that it's this moment that stops Huck dead in his tracks. In other words, Jim's identification of Huck as a *friend* proves decisive in sealing his lips. The next lines read:

I was paddling off, all in a sweat to tell on him; but when he says this, it seemed to kind of take the tuck all out of me.

Huck never changes his explicit moral beliefs – in fact, he feels roundly ashamed of himself afterward, even going so far as to give up on conventional morality entirely. Indeed, Huck acts as he believes he has *no* reason to act, even though the protective course of action hardly seems appealing: it is, after all, very risky. But the matter is nevertheless settled for him at that moment. (“I see I was weakening; so I just give up trying...”; p. 101) Somehow, Huck stumbles into a kind of decency that is never understood by him as such. (“I knowed very well I had done wrong, and I see it warn't no use for me to try to learn to do right...”; p. 103) But it *is* a kind of decency, for all that it is inchoate.

This is all likely to remain puzzling, I suggest, unless something like the possibility I pointed to in the last section obtains. Namely, there are certain superficially descriptive concepts which in fact have normative import, such as that of a *friend*. And despite Huck's obtuseness, even he knows that *one does not turn in one's friends* (just as one *does* return another's rightful property). This, I suggest, is a constitutive norm of friendship. Moreover, he is also aware that friendship is not the sort of thing which one can simply walk away from (another constitutive norm, which renders the others binding). So, in coming to recognize Jim as a friend, Huck finds himself not only unwilling but seemingly unable to blow the whistle on Jim. Doing so would transgress against the ethos of friendship he's internalized. Thus, despite Huck's explicit, misguided moral beliefs, he gains access to a potential source of moral insight, which is tacit but decisive in determining what he does.

This recognition thus marks a crucial turnaround in Huck's moral thinking, a turnaround that he probably deserves at least partial credit for. After all, he could not have had such a moment had he not in fact become Jim's friend: a relationship that turns out to be incompatible with treating Jim as a slave, or as a piece of property. So the capacity to enter into this friendship reflects well on Huck too. He has begun to care about him, in his own limited and patchy way. It is probably not accidental that this incident follows a short time after Huck had to see Jim through a bout of painful illness, following an unfortunate run-in with a snake. Huck shows clear signs of being distressed by Jim's suffering, rebuking himself continually for having planted a snakeskin in his bedroll as a prank. He reasons, somewhat obscurely, that this must have attracted its grieving mate.

So – to circle back to some other accounts of the case – the potential for sympathy might well have been there, and this might be important. (What is a friend if not someone who may be expected to feel badly for you if things go awry in your life?) But, insofar as sympathy was in Huck's heart again as he dissembled to protect Jim, I suggest that this would have been because he had a sort of paradigm shift, in light of which his latent sympathy could now come flooding back. But it is not at all clear that he *was* feeling sympathetic as he acted; he is portrayed as basically numb, and just doing what came naturally. (This happens all the time; one doesn't tend to react to a loved one's crisis by overflowing with tender emotions. In the first instance, you *act*.) Moreover, insofar as Huck acted with a sense of Jim's humanity, this would ultimately be because friendship is one of the things that can put us in touch with another person's humanness, in terms of their individuality, frailty, and basic similarity to oneself. Huck doesn't have abstract or high-falutin' concepts such as that of the human; but friendship is a mode of presentation by which other people can make themselves known to us as individuals. And in this way one might realize that this person in particular doesn't deserve to be turned in. Whether and how one might manage to generalize the lesson is a question for another day.

So we see that the case of Huck Finn bolsters the idea that people can operate according to an ethos which shapes and sharpens their moral thinking – even when no claim to the effect of “I feel obliged to do such-and-such” or “It seems to me that I may do this-or-that” ever comes consciously to mind. We also have here an example of a kind of moral gestalt shift (“slave/friend”) changing someone’s intended course of action radically, simply in virtue of their recognition of what someone means to them. We’ll come back to consider the theoretical implications of these points in the next section.

A second example of a similar sort, also involving the concept of a friend, is Michael’s Stocker’s hospital case:

...suppose you are in a hospital, recovering from a long illness. You are very bored and restless and at loose ends when Smith comes in once again. You are now convinced more than ever that he is a fine fellow and a real friend – taking so much time to cheer you up, traveling all the way across town, and so on. You are so effusive with your praise and thanks that he protests that he always tried to do what he thinks is his duty, what he thinks will be best. You at first think he is engaging in a polite form of self-deprecation, relieving the moral burden. But the more you two speak, the more clear it becomes that he was telling the literal truth: that it is not essentially because of you that he came to see you, not because you are friends, but because he thought it his duty, perhaps as a fellow Christian or Communist or whatever, or simply because he knows of no one more in need of cheering up and no one easier to cheer up. (1976, p. 462)

In this example, Stocker draws our attention to the relatively dim view we’d take of someone who visits their friend out of the motive of duty (or even out of rather less grand and general motives, such as Communist solidarity). Stocker uses the example to suggest that there can be something *alienating* about the Kantian view of moral worth, wherein (in its simplest form) an act is morally worthy only if one’s driving thought is something like: “I must! It’s the (moral) law.” And this certainly seems right, so far as it goes (how far the Kantian must retreat is a question for another day). But the positive characterization of the good case is generally rather lacking.

Someone might take the view that the case bolsters the simple so-called Humean view that an act can be morally worthy only if one *wants* to do it. But this doesn’t seem right either. We don’t generally demand of our friends that they are positively thrilled to be by our bedsides, or that they were salivating at the thought

beforehand. (That is, the appetitive sense of desire seems uncalled for. And we already know they are there of their own accord, so it can't be the volitional sense of desire in play here.) In fact, if one is in rough shape, there might be something a bit odd about the enthusiastic assurance: "Oh, I'm happy to be here – literally; this is great!" What we would generally expect and even hope for is something quite a bit different.

What we hope, I suggest (and here it is easiest to adopt the perspective of the patient) is that our friend comes to visit us *because that is just what friends do*. ("Think nothing of it," he might say.) In other words, visiting me strikes my friend as just the natural course of action. And, insofar as he would feel bound to come anyway, even if his friendly feelings were temporarily on the fritz, I would hope that he would feel obliged to pay a visit not as a moral agent or fellow Christian or Communist, but simply as a friend. For friends are *supposed* to be especially attentive to each other in times of vulnerability and need. Not so with members of large political and religious organizations, it appears. Of course, I might hope (for several reasons) that my friend is not dragging his feet unduly. But, insofar as this is the ideal, his wholehearted willingness to be there should have its basis in an understanding of the meaning of the friendship. Given this understanding, a good friend might even be expected to have little desire to be anywhere more exciting, until he has paid his dues. So he might say, if I were to express excessive gratitude or sheepishness: "Oh come on! Where else would I be at a time like this?," or "Pshaw! You'd do the same for me. I don't mean that it's a *quid pro quo*, of course. But as your friend, you know you can count on me."⁸⁰

As a final example, we move from the requirements of friendship to the generally more exclusive bonds of marriage. Consider Bernard Williams' famous 'one thought too many' thought experiment, which is supposed

⁸⁰ My response is similar in spirit to Stocker's I think, but attempts to embroider it a bit. Stocker writes that what this example shows to be lacking in moral philosophy is "...is simply – or not so simply – the person" (1976, p. 459). I agree completely. But I do worry that Stocker's subsequent suggestion that we want the visitor to act for the *sake* of their friend does not really make progress. 'Sakes' are, after all, a metaphysical nightmare. In fact, they are Quine's recurrent example of a dubious, defective noun that 'nobody wants' in their canonical notation (1960, p. 244). And we don't want to act for the sake of people *qua* bare particular egos, as Rae Langton has pointed out to me. People as they stand to *us* may be more promising subjects, I suggest. I think that this is in keeping with the key ideas in Stocker's wonderful paper.

to make trouble for even partialist versions of consequentialism. Williams takes issue with the idea that the husband who chose to save his wife rather than a stranger from mortal danger (e.g., drowning) needs some further justification for doing so – such as a moral principle which yields the conclusion: “In situations of this kind it is at least all right (morally permissible) to save one’s wife.” He goes on to remark that:

This construction provides the agent with one thought too many: it might have been hoped by some (for instance, by his wife) that his motivating thought, fully spelled out, would be the thought that it was his wife, not that it was his wife and that in situations of this kind it is permissible to save one’s wife. (1981, p. 18)

This seems entirely right to me. Obviously not everyone agrees wholeheartedly, but many people at least feel that there’s something to the idea.

