

THE NORMATIVITY OF PERSONAL COMMITMENT

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment for the requirements for the
degree of
Doctor of Philosophy
in Philosophy

By

Lauren Fleming, M.A.

Washington, D.C.
July 28, 2009

Copyright 2009 by Lauren Fleming
All Rights Reserved

THE NORMATIVITY OF PERSONAL COMMITMENT

Lauren Fleming, M.A.

Thesis Advisor: Margaret Olivia Little, Ph.D.

ABSTRACT

Many of us share the experience of having privately bound ourselves to an end or goal. Through an act of the will, we take on a previously optional end in a way that now normatively constrains our future choices. I call this sort of act a “personal commitment.”

I argue that three attempts to assimilate personal commitments to more familiar normative phenomena each fail to recapture a distinctive characteristic of that initial picture. First, personal commitments are more robustly normative than bare intention. There are substantive constraints on when we get to revoke them, constraints that go beyond considerations of efficient agency. Second, though integrity and personal commitments can both change our reasons for action, integrity-based reasons are generally holistic in their demands. Commitment-based reasons, by contrast, seem to apply each time one’s personal commitment is at issue. Finally, we might be tempted to understand personal commitments as inwardly-directed promises, since both promising and commitment seem to generate non-holistic reasons through the exercise of a normative power. But even if personal commitments are genuinely normative, they do not seem to be strongly moral in character, like promises. A personal commitment to run a marathon may really obligate one to undertake the appropriate training, but it stretches a common sense of the moral to say that it would be unethical to fail to do so. For those

of us who found the initial idea of a personal commitment compelling, accepting any one of these analyses would require a significant revision in our pre-theoretical experience of personal commitment.

I conclude by arguing that we have good reason to resist that sort of revision. Recognizing a *sui generis* normative power of personal commitment adds to a more nuanced picture of normativity as such, and the means by which agents can bind themselves. It also adds to our understanding of the ways in which agents can be self-defining.

TABLE OF CONTENTS

Chapter I: The Normativity of Personal Commitment	1
Personal Commitments: A Proposal	9
The Puzzle.....	16
Chapter II: Personal Commitments as Strong Intentions	22
Humean Models of Intention	24
Sophistication and Pre-commitment.....	34
Resolution	39
A Planning Theory of Intention	43
Chapter III: Personal Commitments as a Matter of Integrity	61
Integrity as Identity	64
Integrity as Fidelity.....	71
Narrative Integrity.....	79
Chapter IV: Personal Commitments as Promises to Oneself	88
First-Order Normative Contours.....	92
Grounding Theories of Promising	103
A Kantian Theory of Promising.....	103
Raised Expectation Theories.....	112
Transactional Theories of Promising	116
Chapter V: Towards a Theory of Personal Commitment	121
A Positive Picture of Personal Commitment	122
The “Normative Footprint” of Personal Commitment.....	123

Personal Commitment as Building on Value.....	130
Personal Commitments and the Virtue of Integrity	142
Personal Commitments and Our Picture of Agency	150
Building a Self-Conception.....	152
Works Cited	158

The Normativity of Personal Commitment

We often talk of being “committed” to various ends. I am committed to teach in the spring. That woman is committed to her husband. Libertarians are committed to small government. But just what we mean by ‘commitment’ can vary.¹ Occasionally, we use the word ‘commitment’ in ways that signify the giving *over* of responsibility, such as when a person is consigned to the care of some institution for the mentally ill. More commonly, however, to say that one is committed involves *taking on* responsibility, as well as *holding oneself* responsible. Here, a person may commit herself to some end, adopting a settled stance toward some value or goal, a stance that involves furthering that value or pursuing that goal, and giving it some significance and weight in her life.

‘Commitment,’ in the sense of taking on responsibility, can itself be ambiguous. Let’s start by distinguishing among three ways in which I might mean that I am committed to some end. Sometimes when we say that someone is “committed”, we are using what we might call an implicative sense of commitment, which marks out those stances that are implied by my other commitments, by my general behavior, and so on. Thus, if I tell you that I was born in August, I’m thereby committed to the claims that I wasn’t born in July, or September, or February, and so on. And if I say that I believe in

¹ Indeed, I think a lot of confusion over the question of “commitment” stems from a growing interest in the general topic coupled with a tendency to use this term in a relatively loose way. Mark Migotti, for example, has discussed the connections between theoretical and practical senses of ‘commitment’ (2003), and even within the practical realm, the term ‘commitment’ can vary from Michael Bratman’s theory of intention (1999, 1987) to more binding practical stances like legal contracts or promises. On the other hand, any analysis of a concept of “commitment” is further complicated by the fact that different discussions of similar concepts have been proceeding using different names, e.g., Susan Wolf on meaning (*forthcoming*), Harry Frankfurt on love, caring, and importance (2006, 2004, 1999), Bernard Williams on ground projects (1973), and so on.

evolution through natural selection, then I am at least committed to denying that creationism is true.

Alternatively, we sometimes employ a motivational sense of commitment, used to describe those goals that we happen to be strongly disposed to pursue. For example, I may say of my friend David that he is committed to his students, in the sense that he puts a lot of time into preparing for class, holds extensive office hours, writes thoughtful comments on their papers, and so on. In this case, if I say that David is committed, I mean that in fact he is strongly disposed to put time and energy into his students (whether or not he has deliberately or reflectively chosen this commitment). If, let's imagine, David burns out this semester and ends up putting a great deal less effort into his teaching, our response would be that he is *no longer* committed to his students. Or if it turns out that all of David's efforts were really just a cynical ploy to win the admiration of his colleagues, I might say that David *never really committed* to his students.

In contrast to the implicative and motivational senses, 'commitment' is sometimes used to mean that an agent has committed herself to some project in a *normative* sense. To say that I have committed myself in this sense is to say that I have, in some way, *bound* myself, put myself under a normative constraint, to pursue the project in question. Now, sometimes agents simply "end up" normatively committed as an indirect result of the ways in which they causally change the world, as when I obligate myself to apologize to you by accidentally treading on your foot.² Other times, however, a normative commitment is the result of deliberately exercising a normative power. As one example,

² This sort of phenomenon gets discussed extensively in the literature on pre-commitment (e.g. Elster), but also in discussions of moral responsibility and reactive attitudes (including Strawson and Wallace).

I can deliberately and directly commit myself by making a promise to you. When I deploy a given normative power, I exercise a capacity to change the normative landscape directly, through my voluntary acts. As Hohfeld puts it, the exercise of a normative power is some voluntary act that causes normative changes, through normative means.³ And while the term ‘normative power’ is a technical one, the concept itself is familiar: for example, I now have the normative powers to make promises and to change my legal name, though each of these powers will require the cooperation of another party, (a promisee, in the first case, and the local government, in the second).⁴

Because my normative commitments change the reasons that I ought to consider in deciding how to act, our response to failure in the normative case should be different than our response to failure in the motivational case. In the motivational case, David being a worse teacher this semester is evidence against the claim that he is committed to being a better one. If, by contrast, David makes a normative commitment to be a better teacher this semester but fails to follow through, this is precisely how we will put it: he *fails*. The failure might be justified – perhaps this commitment is outweighed by some stronger reasons – but the central point is that he has gone wrong *by the lights of this commitment*.

In this dissertation, I focus on the normative sense of commitment, and in particular, on a subset of normative commitments that I call ‘personal commitment’. Often, our normative commitments are to others: we have already mentioned promises,

³ This general description comes from Carl Wellman’s classic discussion of Hohfeld.