The literature on Williams’ thought experiment is now vast. And many authors have justly pointed out that the case is not really a problem for consequentialism as such – consequentialism being intended as a theory of right action, not as a guide to deliberation. But what *about* deliberation? What does the case show us about that? I think few theorists have drawn the correct positive lesson from the case. Namely, concepts like ‘husband’ and ‘wife’ go beyond their legal and conventional foundations. In particular, it is partly constitutive of a marriage (in the relevant sense of an intimate partnership) that *one does not let one’s wife drown* over a ‘competing’ drowning stranger. Pressing norms like these should also have been, in the non-philosophical sense of the phrase, internalized. Having to stop and think about it would not just be bad for one’s spouse (though it would be that too), it would generally reflect poorly on the marriage itself. And the husband who didn’t dive in to save his drowning wife more-or-less automatically would reveal himself not so much to be a subpar husband (though he might well be that too), but rather to be completely out of touch – out of touch with the nature of marriage. This is why a decent excuse might be: “I’m sorry, I was temporarily beside myself. I just shut down.” But in hesitating purposefully (as in “Now let me just have a think”), he would ordinarily show himself to be insensible to what is centrally involved with one of his allegedly biggest

commitments. (Or, it might show that he has in his heart given up on the marriage, or no longer feels any sense of deep connection with his wife, although they are still ‘technically’ living together.) For, in a marriage, one’s partner’s well-being is supposed to be a priority, especially when they are in dire straits. Should it be? The question really boils down to the question of whether or not there should be marriage as we know it. So it is a more radical question than we might initially have thought – and possibly a less difficult one.

One of the things we’ve seen in this section is that, in some of the cases which have recently provoked the most doubt about the adequacy of this or that first-order moral theory (competing theories, as it happens), we are partly bumping up against an inadequacy in our moral psychology. As a result, we lack an explanation for cases where someone acts in a way which is reflective of the ethos underpinning their social relationships – notwithstanding, absent, or even in spite of their explicit moral beliefs. And, as we have just seen, the idea of having an ethos is custom-made to address this explanatory inadequacy. It shows how an agent might behave in instinctive ways, by navigating the social world by means of concepts like friendship and marriage, which implicitly contain rich and detailed codes of conduct. This is why just one explicit thought (“he’s my friend”/“she’s my wife”) is plausibly moral motivation enough. It covers a multitude, and it is an implicitly moral thought.

So people may play by the rules, so to speak, when the game they’re playing dawns on them. But should they? Are these rules in order? This is a further claim, obviously, and we will have to make do with taking it case by case until we have a grip on what makes a practice legitimate (this issue awaits section 5). But in these cases, going the extra mile is not a huge stretch for many of us, since intuitive moral admiration for these actions is pretty widely shared. That is, insofar as we want not only to explain but to *vindicate* the actions of Huck, Stocker’s Grade A friend, and Williams’ ideal husband, there is pressure to say that the ethos of the respective practices has genuine normative force. For if we think that these three amigos not only did the right thing, but that their motives were good (let alone ideal), then it would be puzzling if their motivation had no

foothold in moral reality. (Feel free to substitute your metaphysical surrogate of choice vis-à-vis accuracy conditions.) In other words, if you agree that they seem to have done the right thing for the right reasons, and you buy my explanation of their action, then you'd be hard-pressed to deny that their motivating reason is a normative reason too (to do a little backwards translation from our universal language). To be clear, this doesn't mean that this justification has any special status in the best overall normative theory. Presumably we still want reasons *why* the justification counts as such. And there might be a perfectly mainstream explanation as to its having this normative status. I'll turn to some of these issues in section 5 too.

In this section, I've helped myself to certain intuitions that are fairly widely shared but widely agreed to be theoretically vexing. The intended upshot is that, as moral philosophers, we have been bumping up against inadequacies not only in our moral psychology but also in our metaethics – i.e., in our conception of the structure of practical justification. For the simplest explanation as to why these agents do the right thing for the right reason is that the spirit that moves them – their ethos as a friend or a husband – is genuinely good. And the simplest explanation is likely to be the best, all things being equal.

4: DOING WHAT YOU HAVE TO

In the last section, we saw that there are reasons to do with honoring our intuitions about the moral importance of loyalty to buy into the idea that an ethos can be a source of practical normativity. And, as we are about to see, there are several potential theoretical payoffs in addition to these first-order fringe benefits.

The first point to be made here is that, insofar as the motivation for positing case-specific ethical intuitions is to explain how it is that agents can get some *ex nihilo* purchase on the moral facts in particular situations, or initially become responsive to practical reasons, this picture may be able to provide a very welcome alternative. A comprehensive alternative? Obviously this would be a bold claim, one I would like to make plausible eventually, but I already have my work cut out for me here. However, in section 6, I will give a brief

indication as to why one might go all the way here, holding it to be social practices all the way down, in terms of where our basic moral responsibilities come from. From there, the general claim might be within reach. For people moved to act without really knowing why often say something along the lines of: “It just seemed like the right thing to do. He’s my friend, I had to help him” or (shrugging) “It’s just what you do in these situations with a friend.” Instead of focusing on the first part of the response (as per ethical intuitionism), I am effectively moving to explain what is going on here in terms of the second part. So, if the explanation continued to work, then positing case-specific ethical intuitions would begin to look increasingly like an exercise in unnecessary reification. We could substitute the idea of having an ethos, together with ordinary philosophical intuitions about moral principles and the like.⁸¹

For, rather than saying that Huck Finn had an explicit ethical insight of the form “I am not required to turn Jim in, I should protect him!” I suggested that his thought was more likely an implicit ethical insight along the lines of “Yikes, what was I thinking? He’s my friend, after all.” This thought has the potential to lead directly to action, in virtue of the agent’s awareness that friends *do* protect each other, and they *don’t* turn each other in. This appears to be roughly what in fact happened. Huck probably *couldn’t* have articulated his moral basis for acting, even though he did have one. But, for a less muddled agent, the thought might have been mentally unpacked along the lines of: “As a friend, I have to protect him, not turn him in. That’s just what a friend does.” Thanks to the ‘have to,’ this is an explicitly normative thought now, although of a special form (to be considered next). The point is just that we need not say that Huck performed the moral equivalent of scrying (i.e., seeing around corners) in order to explain what happened, or to explain how someone more mature or sophisticated may have carried on thinking from there on in. Nothing ‘queer’ is going on here.

⁸¹ There would also be the issue of how these principles would be applicable in practice. For example, if we decide that friends may help each other out, then still the task remains of actually identifying our friends. On my view, this recognition task is going to be decidedly non-trivial. But we are natural born classifiers, and some of our classificatory schemes are evidently complex. The subsequent development of this project will focus on these issues.

This is all to the good, I suggest. For the idea of a bare moral intuition is not only unsatisfying, it seems awfully close to marching on the spot. (If the question is ‘How did he do it?’, then the reply ‘By doing it!’ is not so much a lame answer, as Williams once suggested, but rather no answer at all.) We at least want to hear more about what it is that puts Huck in touch with the moral facts in this case. And in lieu of a causal or analytic relation for him to glom (as in vision or mathematical intuition), I effectively suggested a social one. That is, Huck’s moral insight is in my view due to its dawning on him (following Jim’s inadvertent prompt) that he stands in a particular moral-cum-social relationship which mandates a certain course of action, all things being equal. And things are at least *more* equal at this juncture, because the ‘property’ paradigm has been displaced by the ‘friendship’ paradigm he’s now operating under. ‘Displace’ being the operative word – for one of the other things about friendship is that one can’t make friends with a mere object. From which it follows that, insofar as Huck is tacitly aware of this, he can no longer look at Jim *as* a slave or as an ordinary piece of property in the uncomplicated way he did before. Even if he doesn’t start to look at Jim in a whole new light straightaway, maybe the wheels will start to turn and things will ‘click’ for him eventually.

Another possible advantage of the account I’ve sketched is that it might enable us to make better sense of so-called partial reasons (although, in my view, this is actually a bit of a misnomer). There is a burgeoning literature on the question of whether and why there is a special reason to help one’s friends more than others, or to save one’s wife over a ‘competing’ drowning stranger (to now trade examples). The question then becomes: why I don’t have special reasons to help my fellow Jews, or other ex-pat Australians, or people who have brown eyes? Actually, that might not be quite as bad as it sounds: the problem might be more with refusing to help or persecuting the Aryans, Americans, and people with blue eyes. Still, being hidebound *is* parochial and creepy. What would it say about me that I was preoccupied by these sorts of superficial characteristics?

But perhaps the operative question is difficult to answer because it is not actually the right question. For if the picture I've sketched is along the right lines, looking for an individual rationale for saving one's wife is probably a mistake. The point is not just that you're in this together (for life!), and so the rationale is part of a collective pact, although I think that's part of it too. But even more importantly, what would justify this action, if anything, is likely to be the moral-cum-social fact of your intimate union. On reflection, this is just as we would have expected. Much of what is involved in a marriage is ineligible for independent, piecemeal or prior justification. You become fond of many of your partner's little quirks (and blemishes) although you could learn to live with them changing. This arbitrariness is sort of the point: you don't go out and *find* your other half, contrary to popular opinion. To some extent they become one, i.e., they *become* your special friend. (It is not a random process, and it doesn't happen overnight. What happens is essentially that you grow on one another.⁸²) Thus, any possible justification for treating one's spouse in a 'special' way seems likely to derive from the moral-cum-social fact of your special relationship, in all of its particularity and peculiarity. There are a few fundamentals though. Like not letting one's spouse die horribly if one has any real choice in the matter (again, all things being equal – and here it is stipulated that they are). Here are two perfect equals thrashing in the water; they might be identical twins or even clones of one another. But, as far as I'm concerned, I'd still go for the original. The point is not that some people are more equal than others, which is a very suspicious idea. It is that they're not all the same to *me*. And I would indeed hope that I could count on my husband to dive in and save me too, pretty much automatically. Even if that meant us both living with a lifetime of moral regret vis-à-vis the perfect stranger who never stood a chance, owing to the happenstance of love.

⁸² And marriages generally *are* pretty insular. But this is not parochialism; it is then called intimacy. For, the ethos of contemporary marriage encourages treating your spouse as your one and only, for better or for worse. Is this a good or workable ethos? Should we try more iconoclastic and unconventional marital arrangements? (See Raz, 1989.) But this is a different kind of question; marriage is what it is for the most part.

So the view that an ethos can be a source of practical normativity seems primed to make good sense of so-called partial reasons. Does it make too much sense of them though? That is, does it make impartiality a lost cause? I don't see why it should, given that there appear to be legitimate and inspiring social practices at every level of generality. That is, as well as being partners and friends, our lives are rich with social relations and communities governed by various *ethes*. There is an ethos associated with being a parent, a child, a neighbor, an employee, a student, a teacher, a citizen, and perhaps even a member of the human community at large. So the problem is likely to be too many practices, and bad ones, rather than too few. But what happens when two social worlds collide? I won't try to say much about this issue here, partly because it is a problem for most everyone who's not an act utilitarian (or has some similarly monistic, maximizing view). So there are rich resources in the existing literature for dealing with the problem of ostensibly having to do two different things at once, or be in two different places at the same time. These will likely be adaptable to situations in which one is caught between conflicting social obligations, due to one's many hats. And there is also a sizeable literature on moral conflicts that arise directly from one's role. We might be pessimistic about the chances of theorizing away agonizing decisions, where one simply cannot make good on all of one's standing commitments. But that strikes me as just the way life often is, unfortunate as that may be.⁸³

And now for the final point I want to make in this section. One thing that worries me about the reasons-based conception of practical normativity is how little room it makes for a notion of practical necessity worthy of the name. 'Must' and 'have to' are held to be simply more emphatic forms of 'ought to' and 'have a reason,' with no essential difference in the meaning of any of these terms. (There *can* be no conceptual differences, there being no room for it.) But this doesn't seem satisfactory to me. First of all, we often reach for 'have to' in situations in which we are not being especially emphatic, since we don't consider the

⁸³ The main question then becomes how to balance these various commitments. And I agree with Parfit (and many others) that there may be no uniquely reasonable, just or right way to do this, although there will be plenty of unreasonable, unjust, and wrong ways. Which leaves us with the difficult but familiar general question: what are the ethical constraints on balancing and weighing obligations? But leave it for another day, and another theorist: there are lots of resources available in the impressive literature on moral conflicts.

obligation to be especially pressing. It is rather that it is *binding* – I have to go to visit a distant acquaintance in hospital because I said I would, not because she would be particularly upset if I flaked, and not because this is held to be some especially weighty commitment. (The idea is that I wasn't independently obliged to visit in the first place, or even to agree to sign up to visit in response to her mass email soliciting some bedside reading; but now that I have agreed, well, I've made my bed.) In this case, not going when I said I'd go would be relatively inconsequential and a minor offense in the scheme of things. But, *ceteris paribus*, it would still be wrong – after all, I agreed to be there. I might say to myself if I was dragging my feet: “Still, I have to go.” On the other hand, I might not drag my feet at all. It's not that I'm particularly looking forward to it. Still, I recognize that this is just what I have to do today. And we do what we have to do.⁸⁴

This brings me to another problem with the idea of 'have to' as an emphatic normative expression. Like many people, I often find myself executing such commitments quite naturally and cheerfully, more-or-less as a matter of course (even though I can assure you that I'm certainly no saint). Phenomenologically speaking, practical necessity is no barrier to acting with a sense of total freedom. (Cf. Williams' example of practical impossibility: “I promised my wife. I *cannot* let her down.”) Why so? Plausibly because I am aware of my commitment and, as we've seen, a commitment is not the sort of thing which one can simply get out of, as a conceptual matter. And a will implies a way. Why can't I get out of it? Recall the discussion of games in section 1. One can of course defect in the physical sense (practical necessitation is not supposed to be a weird form of spiritual or physical coercion). But in the moral sense, there is no opt-out clause, all things being equal. In other words, dodging a commitment is generally not permitted. You can run, but you can't hide. Or so I have suggested.

⁸⁴ In light of this, we should note the following felicitous/ infelicitous phrase pairs:

- “I ought to – but I won't.” [sheepishly/guiltily]
- * “I have to – but I won't.”
- * “I ought to – but it's impossible.”
- “I have to – but it's impossible.” [regretfully/sorrowfully]

So I believe that an account of practical necessity worthy of the name might be within reach if we endorse the idea of having an ethos. The basic idea is that I sometimes feel bound to do something insofar as I'd otherwise be in violation of some commitment that I'm aware of having made. And there's no getting out of it now, whether I like it or not. This might help explain why our motive for certain actions is seldom naturally expressed by saying: "I felt obliged to do such-and-such, owing to my being in a committed relationship/having made an agreement to that effect." The idea of having an ethos suggests that the more natural way of putting it will also often be the more accurate, e.g.: "As your friend, I have to do such-and-such" or "I have to go – I agreed to be there."

The intended moral of this section is that the idea of having an ethos may better enable us to do justice to the virtuous peasant, who Iris Murdoch rightly pointed out that we professional sophists (i.e., moral philosophers) are liable to forget about or squeeze out (1970).⁸⁵ Not to mention the tongue-tied princess, the noble Cordelia of *King Lear*, who couldn't bring herself to wax lyrical on the distasteful theme of her father's glory. ("Unhappy that I am, I cannot heave/ My heart into my mouth: I love your majesty/ According to my bond; nor more nor less.")

And of course there is also the point that the idea of having an ethos may better enable us to explain what is going on with the artful dodger. He may be refusing to engage in team sports that the rest of us are embroiled in; he may be stalking off the field in an untoward way; and he may not so much as understand the game he's involved with. There are a number of possibilities, and these issues are intimately connected with the debates over both motivational and reasons internalism, as we saw in Chapter Two.

⁸⁵ For illuminating discussion on this point, see Lawrence Blum's "Moral Exemplars: Reflections on Schindler, the Trocmes, and Others" (reprinted in his 1994).