⁴ Obviously, this is not to say that we always have the power to change our normative situation, nor that all agents have the same normative powers. At the moment, for example, I lack the normative power to perform marriages, since I don’t happen to be a member of the clergy, nor a justice of the peace.

and could add legal contracts and plans for coordinated action. We can call these commitments ‘social,’ to mark that they essentially involve an interaction between two or more parties. Other times, however, we seem to make *personal* commitments, putting ourselves under some sort of constraint without having to make a commitment *to* another person. That is to say, we sometimes seem able to bind ourselves to ends, and thereby to change our reasons for action, without necessarily engaging in a social transaction.⁵ In these cases, we might say that I have made a commitment *to myself*.⁶

Take, for example, New Year’s resolutions. Of course, not all New Year’s resolutions are personal commitments: our “resolutions” are sometimes mere intentions, plans, goals, or even fantasies. I may resolve to take up running, just like I do every year, with the full expectation that I will fail this year as I always have in the past. Sometimes, however, New Year’s resolutions seem to be more robustly normative: we feel *bound* by our resolutions. I may say to myself, “this year, I *really* resolve to start jogging.” On these occasions, we seem to be letting ourselves down, or perhaps even failing to respect ourselves, if we don’t take our resolutions seriously. When making this sort of personal commitment, I may happen to inform others of my goals, and this communication may help to causally reinforce my commitment, but the resolution remains, in some sense, “my business.” Nobody else (reasonable reliance on my performance, and independently standing duties aside) has been wronged if I fail in my resolution. If these personal

⁵ As discussed, e.g., by Ruth Chang, Allen Habib, and Connie Rosati.

⁶ Of course, to say that I have made a commitment to myself is not to say anything about its content. While we sometimes make personal commitments that are largely self-regarding, like a serious commitment to take up some hobby, we also seem able to make relatively other-regarding personal commitments, like a private commitment to volunteer at the local soup kitchen.

commitments are normative, then we have to explain this normativity without appeal to a social transaction.

To start to get a firmer grasp on personal commitments, let's look at a few concrete examples. What sorts of acts seem to fall into this category of personal commitment, and which features characterize them? First, imagine someone making a personal commitment to perform some discrete act. Suppose that Annie learns that her local community theater desperately needs \$5,000 for some urgent building repair. Though Annie doesn't currently have \$5,000 lying around, she feels confident that she could raise that money somehow. Annie also has good reason to plan to raise this money – she enjoys going to the group's productions, having the theater helps to build a sense of community in her town, and so on. But while these considerations make it reasonable for Annie to decide to raise the needed money, they don't, by themselves, make it the case that Annie *must* devote herself to raising that money.

Assuming that Annie does decide to raise the \$5,000, she might take on that decision in a few different ways. She might simply form a private *intention* to raise money for the community theater. So long as she has this intention, she ought to act on it, on pain of practical irrationality. But if it turns out that raising the money is just too much work, or if Annie realizes she'd prefer to do something else, she can simply change her plans. Intentions, by themselves, are easily revised or revoked.⁷ Alternatively, Annie might make a public commitment to the theater troupe. She might, for example, promise

⁷ This is a claim we will take up in the next chapter.

them that she will raise the \$5,000. In this sort of case, familiarly, Annie will be *pro tanto* morally obligated to do as she promised to do: she must raise \$5,000 unless doing so would interfere with some more stringent moral duty, or she is released from her obligation by the theater troupe (e.g., if they tell her not to bother, since they've found a different source of money).

There seems, however, be a third course open to Annie, one which is more normatively robust than planning, but without the moral weight of a promise. Annie might make a *personal commitment*. She might privately decide to bind herself, perhaps even obligate herself, to raise \$5,000 for her community theater group. If Annie makes this personal commitment, then it seems like she is in some sense bound to follow through on it. Imagine that several months later Annie tells you, her friend, that she had made this personal commitment and simply never acted on it. Your reaction is likely to be either that Annie never actually made a robust commitment, or else that she is blameworthy in some way. If Annie decided that she wanted out of this commitment, it seems that she should have done something to excuse herself.

Next, let's consider someone making a commitment to a more sustained course of action, like being a vegetarian. After some consideration, Jason comes to the conclusion that, while being a vegetarian isn't strictly a moral duty, it's still a worthwhile end: raising animals is relatively resource-intensive in terms of energy, land, and water; and besides potential health benefits, a meat-free diet will leave Jason more money to devote to other pursuits. Like Annie, Jason could take this end on in a few different ways. He

might merely settle on a plan to be a vegetarian: he's not committing to a vegetarian lifestyle in any robust sense, but just "trying it on for size," so to speak. More ambitiously, if a bit oddly, he might make a promise – to his vegetarian spouse, for example – to eat vegetarian.

But Jason might also make a personal commitment to becoming a vegetarian. As with Annie's commitment, this act of commitment seems to give rise to reasons for action, reasons which go beyond the antecedent considerations in favor of making a personal commitment. That is to say, Jason's personal commitment seems to change what it is he ought to do. Whereas eating a vegetarian meal on a particular occasion may have already been an attractive option, if Jason has really robustly committed himself, then we may see him as *required* to forego meat, or at least not justified in eating it. It also seems reasonable to say that certain considerations – like the flavor of the meat dishes, or the relative prices of omnivorous and vegetarian meals – shouldn't even occur to Jason as relevant to his meal choice. Rather than being outweighed, these considerations should be "off the table" to a committed vegetarian. And if Jason ends up deciding to dissolve his commitment then we would expect him to be able to provide good reasons for the change of heart, and perhaps, to feel some regret.

As a final case, an agent might make a personal commitment to a more general goal, one whose component demands can't be antecedently specified in the way that the demands of Annie or Jason's personal commitments could. Take, for example, Shannon, whose hasty temper causes some problems in her personal and professional relationships,

and which causes her unnecessary physical and emotional stress. Though there are no doubt universalizable reasons to strive for more patience, reasons which apply to all but the most naturally patient individuals, Shannon may decide to give this goal a more central place in her life. Thus, Shannon might make a personal commitment to curb her temper and to cultivate greater patience.⁸ If she has truly committed, then it seems like Shannon must actively work on ways to remain patient (e.g. therapy, meditation, self-help books, etc.), and become more aware of the aspects of situations that tend to trigger her impatience (is it crowds? tiredness on her part?). If Shannon nonetheless lashes out at a friend one day, it seems reasonable to take Shannon as failing in a different way than someone else who acted with similar impatience, but who hadn't made this personal commitment.

Our reactive attitudes in each of the preceding examples – our feelings that each person ought to do as he or she committed to do, and that rejecting this commitment takes some amount of justification – suggest that our personal commitments may genuinely change how it is we ought to act. When we say that someone has *really*, or to put it slightly less colloquially, *robustly* committed, we seem to be saying that her commitment puts her under some important sort of normative encumbrance. This is not to claim that personal commitments generate a moral obligation. We can say that our personal

⁸ Let me note here that I recognize all three of these examples – and indeed, the various other examples I'll use for the next few chapters – are examples in which an agent has committed to some antecedently worthwhile end. As I'll continue to urge, I take it to be very significant that the act of personal commitment changes these ends from something valuable to something that the agent is *required* to pursue. Nonetheless, one might worry that I've "stacked the deck" in my favor by choosing these examples to make my view of personal commitment plausible. In response, let me first say that I take these sorts of examples to be paradigmatic of our experience of personal commitment, and therefore not really an instance of "cherry-picking". More substantively, however, I will address the question of less laudable commitments in the final chapter.

commitments change our reasons for action, while recognizing that their force is somehow different, and generally weaker than that of a promise with the same content.⁹ For example, while a mere preference shift may not be enough to excuse me from a personal commitment, a quite serious and stable change in priorities may. If, while continuing to recognize the importance of becoming more patient, Shannon nevertheless decides in good faith that it is more important at this point in her life to focus on her career, then we may take her to be released from her personal commitment. By contrast, if she had made a promise to her significant other to work on becoming more patient, the fact that she comes to care more about career advancement will not be enough to get Shannon out of her promise. Again, though, recognizing that personal commitments are generally more easily abandoned than promises needn't imply that our personal commitments can be jettisoned for no good reason, or without some sort of residue.¹⁰ The phenomenology gives us at least some *prima facie* reason to think that personal commitments do have some binding force.