5: WHAT ABOUT THE INDIVIDUAL?

Now for the real challenge, which I've conveniently left for last. But now I'm a better position to explain what the problem seems to be, exactly – and hopefully to set your minds at rest before too long. An example will serve as a quick way into the biggest potential obstacle for my view that some social practices, like marriage and friendship, can be a source of practical reasons in themselves. The following discussion is very much inspired by Haslanger's insights (2003) about the range of social constructionist projects, and the notion of a 'debunking' project. I draw directly on her work on the metaphysics of gender too (see, e.g., 2000). But because I'm adapting these ideas for my own nefarious purposes, as well as over-simplifying and skating over a great deal, I obviously deserve the blame if things turn out badly although not the credit if they turn out well.

Imagine that I travel back in time to the 1950s, whereupon someone says to me: "Look, you're a woman. And as a woman, you must make more of an effort. Women bake cookies and wear lipstick and attend to the desires of the men in the household. It's just what you do. And there's no getting out of it. You're a woman for life! So better put on that apron."

Note that the speaker might admit that there's nothing natural about this, exactly. He (or she) might admit that this is just the ways things are. Still, they're the way things are. So better put on that apron.

If there turned out on my view to be nothing wrong with this statement insofar as it was an accurate reflection of the zeitgeist, then the idea of having an ethos would be effectively made ridiculous – not to mention strikingly inconvenient for someone such as myself. And this might appear to be a real danger on account of the apparent inescapability of my womanhood, together with the social norms – the ethos – that surrounded womanhood in living memory. Nevertheless, this statement was hardly all right back then.

Though it might have been an understandable sentiment, one might also have hoped to change it. But what was supposed to be wrong with it? Why did anything have to change?

Happily we are now in a position to see that there are many things which might have been wrong with this statement. This is easiest to see if we consider the many things that I might have said to effectively combat it – feminist resistance has many faces. This will help us to get a sense of the resources at our disposal vis-à-vis the general question as to why some practices seem clearly illegitimate, even if only with the benefit of hindsight. Here are some crude but possibly helpful caricatures of the main possibilities. They are ranked roughly in order of ascending conceptual (and formal) complexity:

The rebutter: “Rubbish. That isn’t what women do at all. Look around you – women are doing all sorts of more interesting things. *You think things are a certain way; but you’re wrong.*”⁸⁶

The rebuffer: “You’re right about the prevailing social norms. And maybe they’re perfectly fine. But I don’t care about being a good woman or a proper lady or whatever. Personally, I just don’t buy into it. I don’t even see myself that way. I have bigger fish to fry. So go bake your cookies yourself. *You say that things should be a certain way; maybe so, but that’s got nothing to do with me.*”

The reformer: “You are right about the prevailing social norms, but this goes into the annals of unreasonable requests. These norms are defective and unjust. They make women’s lives less joyful, less rich, and full of pointless frustration and boredom. In other words, our interests are greatly harmed by them, and those who stand to benefit – i.e., men – are getting an unfair boost from an already privileged position. They would not suffer for the loss of lipstick, cookies, and other such ministrations; they would just be disappointed or mildly inconvenienced. So we must try to change what it means to be a woman. Call me a pioneer. *Things shouldn’t be this way, and I see no reason why they should have to be. So if I have anything to do with it, they won’t be for much longer.*”⁸⁷

The rejecter: “You are right about what women are often made to be, but being a woman in your sense is *not* inescapable. You say that I am not allowed to simply walk away from these social practices. But

⁸⁶ She might continue: “Admittedly, this may have dropped off quite a bit since the war, when women did just about everything that men used to, and were far too busy to bake cookies. But the decline in women’s participation in the labor force is precisely because of these myths about ‘women’s work.’ And very convenient myths they are too, given the current economic downturn, and the subsequent anxiety on the part of men regarding unemployment.”

⁸⁷ There is also the *ad hominem*: “What’s your stake in this, exactly? Are you feeling hungry, perhaps? Is it just that you want some cookies?” Another variant might declare these requests to be incoherent and/or practically impossible. E.g., you can’t expect someone to be an ‘angel in the kitchen and a whore in the bedroom.’ (Compare Mill’s notion that men effectively demand that their wives be willing slaves.) But this response doesn’t look promising in this instance, and I’m running out of r-words. There is also the feminist *reclaimer* (aha!), who tries to adapt the meaning of a term to better fit with existing practices or people’s current experiences. But again, this response doesn’t seem naturally applicable here.

just watch me. May others follow, and let Allah be our judge. For the ethos you espouse is rotten, and it has no power to bind us. *Things shouldn't have to be this way; so it can't be that they have to be. These norms are bad and thus illicit. They are, in a word, bullshit.*"

The resigner: "In a sense, you're right. I am stuck with being a woman, and I am fated to be a bad one it seems. It's not that I don't care – how could I not? Caring about being a good woman is part of the ethos of womanhood itself. It was drummed into me since I was a little girl. I find myself forced to choose between being less of a woman, and being less of a person, or having less of a life. Ultimately, I choose door number one. I am not sure that things will ever change, and I lack my sisters' optimism that we can ring the changes with any ease or simply walk away. But what a pity! What a disgrace! *Things shouldn't have to be this way; so, even if we're stuck with it now, this shouldn't have been the case. For that matter, even if it was always bound to turn out badly, it is still a crying shame. It shouldn't have had to have been this way either. Why must people be so determined to exploit one another's interests?*"

Obviously, these positions are all at odds with one another (perhaps less obviously, real-life feminist positions are generally much subtler affairs which combine elements from more than one of these positions, if only to deal with the impressive variety of sexist and misogynist ideology). Still, the differences here might be expected to lead to inner conflicts, not to mention family squabbles. There is no need to try to weigh in on these issues here; let a thousand flowers bloom. For, assuming even vaguely egalitarian instincts, we can now be quite confident that one of these many apparently coherent lines of reply must be along the right lines in this instance. And this should also give some indication as to how the general question – "Why is this practice bad?" -- invites different possible responses in different particular cases.

Still, in light of the different positions here, the question naturally arises: does the fact that a practice is bad for its participants discharge them of any obligation to keep participating in it (or, from the outside, to support it)? If so, how bad does it have to be? The resigned response in particular raises the issue of whether we sometimes have obligations that it would be better if we didn't have. (As in: "You shouldn't have to; but you do" and "You shouldn't have had to; but now you've agreed...") Even if a practice is sub-ideal from the perspective of human flourishing or justice, this does not rule out the possibility that it may still be a source of obligations, even though it wouldn't have been, ideally (see Chapter Two). Perhaps opting out now will only make matters worse for everyone involved. But these are ultimately general issues: compare a case in

which I made a promise under sub-ideal conditions; e.g., under moderate peer pressure, although nothing that amounts to coercion. Or a case in which I stupidly made a promise, or promised to do something stupid.

In cases of somewhat bad but nevertheless binding social practices, social reform or family therapy will likely be all the more crucial. Otherwise people will be genuinely bound to do things that it would be better (for them and/or others) if they didn't have to. This is very much analogous to cases where legal reform is called for. (E.g.: "There ought to be a law against it," "It's a bad law. Nevertheless, it's the law.") Social reform would be crucial in any case though. For people may also be expected to find it pretty difficult to see why they are *not* required to do something when everyone is insisting that this is just how the game is played. Not to mention the fact that they may be punished or at least shunned if they do figure it out. And we will all miss out on getting to play a better version of the game.⁸⁸ So we find ourselves in some difficult but reassuringly familiar territory here. Philosophers of law as well as feminist philosophers have developed many helpful ways of thinking about these issues.

What do I mean that a practice is bad for its participants? As was hinted at in the little tête-à-tête above, the likelihood of a social practice causing pain, suffering, and frustration to its participants, as well as its limiting their chances of leading a rich and fulfilling life, will surely be at the center of whatever account we end up giving. There is also the issue of how it impacts on people outside – or excluded from – the practice. After all, we were hoping for people at large to get to live as well as they can. So we need to have something to say here about individual human interests.

But if individual human interests are at the heart of all this, then isn't this where the normative requirements ultimately come from? Isn't it really all about me in the end after all? No, I don't think so, and certainly not necessarily. We must distinguish between the grounds of an ethical requirement and the reason *why* it is a

⁸⁸ See Haslanger on 'target concepts' and 'critical analytical projects' (2000).

requirement and justifies certain actions. This can lead to universal and generalized claims which diverge in their truth-conditions. More concretely, practices that are greatly conducive to individual human flourishing may have to be allowed to take on a life of their own in order for some kinds of individual goods to ever be realizable. (As we'll see in a moment, this is plausibly partly because we are deeply social creatures, and partly because we are all in this together.) But, when practices do take on a life of their own, this can prove to be highly inconvenient to the individual. So it isn't all about me, in the sense that I can incur social obligations that I can't get out of even when I'd admittedly be better off that way. The point of the practice *is* ultimately partly about me; but it is also about you. So I can't just shrug off our mutual pact simply because it isn't panning out for me anymore. In other words, this would be wrong – although there may be no point in trying to enforce my performance. You may have to make do with cursing my name in my absence. You may even wish we'd never met, and have no *interest* in our getting back together now.

Why wouldn't things pan out though? Good practices can have bad outcomes if things don't go according to plan for all sorts of reasons. For example, people can fail to thrive even in good environments because they're effectively a bad seed. They may *think* they need more resources than they do in order to survive. Or they may be ruthless about hogging them regardless. They may be subject to the unfortunate effects of cross-contamination from other tainted sources. There are flowers grown in hothouses which now fail to thrive in their usual outdoor environment. And there are natural disasters to bear in mind as well. As Philippa Foot once pointed out and John Stuart Mill effectively demonstrated, moral philosophers do well to spare a thought for plants occasionally. Here we find vivid non-moral analogues of being seedy, greedy, toxic, twisted, weak, damaged, and preyed upon.⁸⁹ There is also the possibility of dying on the inside.⁹⁰

⁸⁹ However, I am not as confident that natural goodness is as analogous as Foot, Judith Jarvis Thomson, and Michael Thompson suggest in their various different ways. This is also why I'm not really inclined towards neo-Aristotelian virtue ethics *per se*. All happy people (and families) may be happy in more or less the same way. And that way does not seem to me to preclude behaving very badly towards others.

Still, we have to be careful here. If I am simply liable to get a special thrill from a certain practice owing to my social nature, then there *is* a temptation to say that it is not the practice itself that gives rise to the practical requirements in the deepest sense. Playing the game would just be my best bet (although no guarantee) of getting ahead in life. Something like: you scratch my back, I scratch yours, we hope nobody starts bleeding profusely. But, if it does turn out this way, it would merely be the cost of doing business. Moreover, if I get the better end of the bargain, then so much the better for me. On this way of looking at things, there is again nothing essentially communal or reciprocal about practical normativity; it's all about me again. Anti-social tendencies would still be on the rise. Unlike some theorists, I wouldn't regard this as a boon. I would take it to be a cheat, as well as objectionably self-serving.

But this strikes me as the wrong way of looking at things too. There are many practices such that the ethos precludes being willing to benefit at your partner's expense. And, if we wouldn't be willing to be so benefitted, then plausibly, we *couldn't* be. To clarify the reasons why, we need to revisit some issues from Chapter One. It's also worth doing this to draw attention to a neat standing possibility. Namely, that one can be a hedonist about value for the individual for all that I've said so far, as long as you are not a buck-passer about practical reasons too. (That is, you obviously can't simultaneously subscribe to the idea that what is valuable are those things which we have normative reasons to bring about or sustain, as appropriate to its nature.) This possibility might be somewhat surprising, although not to readers of Mill and Rawls. What we are going for here is essentially a version of practice consequentialism (as opposed to rule utilitarianism), with

⁹⁰ Cf. Adaptive preferences. This is essentially why I'm not a practice-based consequentialist: I don't see how force-feeding someone can ever be in her immediate interests. (See Chapter One, and also my "On What Matters Not: The Veto Power of Desire.") But good social practices must also, I believe, be attentive to people's *needs* (i.e., the sorts of things a person would have to be crazy not to care about) and thereby recognize their humanity. A good social practice must not render some people susceptible to wither and die, or be liable to make them crazy, thereby making it the case that she genuinely has no interest in what she needs anymore. Sadly, she is too far gone to even care. Nevertheless, we continue to care about her, and we will hope to somehow save her. (She matters as a human being, even if she doesn't matter to herself.) And when practices are like this, we must rage against the machine.

a side of preference hedonism. It's not quite my view, but it's a nice simple view in any case, and it's consistent with the story in this chapter.

Anyway, suppose one says that pleasure is the only thing potentially worth having, and pain and suffering are the only things that can potentially cause people harm (as I argued in Chapter One). We can still say that a pleasurable experience is only *actually* worth having insofar as I want to have it; and some painful experience is only actually worth avoiding insofar as I am unwilling to go through it. (Compare Sidgwick's idea of preference hedonism; the view I've just sketched is functionally equivalent at this level of analysis.) But why would we be willing to forgo pleasure, let alone to embrace suffering? I think there are many reasons, but one has to do with our ability to enter into social relationships which take on a certain meaning in our lives. And someone doesn't tend to want to have *une petite amie* if he really loves his wife – even though it might be fun and he's not one to feel guilty after the fact. (Hence the lament: "How could you? I thought you loved me.")

On the flipside, someone may be unreservedly willing to go through the empathetic distress and the sheer hard work of caring for an ailing loved one because that's part of what they take love to be. So you wouldn't have it any other way, in light of your loved one's sorry state. That doesn't mean you don't suffer badly in the process; but it might conceivably mean that you have no self-interested reason to do things any differently, given the situation. The point is, the fact that we internalize other people's needs can be expected to have a big impact on what we want for ourselves. And this might be expected to have a big impact on our interests, even if it's only a negative impact: i.e., hedonic values offer me certain benefits, but I can always wave them off. So there is the possibility that a relationship remains in the interests of its participants, even though it's certainly no picnic at the moment. But it is in my interests in the non-crass sense that incorporates my sense of other people and the possibility of love.

This is one reason why a hedonist might resist taking things in a utilitarian direction in view of the *interdependence* of people. And there is also the aforementioned issue that social practices can involve

commitments that are binding on us even if things *do* take a genuinely sour turn for some of the relevant parties. (Compare Rawls' remarks in "Two Concepts" on punishment; the basic idea here is pretty similar.) For there are goods in life which simply could not be gained unless one runs the genuine risk of this kind of thing happening. There are kinds of joy and fulfillment (not to mention social cooperation) which are only available to people who've made a certain commitment, even should that prove to work out badly for them down the road. What matters is that (say) marriage is one of the things that can make a human life more joyous and less lonely, all else being equal. (The same cannot generally be said of brief and casual encounters.) But when things are *not* equal, the bonds of marriage still hold – for better or worse. What seems most important here is that the badness is in some sense a stroke of bad luck. It is not that the relationship was doomed from the beginning, or would always in all likelihood have led to somebody getting hurt.

So we might countenance the strong claim that, in order to be good, a practice or relationship must be in the interests of *all* of its participants if everything goes to plan.⁹¹ In this way, a hedonist may resist taking things in a utilitarian direction in view of the *independence* of people (or what Rawls called their 'separateness'). This idea is at the heart of non-utilitarian thinking; as well as being very Rawlsian-sounding, it should also remind us of Scanlon's contractualism and of course the Kantian tradition that they both in some sense belong to. The basic intuition might be described this way: everybody matters, so although the numbers certainly matter, the proportions really don't – a high ratio of winners to losers is cold comfort if their gain is another person's mortal loss. People may be counted, but not weighed. Why so? I'll come back to this question in the next section.

There are umpteen residual issues, of course. We have yet to even consider the people outside the practice, especially those excluded from it. (We might want to add a necessary condition to the effect that a good

⁹¹ Obviously good social policies need not satisfy this strong claim; but policy-makers are in the vital but gloomy business of making the best of a bad lot (and trying to improve the current situation without making things worse in the process). Occasionally, they might also get to choose between fostering social practices which are good in different ways. But this is probably not going to be a pressing concern anytime soon.

practice is not allowed to be liable to be bad for people at large, though it needn't be expected to be positively good for them, as it plausibly might for its participants.) There is also the issue of whether or not people have communal obligations to enter *into* certain practices. (I suspect not; relationships are not supposed to be acquisition mergers; rather, two individuals are meant to come together in *forming* the relationship. On that way of looking at things, I'm a free agent until I'm not, or unless I never was.⁹²) We might also spare a thought for other sentient creatures, and for the natural world we have to move about in.