Personal Commitments: A Proposal

Given the examples of Annie, Jason, and Shannon, we can now make a few tentative remarks about the characteristic nature of personal commitments. One distinguishing mark of personal commitments is that they seem to give rise to new reasons for action. By saying that personal commitments give rise to new reasons, I

⁹ For a discussion of different flavors of normativity that fall between the instrumental and the moral, see Susan Wolf, Harry Frankfurt, Bernard Williams.

¹⁰ We'll return to what sorts of justification are needed to drop a personal commitment, and under what conditions normative residue will remain, throughout this dissertation.

mean to distinguish them from two other ways in which agents can change the normative landscape. In the first place, personal commitments give rise to a more robust normativity than what John Broome has called ‘normative requirements’¹¹. We might describe normative requirements as consistency constraints: given the fact that an agent has taken some stance, she must now either abandon that stance or else accept its normative consequences. Broome takes this to be the sort of normativity involved in intending: if I intend to *phi*, I must either take the steps necessary to *phi*, or else I must abandon the original intention. These two options are symmetrical. By contrast, our examples suggest that acting on one’s personal commitment and rejecting that commitment are not, generally speaking, equally good options. Rather, making a personal commitment gives one reason to follow through on one’s commitment. We’ll be exploring this idea more fully in chapter 5.

In saying that personal commitments seem to generate *new* reasons, I mean to say that they do not just specify standing reasons. In contrast to things I happen to say, or ways in which I happen to present myself – which may specify the standing demands of integrity, or consistency – the normativity of personal commitments isn’t just a matter of filling in the details of some more general extant norm. Personal commitments create new reasons for action through the *act of commitment*, not by piggybacking on some other normative demand. (We will come back to this contrast in Chapter 3, in terms of the distinction between specifying a virtue like integrity, and creating a new reason through the exercise of a normative power.) While personal commitments will turn out

¹¹ John Broome. Normative Requirements. *Ratio*, 1999, 12: 398–419.

to be intimately related to the virtue of integrity, I will argue that integrity alone does not exhaust their normativity: personal commitments also give rise to novel reasons for action.

Furthermore, commitment-based reasons seem to be structurally complex in a particular way. As I have been describing them, personal commitments clearly give rise to positive first-order reasons for action. Positive first-order reasons are familiar: they are considerations that count in favor of doing (or not doing) something. That personal commitments seem to change our first-order reasons for action is perhaps the most clear when an agent commits to one end out of a group of equally eligible but incommensurable ends. Before making a personal commitment to raise money for her community theater, for example, raising \$5,000 to the group would be an end worth pursuing for Annie, but not something she seems to be *required* to do in any sense. She also could have chosen to spend her free time volunteering for a different organization, or pursuing some hobby. Once she has committed, however, she is violating her commitment if she never gets around to raising the necessary funds. Other times, it may be the case that an agent already ought (all things considered) to pursue the end to which she commits. In the case of the fairly common resolution to exercise more regularly, for instance, it is no doubt true that most people have prudential reason to exercise more often anyway. By making a commitment, however, one adds to, or solidifies the reasons that one has for exercising, making the need to exercise a demand that constrains other deliberation.

In concert with these first-order reasons, personal commitments also seem to give rise to second-order reasons like the sort that Joseph Raz calls “exclusionary reasons”.¹² Exclusionary reasons are second-order reasons directing us not to act on considerations that would otherwise count as relevant first-order reasons.¹³ So, when Jason commits to being a vegetarian, the fact that he would really enjoy eating a hotdog is simply irrelevant to what he should decide to eat tonight. Acting on this preference is not just *outweighed* by Jason’s commitment; rather, it is *excluded*. Though Jason might think to himself, in the presence of very strong temptation, “my commitment to being a vegetarian is simply much more important than how delicious that hotdog would taste!”, this seems like a second-best way of approaching the situation. Presumably, most of us will think that the truly committed vegetarian will simply not take the potential taste of the meat to be a relevant consideration, and that, if he does, he is being weak willed relative to that commitment.

A second characteristic of personal commitments is that they seem to be accompanied by some substantive constraints on how and when they can be fully revoked. The demands of a personal commitment could, of course, be defeated on particular occasions by more weighty or stringent considerations. Even if Annie successfully makes a personal commitment to raise money for the community theater, she

¹² Raz calls this particular combination of reasons a “protected reason” (*The Authority of Law*, 18). While I won’t get into the details of Raz’s own account in this dissertation (see, though, Emran Mian for criticism), I want to borrow his general structural suggestion: personal commitments, in generating demands, give rise to complex reasons that both direct us to act on the commitment, and forbid us from acting on a range of competing ends.

¹³ Joseph Raz. *Practical Reason and Norms*. Princeton, N.J : Princeton University Press, 1990, p.39.

may need to put off the actual fundraising if she needs to go care for a sick family member. The question here, however, is when someone who has made a personal commitment can rightly count as *no longer* being committed.

As we have already seen, there are senses of ‘commitment’ on which being committed just is following through on one’s commitment, and hence, to drop a commitment is simply to stop acting on it. Another sense of ‘commitment’ is that of an intention, and as we will see in the next chapter, intentions can typically be revised or abandoned with relatively few constraints. Dissolving a personal commitment, by contrast, seems like it may require meeting more weighty exit constraints, not just failing to act on the commitment, or changing one’s mind on a whim.

If Shannon merely intends to work on becoming more patient, simply waking up one day without a desire to be more patient would be enough to justify giving up that intention. Shannon capriciously dropping a mere intention to become more patient might be flaky, or ill-considered, (or perhaps a very good idea), but simply changing her mind doesn’t violate the norms of intending as such. (I’ll return to flesh this claim out in the next chapter.) But while an intention is just a practical attitude that allows us to settle on one course of action, personal commitment is a way of conferring a new sort of significance on certain ends and goals. Once Shannon makes a personal commitment to developing patience, more seems to be needed to excuse herself fully: for example, perhaps the time and energy needed to develop patience are needed more urgently for some other important end, (e.g., caring for her family, or meeting her professional

obligations). If one fails to meet the higher standard for revoking a personal commitment, one will be cheating on the commitment, rather than freeing oneself from it.

In addition to demanding relatively significant justification, we might also regard as apt a sort of regret from the committed agent if he is to exit his personal commitment honorably.¹⁴ For example, it seems reasonable to expect that Jason be disappointed if he is forced to renounce vegetarianism, either because he made a personal commitment that turned out to be unreasonable, or because circumstances compel him to revoke his commitment to an end that he had found so valuable. He will seem blameworthy, at least to some extent, if he fails to experience this sort of regret. We will return to the question of what sorts of exit constraints can be attached to personal commitment in more detail in Chapter 5.

Finally, and related to the last point about regret, personal commitments seem to be accompanied by certain characteristic affective shifts. In the paradigm case, the content of a personal commitment should feel more important than it did before committing: a committed agent ought to feel a sense of satisfaction when she follows through on her personal commitment, some sort of guilt when she fails, and regret if she ends up having to abandon her commitment. So if Shannon makes a personal commitment to cultivating greater patience, we feel that she should recognize, but also *feel* the normative weight of this commitment. We may expect her to feel apprehensive

¹⁴ Of course, feeling regret is compatible with simultaneously experiencing other emotions – one may regret the need to give up a time-consuming and unrewarding commitment to a local charity, all the while feeling profound relief to be free from a draining obligation to a poorly organized group.

in situations that could provoke her temper, proud when she maintains her patience, and disappointed or even guilty if she gives in to frustration. Ideally, personal commitments seem typically to involve a sort of attunement, an interest in the details of one's world that may be relevant to one's commitments. Thus, for example, Shannon should begin to notice which aspects of situations trigger her temper, and to look out for those triggers ahead of time. If she really cares about her personal commitment, these details ought to become salient for her, in a way that they likely weren't before.