So to say there's a long road ahead is quite the understatement. But I hope to have shown in this section is that it's not an unmarked road or even unfamiliar territory. That is, I hope to have made plausible the possibility of combining two different views: (1) that social practices can themselves be sources of collective or relational practical normativity, and (2) that they give rise to practical normativity in virtue of their potential to be conducive to individual human interests when things are going well. And, as we also saw, the only thing of any ultimate good to the individual might be pleasure, and suffering the only ill, for all that has been said so far. The basic picture would then be something like this: I generally have a vested interest in having my fun unless I don't want to for whatever reason, and a personal stake in avoiding pain unless I'm genuinely willing to go through it (ditto). Good practices are then conducive to human interests, *inter alia*; bad ones conducive to people having their interests trampled or undermined (ditto). Now, as an individual, I have personal reasons to look after my own interests, and cooperative responsibilities to participate properly in the good practices with which I'm involved – unless it's no longer in everyone's best interests that I do, in which case something's gotta give. Similarly, I am obliged to resist participating in the bad practices I'm caught up in unless it's in everyone's interests that I play along. There will presumably also be humanitarian obligations to help the 'distant needy' (not to mention the battered wife across the way) unless we're not in a position to

⁹² And obviously the rules of a practice only apply to me after I have elected to start playing. I remember some of these issues coming up in the discussion following Ruth Chang's stimulating colloquium at MIT last year, which first got me thinking about the relationship between reasons internalism and the idea of an explicit commitment to another person. I am grateful for her many insights and encouragement on that occasion.

help them, although that's another story. But notice that, instead of the bipartite distinction between self-interested and moral reasons (i.e., "me" versus "them") we are effectively moving towards a tripartite distinction between individual, collective, and purely altruistic responsibilities (i.e., "me," "us," and "them"). As I said, this leaves out further elements which I think are crucial, although not everyone would agree (see *n.* 90). Nevertheless, it seems to me to be a good if rough beginning. Doubtless the devil will be in the details (e.g., how in the social world am I supposed to see my way out of trouble, exactly? And what to do if this practice is an indifferent or mixed bag?). And there is also the obvious potential for serious conflicts of interest. But this may be a problem in practice rather than with the theory, I suggested earlier.

Anyway, the intended moral of this section (following on from the opening tête-à-tête) is that bad practices make for bad laws, and as far as the question of what constitutes a bad practice goes, we have promising leads in the existing literature. Furthermore, it's not crazy (cross fingers) to hope to eventually incorporate the idea of having an ethos into a full picture of practical normativity which does justice to our individuality and our sociality too.

The issue remains as to how people embroiled in a practice can *tell* whether or not they're playing a decent game. Really that should be an 'if.' Sometimes it will be difficult or impossible for them to tell; other times, not so much. Although there may be special impediments to people caught up in the action, this is recognizable as a variant of that general epistemological problem: if I'm dreaming, how could I possibly tell? But we sometimes figure it out somehow (e.g., our bedfellow says: "You're kicking me in your sleep! Cut it out"). Arguably we may be more confident in being awake when we actually are. And it is probably more important, as well as more feasible, to collectively try to determine whether we might be living in a culture of sleepwalking – or worse – than to establish our own innocence.

6: THINGS FALL APART

In the last section we saw how practices might have to change; but sometimes they have to end. Sometimes they are not bad so much as evil. How can we deal with this fact?

We can make a start by considering why we might think that, in a good social practice, *everyone* has to be at least no worse off when everything is going well. Does this amount to an independent side constraint on good social practices? Not necessarily. There is also the standing possibility that it is practices all the way down. That is, suppose one thinks that one is subject to certain moral requirements simply in virtue of being human. That is, there is a general ethos associated with our common humanity, with 'human being' then referring to an implicitly moral-cum-social rather than a biological kind. This would go a long way towards establishing an absolute baseline that social practices must meet in order to be legitimate. A practice that will inevitably be bad for some people (inside or outside the practice) even under the best of circumstances would be bad *simpliciter* because part of what it means to be human is that one's interests will not be dismissed or overlooked. A practice that actually *destroys* people even under the best of circumstances (perhaps that was the whole point) would be simply unconscionable. Under this view, dehumanizing and soul-destroying practices such as the enslavement of Blacks and the trafficking of women would also be ruled out as sources of normativity from the outset – let alone as practices to consider readopting or supporting. We just don't do those sorts of things to other human beings, in light of their capacity not only to suffer but to comprehend their situation and lose their minds as a result. This is different from cases where the practice is excluded after we've had a brief think about its likely effects on people, as with 1950s housewifery.

I hope that these ideas might seem in some ways attractive (if decidedly vague) to those who are sympathetic to notions such as basic human dignity and moral decency. Obviously undertaking to say what it means to be human these days, and what it means to overlook someone or treat them as subhuman, is a pretty daunting task. What can we say here? A starting point might be that we don't make sacrifices of the innocent anymore.

(God is dead, or has lost his taste for fresh blood.) A ‘sacrifice’ would now just be homicide, one of the worst crimes there is, modulo a plausible *ceteris paribus* clause. Recall Anscombe’s remarks on those who would seriously consider implementing a system of justice which deliberately orders the execution of innocents periodically, to grease the wheels of utility in some elaborate way. She wrote: “I do not want to argue with him; he shows a corrupt mind.” Although I can certainly understand people’s irritation at the supercilious tone, I have more sympathy with the underlying sentiment than most. I’m certainly open to having a *discussion*, but arguing about the issue directly might well be misleading and possibly inane. To my mind, someone who thinks that an innocent might deserve to die (i.e., that this might be a perfectly just, if unfortunate, result) doesn’t really believe in innocence *per se*. (As lawyers sometimes remark: “Everybody’s guilty of something.”) They are entitled to this view of things, of course, but it will be hard to have a straightforward disagreement about the subject as things stand. We might be better off having the conversation: “What is it about the idea of innocence that you find problematic or unconvincing?” They might end up convincing me that the idea *is* more problematic than I’d initially appreciated – it is irrevocably mired in its religious roots, say, and is incoherent without that metaphysical backdrop. Or in the end, it just isn’t helpful. In which case, we might decide to go for a revisionary account or move to scrap the idea entirely.

The locution ‘believing in’ is worth pausing over for a moment. Again, we use it all the time, but we don’t work with it as much as we might in ethics. And it is interesting on a linguistic level too: compare the idea that I don’t believe in miracles to the idea that I don’t believe in organized religion. (Similarly with not believing in love at first sight, versus not believing in marriage.) Again, a normative reading is possible – indeed, virtually mandatory – as soon as extant social mores are in question. Some people don’t believe in banning abortion, eating meat, the nuclear family, the concept of ‘chastity,’ or wearing the latest fashion. Of course they believe such things *happen*, that the relevant practices exist, or that these ideas are out there; the point is that this is supposed to be a problem of some kind. Other people feel differently, or don’t believe in *not* believing in these things. These are some of the key lessons which I find in Haslanger’s (2007) and (2010)

papers, although it is admittedly my own way of putting things. Anyway, I think it might be helpful sometimes to orient discussions in ethics around these ordinary sorts of disputes. For the question of whether or not the individual actions the practice involves are permissible may sometimes be less interesting and relevant (not to mention less tractable) than the question of whether or not the practice as a whole is legitimate. (This is not a question of social policy, exactly. Or at least it is not ‘just’ a policy question.) This is another of Haslanger’s points (2003). Here’s a quick way into what I take to be the basic idea: there are some actions which aren’t properly intelligible except in a wider social context. As Rawls said, we cannot ask: “Is someone allowed to run to third base now?” out of the context of baseball. The question would make no sense. As Anscombe taught us, all action is action under a description. As Haslanger has shown, that description may have to make mention of a social practice.

Meanwhile, as human beings, here we are trying not to kill each other. We generally decline to eat each other too – people are not food. And, as Haslanger argues, to say that this stuff is fit for eating is plausibly not just a nutritional claim, to the effect that this is edible. It is also, Haslanger observes, a normative claim to the effect that this stuff is generally *appropriate* for a human being to eat (2010, p. 18). Thus, human carcasses may easily clear the bar for edibility under certain conditions, but human beings are not food, at least under ordinary circumstances. So we need not exclaim “Eating people is wrong!” in order to voice our misgivings about doing so. It is enough to say: ‘We don’t do that,’ ‘I don’t believe in cannibalism,’ or ‘Um, that’s my leg.’ Indeed, the wrongness claim sounds more than a little absurd – a joke which Malcolm Bradbury’s comic novel of that name effectively trades on. Is this an effect wrought by understatement? I’d argue not, or at least not merely. One shows oneself, in raising the issue, to have sailed right past the point.

But if good practices can give us obligations, why can’t they also take them away? Why can’t I form a Cannibal Club with other like-minded folks, where we nibble away at each other in secret (always with permission, and nobody ever *dies*). Still, this seems considerably more problematic in light of the foregoing. So

we should distinguish between practices that propose to add to our stock of reasons, versus those that propose to diminish it, or eliminate certain parts of it. While the former can lead to conflicts – a potent source of practical dilemmas – the latter are troubling at the outset, especially if there is indeed a general human ethos which the practice threatens to override, corrode, eliminate, or obscure. We have less reason to worry in advance about practices which merely propose novel norms with no independent point. Sometimes these practices even turn out to be good: emphasis on ‘sometimes.’ For initially arbitrary practices can go in more or less value-adding directions. It can be difficult to tell which is which, i.e., whether this is all going downhill, or at least getting nowhere.⁹³ Still, there is little chance of eating one’s peas with the correct dining utensil turning out to be the ethical equivalent of refraining from eating one’s pets, say. For, unlike the arbitrary rules of dining etiquette (which may not even count as an ethos anymore), the practice of having pets tends to enrich people’s lives significantly, at least in relatively prosperous households.

But what’s arbitrary about having a pet, you ask? For one thing, as Cora Diamond points out, the very fact that one *does not eat one’s pets*, even though it is not as if they constitute a special class of animals which has an independent right or additional interest in not being eaten. It is not only to some extent arbitrary which *types* of animals we treat as pets rather than food (if any). It may also be entirely arbitrary that this little piggy was chosen by the farmer to *be* the family pet. But he is part of the family now; he has a name and everything. And this might be expected to make it practically impossible for the family to chow down on dear old Napoleon now, unless there was an imminent danger of someone going hungry otherwise. (The one we called ‘Snowball’ was not so lucky.⁹⁴) As well as the possibility of not having the luxury of having pets anymore, they

⁹³ This goes for many filial and sexual taboos as well, although the social costs of change are relevant here too. And it is partly why fashion is a controversial case; I suspect though that it also has to do with its being based on an uneasy mixture of response-dependent and stipulated truths about what is ‘in’ and what is ‘out’ this season.

⁹⁴ Compare the sentence: “Cows are not food” (Haslanger, 2010). One of Diamond’s key complaints is that, in contemporary ethical writings on vegetarianism, “there is nothing in the discussion which suggests that *a cow is not something to eat* (1978, p. 468). So I take my two allies to be on the same page here.

may also change their minds. Napoleon is starting to look pretty tasty all of a sudden. If one thing leads to another, then they wouldn't be dining this evening on their pet *per se*; they would be dining on their ex-pet. What's for dinner would be Napolean *de re* but not *de dicto*. That is, they would have effectively ceased to have a pet at the precise moment when they decided to turn him into a side of bacon.

These are the key lessons to be found in Cora Diamond's "Eating Meat and Eating People." Diamond writes:

It is not 'morally wrong' to eat our pets; people who ate their pets would not have pets in the same sense of that term. (If we call an animal that we are fattening for the table a pet, we are making a crude joke of a familiar sort.) A pet is not something to eat, it is given a name, is let into our houses and may be spoken to in ways in which we do not normally speak to cows or squirrels. That is to say, it is given some part of the character of a person. (This may be more or less sentimental; it need not be sentimental at all.) Treating pets in these ways is not at all a matter of recognizing some interest which pets have in being so treated. There is not a class of beings, pets, whose nature, whose capacities, are such that we owe it to them to treat them in these ways. (1978, p. 469)

The upshot is that the hopeful ethical question: "May we please eat our pets?" rests on a mistake. We will have to rest content with queries like: "May we please eat our 'pets?'" and (more to the point, perhaps) "May we please have a pet?" The second in particular is a perfectly good question, and as difficult as it deserves to be.

Obviously good practices can come to an end, just as much as bad ones. Should we be as sad about this as we are glad about that? I don't think so, not from the outside, and certainly not in the abstract. Social practices, like other things, are liable to change and fall apart. Relationships break up; friends drift away; new ones form in their place. Practices may have no natural place in the social world anymore. Sometimes practices end because their ethos has changed over time ("I don't know what's happened to us"). Or, it has been irretrievably infringed upon ("What has happened is unforgivable; you've destroyed our relationship.") And one can find oneself simply unable to participate in a practice any longer. ("I'm sorry, but I can't do this anymore. My heart wouldn't be in it. So it's just not going to work.") Electing to withdraw from one's commitments can be a very serious business. One doesn't just walk away from a marriage, for example,

without a backwards glance. But parting company can be for the best if neither of you remains so much as capable of getting anything out of it anymore. Personal happiness is a legitimate end, even if it is not a necessary one. And walking away is sometimes positively required, as opposed to merely being allowed or recommended. But we should be open to the possibility that this is because the person has even more important commitments, which would lead to fatal conflicts of interest if things were to continue. Or she may be preparing herself to enter into other games, which are expected to be as or more important. (Channeling Cordelia, a single person might say: “Sure, I shall never marry like my sisters,/ To love my father all.”)

In this chapter, I’ve considered questions like “What do we do for each other?” and “What do we do with each other?” I’ve also touched on questions like “How do we see each other?” and “Who, exactly, are ‘we?’”, although these are really questions for another day. I’ve suggested that considering these issues might shed light on the question of what we owe to each other. But although this way of putting things is certainly civil, it is not exactly expansive. (The question “What do I owe you, exactly?” sounds more than a little tight-fisted.) Moreover, if what I’ve suggested here is along the right lines, then the question answers itself. What I owe you is what I somehow contracted to give you, in some kind of social exchange. So, at the risk of sounding stropy or naïve, I have to ask in conclusion: when it comes to doing metaethics, why can’t we try to be a bit more social?⁹⁵

⁹⁵ I want to thank some of the friends who’ve most inspired this chapter, in ways that aren’t always easy to communicate via direct quotes. One of the problems here is that I’m often riffing on their ideas in contexts ostensibly far removed (also known as stretching and possibly even distorting). But along with Sally Haslanger, I’d like to thank Lawrence Blum and Raimond Gaita for inspiring me to think about these issues. All three of these thinkers might be understood as advancing, in different ways, a conception of the interconnection between the ethical and the social which is informed by a deep sense of what life is actually like.