Now, it may sound odd to claim that whether we can successfully make personal commitments depends on contingent psychological changes. To compare, whether I'm bound by my promise to meet you for lunch is precisely not a matter of my contingent psychology: once I've made this promise, I have to meet you whether or not I care about doing so. Indeed, I am not claiming that this sort of affective shift is a necessary condition of personal commitment, (nor could it be a sufficient condition!). Having the appropriate affect, however, may be one mark of success. This sort of caring can be a sign that one truly has committed, and also that one recognizes the import of having made a personal commitment. It would certainly seem to be unusual that an agent would make a personal commitment without also *feeling* its importance. Though Shannon simply might not come to care about her personal commitment, this would strike us as odd. More strongly, we may be inclined to say that Shannon's commitment is defective in some way.

Again, though, the normative force of personal commitment doesn't seem to be dependent on that affective shift. Thus, just as happening to care (in a psychological

sense) about some end is not sufficient to constitute a personal commitment, so finding that one no longer cares can't be equivalent to fully exiting a personal commitment. A committor's affective attunement may fall off as time goes by; where she once cared deeply, she may simply find herself apathetic about whether she acts on her personal commitment. But while this would certainly be a sign that the agent should reevaluate her commitment, it does not, by itself, seem to mean that she is no longer committed.

The Puzzle

Reflecting on the examples above seemed to suggest that personal commitments really are normatively binding. But does this intuition survive critical reflection? Though personal commitments may *seem* to be part of our lives, are they genuinely normative in the way I have described, or is this normativity merely apparent?

There are at least a couple of in-principle reasons to be wary of the claim that personal commitments are binding: one, that this claim runs afoul of Wittgenstein's private language argument, and two, that it is guilty of bootstrapping. Starting with the private language argument there are, familiarly, various ways of understanding just what Wittgenstein was arguing. One way of interpreting Wittgenstein is as arguing that without independent standards of correctness, we can't count as engaging in a normative practice (like language) at all. And just as languages are inherently social, one might argue, so private systems of normativity more generally don't make sense. That is to say, if normativity essentially only makes sense in a social context, in which individuals can

hold each other accountable to public norms, one might be skeptical of the possibility of personal commitment, as a normative move that I've claimed can be made in full privacy.

Nonetheless, I think this worry is misplaced. Even if it is the case that normativity essentially involves agents being able to justify reasons to, and demand reasons from, one another, this needn't impugn the possibility of personal commitment. Rather, one could simply see personal commitments as one agent's moves within a socially-defined system. Just as we can talk to ourselves, while denying the possibility of private languages, so we can commit ourselves, without implying there can be privately-generated normativity. Though personal commitments are always (at least in principle) privately enacted, we can always (at least in principle) be judged and held responsible for those commitments by one another.

Alternatively, one might worry that personal commitments as described above are guilty of bootstrapping. To object that something is an instance of bootstrapping, in the problematic sense, is to allege that it illicitly tries to create reasons "out of nothing". Put colloquially, the objection is that we can't give ourselves new reasons for action simply because we *decide* to "commit" ourselves to having those new reasons for action. As T.M. Scanlon points out about taking on a new course of action, "To see something as one's end is to see it as something one has reason to promote. But it does not follow that to see something as one's end is to see it as something one has reason to promote *because* it is one's end."¹⁵

¹⁵ In his *Moral Dimensions*, p.93, though Scanlon makes a similar point in "Reasons: A Puzzling Duality?".

Of course, we must keep in mind that there are various ways to take on an end. It is perfectly true that there are all sorts of situations in which a mere decision can't change what reasons we have. Familiarly, it would be wrong to think that intentions, for instance, give rise to new reasons for action: the mere fact that I have made some plan does not itself change what I objectively have most reason to do. To use Bernard Williams's familiar example,¹⁶ the fact that I intend to drink this glass of gasoline under the mistaken impression that it is a glass of gin does not actually give me a *reason* to drink this glass of gasoline. Of course, forming an intention does trigger certain norms of consistency. If I intend to *phi*, the following disjunct is true: it is the case that I ought to [take the necessary steps toward *phi*-ing, or drop my intention to *phi*].¹⁷ Most will agree, however, that intentions, along with practical attitudes such as wishing, fantasizing, and taking under consideration, fall short of giving rise to new reasons for action.

To say that personal commitment involves bootstrapping would be to argue that even our most serious private resolutions cannot give rise to new reasons for action, and can at most trigger these sorts of consistency constraints. But simply to assert "bootstrapping" as an objection is too fast. After all, unless one is a skeptic about normative powers as such, there are in fact some times when we can change what we ought to do through a volitional act. For example, I can bind myself simply by making a promise to you (I now have a new reason to be at the café), and a colonel can obligate her subordinates just by issuing an order to them (they now have a reason to shut the door as

¹⁶ From his *Moral Luck*, p.102.

¹⁷ John Broome, of course, has discussed the distinction between reasons and this sort of constraint at some length.

commanded). On these occasions, it is precisely a volitional act that changes what somebody ought to do – I ought to do as I have promised, and the soldiers ought to obey their commanding officer. There are, of course, some constraints on these normative powers being executed successfully; the recipient of the promise, for example, must have what we might call “uptake”, accepting the promisor’s offer to obligate herself. Nonetheless, the promise is binding because the promisor wills that she be obligated in that way.

Thus, to say that recognizing a power of personal commitment would entail accepting bootstrapping is to jump the gun a bit. The real question is: can we settle on an intelligible notion of personal commitment, or are we unjustified in claiming that personal commitments give rise to new reasons for action? That is to say, should personal commitments be included among those normative powers that change our reasons for action, or are they really more like intentions, governed only by norms of practical consistency? And in putting together a positive view of personal commitment, we seem to face a tension. On the one hand, it often seems as though we can bind ourselves through personal commitments, in ways that need not involve any social interaction at all. The demands that Jason’s personal commitment makes on his time and resources can seem to defeat mere preferences, or even policies, like eating what he would find most tasty, or eating whatever his hosts at a dinner party put in front of him out of politeness. When someone like Jason fails to live up to his personal commitment,

he may feel guilty.¹⁸ Furthermore, given that he has committed, this guilt may seem reasonable from the perspective of a third party, or even demanded by the situation. An observer might aptly say about his feelings of guilt, “After all, he *did* make a commitment.” Yet on the other hand, despite these intuitions, the idea of binding ourselves apart from a social transaction may seem mysterious. If I have not committed *to* someone else, someone who can hold me to my word, the worry is that I have not committed at all. Unless I *owe* this commitment to someone else, it seems like I can let myself off the hook whenever I want – at most, I am under the sort of consistency constraints imposed by intention. After all, if I decide that I no longer want to be committed, who is to gainsay me? Who am I hurting? If, however, I can exit a personal commitment “for free”, then I cannot really violate it; it is not robustly normative.