REFERENCES (CHAPTER ONE)

- Bergmann, Michael. (1997) "Internalism, Externalism and the No-Defeater Condition," *Synthese* 110(3): 399–417.
- Brandt, Richard B. (1979) *A Theory of the Good and the Right*. Oxford: Clarendon Press.
- Chang, Ruth. (2004) "Can Desires Provide Reasons for Action?" in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (eds.), Oxford: Oxford University Press, 56–90.
- Cohen, Stewart. (2002) "Basic Knowledge and the Problem of Easy Knowledge," *Philosophy and Phenomenological Research* 65(2): 309–329.
- Dancy, Jonathan. (2004a) *Ethics without Principles*. Oxford: Clarendon Press.
- _____. (2004b) "Enticing Reasons," in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (eds.), Oxford: Oxford University Press, 91–118.
- _____. (2000) *Practical Reality*. Oxford: Clarendon Press.
- Darwall, Stephen. (2003) "Desires, Reasons, and Causes," *Philosophy and Phenomenological Research* 67(2): 436–443.
- Finlay, Stephen. (2009) "The Obscurity of Internal Reasons," *Philosophers' Imprint* 9(7): 1–22.
- Hart, H.L.A. (1961) "The Ascription of Responsibility and Rights," in *Freedom and Responsibility: Readings in Philosophy and Law*. Herbert Morris (ed.), Stanford, CA: Stanford University Press, 143–48.
- Hill Jr., Thomas E. (1973) "Servility and Self-Respect," *Monist* 57 (January): 87–104.
- Humberstone, I. Lloyd. (1990) "Wanting, Getting, Having," *Philosophical Papers* 99(2): 99–118.
- Kant, Immanuel. (1996) *Groundwork for the Metaphysics of Morals*. Trans. Mary J. Gregor, Cambridge: Cambridge University Press.
- Kraut, Richard. (1994) "Desire and the Human Good," *Proceedings and Addresses of the American Philosophical Association* 68: 39–54.
- _____. (2007) *What is Good and Why*. Cambridge, MA: Harvard University Press.
- McDowell, John. (1998) "Virtue and Reason," reprinted in his *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press, 50–76.
- Mill, John Stuart. (1869) *The Subjection of Women*. New York: D. Appleton and Co, 1975.
- _____. (1861) *Utilitarianism*. Roger Crisp (ed.), Oxford: Oxford University Press, 1998.
- Moore, G.E. (1903) *Principia Ethica*. Amherst, New York: Prometheus Books, 1988.
- Nozick, Robert. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
- Nussbaum, Martha. (2001) Symposium on Amartya Sen's Philosophy: 5 "Adaptive Preferences and Women's Options," *Economics and Philosophy* 17: 67–88.
- Parfit, Derek. (Forthcoming) *On What Matters*. Oxford: Oxford University Press.
- _____. (1984) *Reasons and Persons*. Oxford: Clarendon Press.
- Plato. (1997) "Philebus," in *Plato: Complete Works*. John M. Cooper (ed.), Indianapolis: Hackett.
- _____. (1997) "Phaedo," in *Plato: Complete Works*. John M. Cooper (ed.), Indianapolis: Hackett.
- Quinn, Warren. (1993) "Putting Rationality in its Place," in his *Morality and Action*. Cambridge: Cambridge University Press, 228–255.

- Raz, Joseph. (1999) *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press.
- Scanlon, T.M. (2009) *Being Realistic About Reasons*. Locke Lectures, Oxford.
- _____. (1998) *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Schroeder, Mark. (2007) *Slaves of the Passions*. New York: Oxford University Press.
- Sen, Amartya. (1999) *Development as Freedom*. New York: Knopf.
- Setiya, Kieran. (Forthcoming) "Internal Reasons," in *Internal Reasons: Contemporary Readings*. Kieran Setiya and Hille Paakkunainen (eds.), MIT Press.
- Street, Sharon. (2009) "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters," *Philosophical Issues* 19(1): 273–298.
- Vogel, Jonathan. (2000) "Reliabilism Levelled," *Journal of Philosophy* 97: 602–623.
- Williams, Bernard. (1995) "Internal Reasons and the Obscurity of Blame," reprinted in his *Making Sense of Humanity*. Cambridge: Cambridge University Press, 35–45.
- _____. (1981) *Moral Luck*. Cambridge: Cambridge University Press.
- Wolf, Susan. (2010) *Meaning in Life and Why It Matters*. Princeton: Princeton University Press.

REFERENCES (CHAPTER TWO)

- Barcan Marcus, Ruth. (1966) "Iterated Deontic Modalities," *Mind* 75: 580–582.
- Carmo, Jose and Jones, Andrew. (2002) "Deontic Logic and Contrary-to-Duties," in D. M Gabbay and F. Guentner (eds.), *Handbook of Philosophical Logic* Vol. 8, Springer.
- Chisholm, Roderick M. (1963) "Contrary-to-Duty Imperatives and Deontic Logic," *Analysis* 24: 33–36.
- Forrester, James William. (1984) "Gentle Murder, or the Adverbial Samaritan," *Journal of Philosophy* 81: 193–196.
- Harman, Gilbert. (1975) "Moral Relativism Defended," *Philosophical Review* 84(1): 3–22.
- Humberstone, I. Lloyd. (1971) "Two Sorts of 'Oughts,'" *Analysis* 32: 8–11.
- Jones, Andrew and Porn, Ingmar. (1985) "Ideality, Sub-Ideality and Deontic Logic," *Synthese* 65: 275–290.
- MacKinnon, Catharine A. (2006) *Are Women Human?: And Other International Dialogues*. Cambridge, MA: Harvard University Press.
- McNamara, Paul. (2008) Deontic Logic, *Stanford Encyclopedia of Philosophy*.
- Parfit, Derek. (Forthcoming) *On What Matters*. Oxford: Oxford University Press.
- Prakken, Henry and Sergot, Marek. (1996) "Contrary-to-Duty Obligations," *Studia Logica* 57(1): 91–115.
- Stern, Robert. (2004) "Does 'Ought' Imply 'Can'? And Did Kant Think It Does?" *Utilitas* 16: 42–61.
- Strawson, P.F. (1962) "Freedom and Resentment," *Proceedings of the British Academy* 48: 1–25.
- Thomson, Judith Jarvis. (2008) *Normativity*. Chicago: Open Court.
- Williams, Bernard. (1995) "Internal Reasons and the Obscurity of Blame," in his *Making Sense of Humanity*. Cambridge: Cambridge University Press.
- _____. (1985) "Morality, the Peculiar Institution," in his *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
- _____. (1981) "Internal and External Reasons," in his *Moral Luck*. Cambridge: Cambridge University Press.

REFERENCES (CHAPTER THREE)

- Anscombe, G.E.M. (1958) "Modern Moral Philosophy," *Philosophy* 33(124): 1–16.
- Arpaly, Nomy. (2003) *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford: Oxford University Press.
- Bennett, Jonathan. (1974) "The Conscience of Huckleberry Finn," *Philosophy* 49(188): 123–134.
- Blum, Lawrence. (1994) *Moral Perception and Particularity*. Cambridge: Cambridge University Press.
- Diamond, Cora. (1978) "Eating Meat and Eating People," *Philosophy* 53(206): 465–479.
- Haslanger, Sally. (2010) "Ideology, Generics, and Common Ground," in *Feminist Metaphysics*. Charlotte Witt (ed.), Dordrecht: Springer, 179–208.
- _____. (2007) "‘But Mom, Crop-Tops Are Cute!’ Social Knowledge, Social Structure and Ideology Critique," *Philosophical Issues* 17(1) *The Metaphysics of Epistemology*: 70–91.
- _____. (2003) "Social Construction: The ‘Debunking’ Project," in *Socializing Metaphysics*. Frederick Schmitt (ed.), Lanham, MD: Rowman and Littlefield.
- _____. (2000) "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Noûs* 34(1): 31–55.
- Leslie, Sarah-Jane. (2008) "Generics: Cognition and Acquisition," *Philosophical Review*, 117(1): 1–47.
- Mackie, J.L. (1977) *Ethics: Inventing Right and Wrong*. New York: Penguin.
- Markovits, Julia. (2010) "Acting for the Right Reasons," *Philosophical Review* 119(2): 201–242.
- Murdoch, Iris. (1970) *The Sovereignty of Good*. London: Routledge & Kegan Paul.
- Parfit, Derek. (2011) *On What Matters*. Oxford: Oxford University Press.
- Quine, W.V. (1960) *Word and Object*. Cambridge: The MIT Press.
- Rawls, John. (1955) "Two Concepts of Rules," *Philosophical Review* 64(1): 3–32.
- Raz, Joseph. (1989) "Liberating Duties," *Law and Philosophy* 8(1): 3–21.
- Scanlon, T.M. (1998) *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Searle, John. (1964) "How to Derive ‘Ought’ from ‘Is,’” *Philosophical Review* 73(1): 43–58.
- Stocker, Michael. (1976) "The Schizophrenia of Modern Ethical Theories," *Journal of Philosophy*, 73(14): 453–466.
- Twain, Mark. (2010) *The Adventures of Huckleberry Finn*. New York: Vintage Classics.
- Williams, Bernard. (1985) *Ethics and the Limits of Philosophy*. London: Fontana.
- _____. (1981) *Moral Luck*. Cambridge: Cambridge University Press.