In this dissertation, I address this puzzle in two ways. First, I continue to develop this initial picture of just what personal commitments look like. What are their characteristic normative contours? What functional role do they serve? How do their demands interact with other norms? I will contend that, given our characteristic experience of personal commitments, explanations that purport to assimilate personal commitment to some other phenomenon are deeply revisionist. Though ultimately some may be willing to accept this revision, we must at least recognize it as a cost. Second, I

¹⁸ It is sometimes argued that guilt is always other-regarding (a response to failing others), while shame is self-regarding (a response to failures to meet one’s own ego ideals). Gabriele Taylor, for example, argues that one only needs to imagine an audience in order to feel shame, whereas feelings of guilt require the belief that one has violated some external norm. Even if, however, we accept these analyses of shame and guilt, I think my final analysis of personal commitment will show both sorts of reactive attitudes to be appropriate in the case of personal commitment – an act of personal commitment gives rise to new norms, whose demands one may feel guilty in violating, but also centrally implicates one’s integrity, making failure a matter for shame as well.

address the way in which the picture I develop can shed light on our broader conception of agency.

In the next three chapters, I will consider three different attempts to analyze personal commitment to some more familiar phenomenon. In chapter two, I consider the possibility that personal commitments are simply a particularly psychologically strong subset of intention—an attempt that, essentially, abandons the intuition that personal commitments are robustly normative. In the third and fourth chapters, I consider attempts to retain the robustly normative nature of personal commitments, but by assimilating them to other, more familiar phenomena, whether as merely a matter of the virtue of *integrity*, or else a limiting, self-directed sort of *promise*. Ultimately, I will argue that all three of these attempts are insufficient, and that we needn't take for granted that there is something contradictory about privately binding oneself to an end or goal. The final chapter will offer a more detailed account of personal commitment, and argue for understanding personal commitment as a distinctive normative power.

Personal Commitments as Strong Intentions

In the introduction to this dissertation, I began to describe what I take to be the characteristic contours of personal commitment, both phenomenological and normative. A personal commitment seems to be a way of first-personally, and privately, putting oneself under some sort of obligation. When we make a personal commitment, we give that end a particular sort of significance in our lives, making the previously optional end something which issues demands, which we now *must* pursue. Yet even if this general description sounds familiar, in part or in whole, one might wonder whether personal commitment is really a distinct *kind* of practical attitude. Rather, perhaps personal commitment is just a sub-genre of some other sort of normative phenomenon. In this chapter, I will ask whether what I have described as personal commitment is simply what an intention looks like when it is very strong or particularly persistent.

Let me be clear about this suggestion. On any account, personal commitments will obviously entail intentions as upshots, or components. If I make a personal commitment to learn to read German, for example, this seems to entail that I intend to study grammar at some point, to learn new vocabulary, and so on. And if I abide by this personal commitment, that will involve acting on the relevant intentions. Since personal commitments involve intentions plus some additional mental states and associated norms, (though we don't know yet just what these will turn out to be), there is a clear sense in which one could say that personal commitments are, in fact, "souped-up" intentions, or "intentions-plus". Our question here, however, is whether personal commitments are

really *just* intentions. Are personal commitments different from garden-variety intentions only because everyday intentions happen to be more fleeting, or less weighty? Or is personal commitment partly defined by normative contours that outstrip those of bare intention?

Understanding personal commitment as simply one species of intention would come with a number of benefits. For those who prefer desert landscapes, it wouldn't require adding any elements to our ontology. The significant philosophical work that has already been done on intention would give us a starting point for understanding the way in which our commitments bind us, and how this normative force is grounded. More substantively, intention does involve a certain sort of "commitment" to a course of action; this fact is picked up by the many philosophers who use the word 'commitment' to describe intention.¹⁹ Having settled on an intention, I've committed at least in the sense that, absent further reflection, this is the course of action I'll pursue. Now, unlike personal commitments, which seem typically to be relatively strong and persistent, intentions can vary pretty widely in both their motivational strength and how long they persist across time. But this difference would be perfectly compatible with recognizing personal commitment as a subset of intentions, one marked out by greater degrees of motivational strength and diachronic persistence, but not giving rise to a different *kind* of norm, namely a requirement on action.

¹⁹ Including Michael Bratman and David Gauthier, among the authors cited in this chapter.

Ultimately, however, I don't think we can understand personal commitment simply as a very strong, persistent, intention. If we are to do justice to our everyday experience of personal commitment – the way in which we form and revoke our personal commitments, and how they characteristically direct our action while they are binding – then personal commitments come apart from mere intention.²⁰ Of course, there are many different models of intention, but for our purposes, I'd like to focus on a few leading attempts to understanding the way in which intentions can have diachronic force, since I take a central problem for personal commitment to be how they can bind us over time. As we consider these various theories, including Humean approaches, a few decision theoretic models, and the “planning theory” of intention, we will also try to flesh out our picture of the normative contours that define personal commitment. If personal commitment is not fully reducible to intention, then what is it that is missing – what more is involved in personal commitment?

Humean Models of Intention

The starting point for any discussion of intention is the belief-desire, or “Humean”, model. Familiarly, on this model an intention is made up of a cognitive component (a belief), a motivational component (a desire), and some sort of connection between the two.²¹ (At the very least, the belief has to be somehow about the object of

²⁰ As I already suggested in the last chapter, I take the overarching structure of this dissertation to be a conditional one – *if* our initial description of personal commitment is veridical, then what kind of analysis of its binding force can we give? As we go through the middle chapters, we will elaborate on this initial picture, but we won't really get into whether we *ought* to take this picture as accurate until the final chapter.

²¹ Though there are a variety of interpretations as to just what counts as a “Humean” model of attention, I take the general picture I sketch out here to recapture the most commonly shared aspects of a Humean

the desire). Though we are not primarily interested here in ontological questions, (i.e. whether intention can be reduced to the states of belief and desire, or has to be recognized as a psychological primitive), Humean models are also committed to certain claims about the normative demands that having an intention places on an agent. Rather than delve into each of the numerous variations that have been developed on this general picture, I will proceed with a basic sketch of the normative constraints under which intention, so understood, places us.

Because it analyzes intention into two more basic components, the Humean model also derives the norms associated with intention from those governing belief and desire. Belief entails relatively robust normative constraints, including mutual consistency, respect for evidence, and explanatory coherence; in having belief as a component, an intention is subject to the same constraints. Thus, intentions must be mutually consistent, so that I am criticizable if I intend both to *phi* and not to *phi*, or intend to do both *phi* and something I know to be mutually exclusive with *phi*-ing. Intentions must also be responsive to my other evidence about the world, (e.g., I cannot rationally intend to change the past, given my knowledge of how time and causation work), and are preferably grounded in beliefs that themselves are simple, plausible, and explanatorily powerful.

By contrast, desire alone, understood roughly as a disposition to pursue the object of one's desire, does not put us under any normative constraint. As one might put the point, we can have desires "for free." This is not to say that we can desire at will, but

theory, and also to serve as a foil, being an essentially motivational entity, to the essentially normative personal commitment.

rather than that however we arrive at a desire, having it doesn't bring normative constraints with it. The nature of desire itself does not require that one's desires be mutually consistent, morally appropriate, or even realistically achievable.²² Given that the Humean picture defines what it is to have an intention functionally – a belief-desire pair can count as an intention only if the desire component is *ceteris paribus* strong enough to move one to act (at the least, in the absence of a stronger desire) – the strength of the component desire cannot fall below a certain threshold if one is to count as having an intention.²³ To be an intention is, essentially, a matter of having the right sort of *motivational* force and content. Given this connection to motivation, Humean intentions are governed by a relatively thin sort of normativity: if an agent claims to have an intention to *phi*, yet fails to take the appropriate steps toward *phi*-ing, then one of three things is the case. Her intention to *phi* may be outweighed by some stronger intention, she is engaging in some self-defeating behavior, or else does not really have an intention to *phi* at all.

Given this picture, can personal commitments be identified with intentions? Can we understand personal commitments as a motivationally strong subset of the rough sense of intention we've been drawing? I think not: personal commitments are more than just intentions understood as strong belief-desire pairs. In particular, personal commitments seem to be characterized by a relatively robust, diachronically extended

²² This is precisely why some of our desires are called hopes, wishes, and even fantasies.

²³ Clearly belief, like desire, can come in degrees, and there is also a minimal strength of belief below which one no longer counts as having an intention. Just as I do not intend to *phi* if my desire to *phi* is only slightly warmer than indifferent, I do not intend to *phi* if I'm quite skeptical that such a thing is possible.

sort of normativity, one which outstrips the relatively thin normativity that characterizes a basic Humean picture.

Surely personal commitments, like Humean intentions, are subject to the norms that guide belief. It would be irrational, for example, to make a personal commitment that conflicted with some standing intention and yet have no plans to revise either the commitment or the intention. Similarly, we could criticize a personal commitment to building a perpetual-motion machine as irrational if the would-be committor also recognizes that friction makes such a thing impossible.

Where the norms guiding personal commitment and intention seem to come apart, however, is when we turn to the practical side of things. Again, for the Humean, intentions lead to action because ‘intention’ is functionally understood: your mental state counts as an intention only if it will (under the right circumstances, and in the absence of a stronger competing intention) lead you to act. If you “fail” to act on an intention in these circumstances, the appropriate response is not actually that you erred in some way, but rather to say that you didn’t actually have the intention in question. For example, if last night you formed the intention to go running this morning, and on waking up you find that you no longer feel like going for a run, then you’re off the hook – you no longer have the intention in question, and so can’t be said to be failing by its lights. A Humean intention has normative force only so long as the desire continues to persist.²⁴

²⁴ It may seem here that I am being unfair to Humean theorists, who of course recognize that it is often more rational to act on one’s long term interests or preferences, not just one’s strongest occurrent desire. For the moment, I am just focusing on the basic Humean framework. The next section will consider some strategies Humeans have suggested for explaining how – given their framework – it can be rational to make decisions that go beyond an agent’s strongest occurrent desire, to take into account her own past and future desires.

This, however, doesn't sound like the sense of 'personal commitment' that we've been using. If I've made a personal commitment to go running each morning, for example, then it does not seem to be the case that I can, with no remainder, simply change my mind and reject this commitment. Similarly, if Jason can simply excuse eating a hamburger for lunch by saying that he "felt like it", that feels incompatible with the sort of commitment we've been describing. At least if our pre-theoretical experience is right, there is a difference between *breaking* a commitment and *exiting* that commitment.²⁵ If I'm genuinely committed, then at the very least getting out of that commitment seems to require thoughtful reconsideration, as well as having some good reason to forego one's commitment.²⁶ And if I lack good reason and yet simply ignore the demands of my commitment, I should feel some sort of guilt or shame at failing by my personal commitment. These sorts of claims, however, are far too strong to characterize Humean intention, as such.

Of course, it is important to remember that our original suggestion was to define personal commitment as a subset of intention, one made up only of particularly strong, persistent intentions. That is to say, we were asking whether personal commitments can be identified with those intentions that happen to be fairly effective at motivating an agent, and whose motivational force tends to maintain its strength over time. Whereas

²⁵ Though, as it turns out, I think that there is a complicated relationship between extended failure or cheating, on the one hand, and exiting a personal commitment, on the other. This is a claim to which we'll return in the final chapter.

²⁶ As we've already noted being able just to change one's mind may seem to be compatible with certain senses of 'commitment'. In the right context it makes perfect sense to say, for example, "yes, I had committed to going running this morning, but then I changed my mind." And this is precisely the sort of 'commitment' that various authors employ when they talk about intentions. As I've stipulated, though, our focus is on those times when we seem to use the word 'commitment' in quite a distinct sense, one which is wrapped up with a certain sort of diachronic normativity. That is to say, we sometimes use the word 'commitment' in a way that at least suggests that a committed agent has put his future self under obligation.

Shannon's intention to become more patient certainly could be a relatively fleeting one, we could also imagine that a scenario in which Shannon's intention happens to be quite persistent, and carries a lot of weight in her decision-making. Identifying personal commitment with this subset of Humean intention seems to come closer to our initial description, on which personal commitments are rather significant in deliberation, and not just practical attitudes that one can adopt and abandon on a whim.

Notice, however, that this proposal – identifying personal commitments with those intentions that happen to be more strongly motivating across a reasonable stretch of time – means that I won't be able to tell whether I've actually made a personal commitment (as opposed to a weaker intention) until quite some time has elapsed. Though I may say that I am making a personal commitment to *phi* at some time *t*, I will not be able to tell whether this is actually a personal commitment or just a weaker intention until enough time has elapsed to see whether the “commitment” ends up having the requisite motivational force and persistence.

As we've noted many times before, the language of “commitment” is fairly vague, and I think that there definitely are some sorts of commitment whose success can only be assessed retrospectively. For example, I take this to be the case with what I called, in the introduction, a ‘motivational commitment’, which is really just a pattern of reliable action toward some end. Though David may genuinely take himself to be committed to teaching, feeling the objective importance and subjective pull of helping his students in some moment of decision, whether this feeling is something relatively deep

and persistent or just particularly strong but fleeting is something that will take time to determine.

In other kinds of commitment, however, what determines whether I've been bound is simply whether I've made the commitment – whether I've executed the appropriate act to bind myself – something which can be determined synchronically, as opposed to diachronically. In the case of promises, for example, whether the commitment comes off is something that can generally be determined at the moment of promising. Similarly, I think that our initial examples (of Annie, Jason, and Shannon) seem to show that there is a kind of private commitment whose binding force can be determined prospectively, rather than retrospectively. In these sorts of cases, one says, “I commit myself to donating to the community theater”, not “I hope this intention to be more generous turns out to be a serious commitment”, nor “I really mean to donate the money, but we'll see what happens.”²⁷

At least as many of us have experienced these personal commitments, the act of commitment is a prospective one, laying normative constraints on one's future behavior, constraints that are often put in place precisely to counteract the possibility of waning motivation. Saying that Shannon has made a personal commitment seems to go beyond

²⁷ Even in the sense of commitment that is our focus, it may be the case that *sometimes* I don't know whether I have made a personal commitment until some amount time has passed. For example, perhaps I thought I was really committing myself to fighting animal cruelty in the moment where I decide to commit myself, only to recognize later that I had simply been overwhelmed by powerful emotions while watching a public service announcement – I was being passively, if strongly affected by emotions rather than actively undertaking a commitment. In such a case, strong emotions and environmental factors may cloud my ability to determine whether I've actually, as a volitional matter, committed myself to some goal. I think it's a mistake, however, to take this sort of experience to be a paradigm of “personal commitment”. Rather, the cases in which I turn out to be wrong as to whether or not I have made a personal commitment seem intuitively to be a sort of anomaly, interesting precisely because they demonstrate a failure of the normal functioning of personal commitment.

just pointing out that she happens to believe that an improved temper is worth pursuing, and that she also happens to be strongly motivated to achieve that end. It is to say that she has decided to constrain herself, to make it the case that she is *required* to pursue the end of patience. And it is from the moment of commitment that we begin to assess whether Shannon is living up to her commitment, not just retrospectively. If Shannon makes this personal commitment and then has a spotty record of patience over the next month, our reaction is likely to be that she needs to take her commitment a little more seriously. Or, if Shannon indulges in a temper tantrum the day after making her personal commitment, we may think it's the case that Shannon wasn't really ready to take on this particular commitment. These reactions make sense because the normative force of Shannon's commitment results from her act of commitment, her willing herself to be so committed, not just because of a pattern of caring about patience and striving for it.²⁸

Now, claiming that the normative weight of personal commitment can come apart from motivational strength in this way, (such that we can determine whether someone is committed prospectively without knowing what her motivation will look like over time), is not to say that psychology isn't relevant here at all. Whether the object of my commitment feels compelling to me can be *part* of the story about how I make, revise, and even reject a personal commitment. For example, at least part of one way we might honorably exit our personal commitments would be to find that they are no longer

²⁸ I will continue to develop a story as to what distinguishes descriptive cases of commitment (in which non-performance is a sign of non-commitment) from normative cases of commitment (in which non-performance means that one is failing on a commitment that one has really made), as well as to make the case that there are private instances of normative commitments. For the moment, I'd just like to focus on the fact that, pre-theoretically, we seem to recognize both sorts of commitment among the sorts of things an agent can do.

meaningful to us. If Jason really ceases to care about being a vegetarian, this could be one step toward dissolving his commitment. If this personal commitment ceases to add some meaning to his life, or to show up as important, then it would seem to be a sort of fetishism to insist that he remain committed nonetheless. But this seems, at most, to be part of the story – whether or not Jason is legitimately released from his personal commitment will also depend on other factors. For example, psychological facts about his state of mind (is he simply distracted today, or in a cynical mood about the importance of his contribution), as well as normative facts about this change (does he have a good justification for this shift in meaning), and so on, may also be relevant. (We will return to questions of the relationship between psychology and personal commitment in the final chapter.)

However the details of our positive account work out, for now there are two main ways in which a belief-desire model of intention fails to recapture our pre-theoretical experience of personal commitment. First, casting personal commitment as a particularly strong sort of belief-desire intention seems to tie our commitments too closely to our contingent psychology, whereas personal commitments seem to be something that we can actively will into being. An agent will, no doubt, have some story as to why she is motivated to make a given personal commitment, but the binding force of the commitment itself comes from the act of commitment, not from the underlying motivation. Second, the sort of normativity that characterizes personal commitment seems to be more robust than that of a basic belief-desire model of intention. Rather than triggering rather minimal consistency constraints – one must follow through on one's

personal commitments, at the pain of no longer being committed – our personal commitments seem to give us new reasons to follow through on those commitments. The act of commitment takes the choiceworthy value in Annie’s generosity, Jason’s lifestyle choice, and Shannon’s plans for self-improvement, and makes those ends that each agent is *required* to pursue. If we are going to offer an account of personal commitment that does justice to this experience, we will need a model that incorporates this more robust sort of normativity.

Now, the vast majority of theories of intention agree that intentions do not give rise to new reasons for action, in large part because this claim seems to entail the problematic sort of bootstrapping that we mentioned in the last chapter.²⁹ Given that I’m pushing the idea that personal commitments bind us by giving rise to reasons, and indeed reasons that are in particular *requirements* on how we ought to act, one might wonder why I would bother to consider other theories of intention. Why not skip ahead and see if I can make sense of reason-generating personal commitments in other ways? While it will be no surprise that ultimately I don’t think we can give a theory of personal commitment solely in terms of intention, I nonetheless think it’s worth looking at more sophisticated theories of intention for two reasons. First, it could turn out that we find one of these models fits pretty closely with our intuitive view of personal commitment, and that ultimately we should revise our initial ideas about just how binding personal

²⁹ See, e.g., Bratman (1987, 25-6) and Broome (2001, 98). There are, however, notable exceptions. David Velleman and Christine Korsgaard, for example, argue that intentions give rise to new reasons for action, (and den Hartogh argues that Bratman implicitly has this kind of view as well). Though I won’t explicitly take on either of their specific views, I think that the other objections I raise to considering personal commitment as a particularly strong kind of intention will extend to their theories as well.

commitments are. Second, even if we don't revise our original intuitions, using intention as a foil may help us to get a clearer picture of how personal commitments bind us.

Sophistication and Pre-commitment

As it happens, providing an account that shows how certain intentions can continue to bind us even in the absence of an occurrent supporting desire has been one central goal of recent work in decision theory. After all, there are times when it seems like instrumental rationality alone (in the absence of anything like a personal commitment) requires more stability in our intentions than the stability that can be provided by what we happen contingently to want most. Weak will, for example, seems to be one paradigm case of practical irrationality. Consider a typical case of weak will: a settled preference to lower my cholesterol may lead me to form an intention to pass on rich desserts. When offered a particularly tempting piece of cheesecake, however, my preferences may temporarily shift, such that I now prefer the cheesecake much more than lowering my cholesterol. After eating the cheesecake, my preferences will revert, such that I will once again care more about lowering my cholesterol, and will regret having given in to temptation. If my intentions are just a function of my relevant beliefs and my strongest occurrent desire, or preference³⁰, then it seems like what I ought to do is eat the

³⁰ 'Preference' is a theoretical term of art, as indeed is 'desire' in the sense we have been discussing. Because the literature that deals with what counts as rational intention stability uses the term 'preference,' in this section I will shift from the language of 'desire' to that of 'preference'. Nonetheless, because both desires and preferences, in their respective senses here, are just a matter of one's brute motivational dispositions, I take this section to be responsive to the themes and concerns of the last section, and to what follows in my discussion of Bratman.

cheesecake, and then regret my decision afterward. But giving into this sort of temptation seems like a classic case of irrationality.

In order to explain why giving into this sort of temptation is irrational, we need to broaden our perspective in assessing instrumental rationality, so that the rational formation and revision of intentions isn't just a matter of one's strongest occurrent desire. One such model is 'sophistication': the sophisticated chooser tries to ensure performance of her future-directed intentions by settling only on those intentions that are both supported by her current preferences and reasonably foreseen to be supported by her preferences at the time of action.³¹ That is to say, the sophisticated agent does not form intentions, however well supported by her current preferences, if she foresees that she will fail to follow through on them. A sophisticated agent, for example, will only form an intention to jog daily if this intention can help to satisfy her present preferences, and if she foresees her future preferences allowing her to make good on the intention to run.

At first glance, sophistication may sound like it holds agents hostage to their future whims – I can't intend now to jog daily, given that I can reasonably predict that I will encounter temptation to skip a run at some point in the future. But note that agents are also capable of affecting their own future preferences, for example by rigging the future choice situation. Jonathan Elster, for example, calls one such tactic 'pre-commitment'. In pre-committing, agents affect their own future behavior by causally changing the choice scenario, most commonly by eliminating future opportunities to give into temptation, but alternatively by altering the expected outcomes so that resisting

³¹ Gauthier, 223

“temptation” will actually be more rewarding. Thus, Ulysses binds himself to the mast so that he can listen safely to the Sirens, the tippler locks the liquor cabinet to avoid bingeing, and the spendthrift puts away his money in a Christmas club account. Pre-commitment allows an agent to control his future behavior by making sure his future preferences will be in line with what he wants now.

Pre-commitment is clearly an important tool for temporally extended agents, since there are times when coordinated agency requires causally influencing our future selves. There are many times when it is important to control our future action, but when a mere decision will be insufficient to guide us: Ulysses could make the most solemn vow to his crew that he’d ignore the Sirens’ song, but unless he *causally* restrained himself, he would jump into the sea nonetheless. Pre-commitment strategies are also familiar to us from more mundane examples: I put the alarm clock across the room so that I will have to get out of bed to turn it off, I find a running buddy so that my desire not to inconvenience a friend will bolster my weaker desire to go for a run, and so on.

While precommitment strategies are important practical tools, however, this is not the way in which personal commitments operate. Rather, as we noted briefly in the introduction, if personal commitments do exist as a distinct sort of phenomenon, then they change our reasons for action in a directly normative way, not by indirectly piggybacking on causal changes. Put in more technical terms, agents make personal

commitments by deploying a normative power.³² If I can bind myself through a personal commitment, then I can bind myself more or less just by saying to myself “I commit to...”. To make this point more clear, compare the two ways of changing one’s reasons for action (causally versus through a normative power) in the interpersonal case. I can put myself under obligation to buy you a new pair of shoes if I ruin yours by accidentally spilling a cartridge of printer ink all over them. In such a case, I have indirectly made a change in my normative status through the way I have causally changed the world; I owe you a pair of shoes because my actions trigger a standing duty to make up for my harmful actions. Alternatively, I could put myself under obligation to buy you a new pair of shoes if I *promised* to do so. In this case, I have directly changed my normative status (I now owe you a pair of shoes) but not in virtue of some causal change I’ve made. Whatever one’s theory of promising (and as we’ll see in Chapter 4, there’s a fair amount of diversity in philosophical thought on promising), our promises are binding simply because we have given our word.

Precommitment seems to be like an intrapersonal analogue of piggybacking a normative change on a causal one. In both cases, an agent’s reasons for action are changed indirectly, in virtue of physical events affecting which reasons are relevant to the situation at hand. Personal commitment, by contrast, seems more like an intrapersonal analogue of the promise; in either case, the agent is able to put herself under normative

³² As a reminder, in the introduction, we defined a normative power as the performance of “some voluntary act that causes normative changes, through normative means”. Examples of normative powers include the power of clergy or judges to marry couples, the power of legal persons to enter into contracts, and the power of moral agents to make promises.

constraint just by willing that to be the case.³³ While a precommitment is effective by hooking into standing reasons, narrowing one's options (as when the tippler locks the liquor cabinet and throws away the key) or changing the expected payoff for a given action (as when I enlist a running buddy to hold me to my plan to go jogging), personal commitments with the same content operate by creating new reasons for action. I ought to follow through on a personal commitment simply because I have committed to do so.

When Annie, for example, makes a personal commitment to raise money for her community theater, she seems to bind herself in virtue of that act of commitment, whether or not she has made any causal changes in the world around her. Perhaps this statement of commitment (whether made inwardly or actually verbalized) needs to be accompanied by certain psychological, affective, or volitional states in order to bring off a successful normative change.³⁴ After all, to compare, there are similar constraints on whether a promise is successfully made – in particular, the promisor (the maker of the promise) must recognize that she is doing the sort of thing by which she can bind herself, and the promisee (the recipient) must take the promisor to be attempting to bind herself. But though there may be these sorts of conditions on the success of the normative act, neither the personal commitment nor the promise seem to have their normative upshot through the sort of causal piggybacking we see in precommitment.

³³ Indeed, this is a parallel we will consider in much more detail in Chapter 4.

³⁴ This is a question to which we will return throughout the dissertation, particularly in the final chapter.

Resolution

I've just argued that sophistication is unlike the picture of personal commitment with which we've been working in two ways: first, sophistication is a strategy for insuring that one's intentions are more *stable*, whereas personal commitments seem to go beyond this in actually giving rise to new reasons, and second, sophistication operates by causally changing the costs and benefits of performance, not directly through an act of the will like commitment. There are other ways, however, of modeling intentions that are more stable and normatively binding. The model of resoluteness put forth by theorists like Edward McClennen and David Gauthier, for example, attempts to recapture the rational stability of intentions by broadening the stretch of time over which we assess rationality. "Whereas the ex ante sophisticated agent is oriented to the idea that the ex post self is an independent self, the ex post resolute self is oriented to the idea of the ex ante self as a controlling self and hence, to the idea of his ex post self not being completely independent."³⁵ In other words, rather than only making choices that we can reasonably believe our future selves would also prefer, we might make choices that we can determine we would prefer *overall*. One way of explaining why we sometimes ought to follow through on an earlier plan, even though we may not currently feel like it, is to assess rational action by assessing one's choice against one's preferences over a longer window of time.

Consider again Ulysses' choice with respect to the Sirens. On a "myopic" theory of rational choice, Ulysses will simply form an intention based on his strongest desires at

³⁵ McClennen, 160

each moment. His strongest desire at the initial moment of choice is to sail by the Sirens while listening to their songs; acting on his strongest desire as he actually hears the Sirens, Ulysses will end up jumping into the sea. By contrast, a “sophisticated” Ulysses, as we have seen, will bind himself to the mast in order to make sure that his (physically constrained) preferences at the time of action are in line with his preferences when he forms his original intention. Sophistication complicates the initial scenario by asking Ulysses to make plans in light of both his present and future preferences, but like the myopic Ulysses, the sophisticated Ulysses always acts on his strongest current preferences. And as Gauthier points out, in neither of these cases does Ulysses get exactly what he would like: the myopic Ulysses drowns, while the sophisticated Ulysses has to be tied to a ship’s mast. The option of resolution asks whether it wouldn’t be more rational, instead, if an agent were able to settle on a plan and *stick to it*. Taking some liberties with the myth, wouldn’t it be more rational of Ulysses simply to decide to listen to the Sirens without leaping out of the boat, and then to follow through with that plan?

The resolute agent settles on plans that she expects will best serve her interests across time, and then follows those plans, even though her occurrent preferences may change. The intuition here is that part of what defines rational agency is pursuing one’s interests as a diachronic agent, not just as a chooser at a discrete moment in time. On the model of resolution, a rational choice is one that serves the interests of both the present chooser and her future self. To put the point more metaphorically, an agent is more likely to achieve the best outcomes if her various temporal “slices” cooperate with one another in making and acting on decisions; one should simply continue to follow one’s intentions

unless there is sufficient evidence to suggest that these (resolute) intentions are inferior to sophistication.

Resolute choice seems to come closer to personal commitment than sophistication. Whereas relying on sophistication would require coercing her future self causally, a resolute model of intention can recognize that answering the question of whether Annie should follow through on her commitment can just require stepping back and looking at a larger chunk of her life. Since human agency is essentially diachronic, determining what Annie ought to do intuitively involves looking at her life across a span of time, not just at some discrete moment. And indeed, both resolute intentions and personal commitments are the sort of practical attitudes that highlight the diachronic nature of agency. Part of what defines personal commitment is the way that it allows us to project normative demands into the future, to levy demands upon our future selves. Both resolution and personal commitment are built on the recognition that we are not actually time slice agents who have to figure out how to coordinate ourselves over time, but rather agents whose existence importantly extends over a period of time.

But despite these similarities, resolute choice does not line up with our picture of personal commitment. Although resolution explains why I must sometimes retain an intention, even against the balance of my current preferences, this explanation is grounded solely in instrumental benefit. My interests as a diachronic agent are best served by sticking to this intention. I may also, sometimes, make a personal commitment as a means to greater instrumental benefit: for example, better health would be one reason for me to make a personal commitment to jog each day, or for Jason to make a personal

commitment to vegetarianism. And indeed, it might be smart to commit oneself only to plans that would be rational on the model of resolution – surely it would make the most sense to reserve personal commitment for those ends that leave one better off at the end of the day. But while sometimes I may make a personal commitment in order to up the odds that I will pursue some instrumentally beneficial course of action, other times (and indeed, I think it will turn out, many more times) I seem able to make a personal commitment as a way of giving some end *significance* in my life.

Again, this is still just a claim about our pre-theoretical experience of personal commitment, but while instrumental benefit may count as some sort of side constraint (we might understandably question whether commitments that are genuinely harmful to oneself are truly binding – a question we will revisit in the final chapter), personal commitments as such don't seem to be rooted in benefit to the committor. Annie's personal commitment isn't primarily about the value it would add to her life, and the same goes for Jason's personal commitment – Jason and Annie have bound themselves, even if their respective commitments fail to serve any purpose, and indeed even if those commitments come at some personal cost. To make the normativity of personal commitment dependent on personal gain would seem to make the commitment too much about the *agent*, as opposed to the end to which she is committing.³⁶ When Annie commits to raising money for the local community theater, this feels like a way of saying just how *important* the community theater is to her, not a claim about whether she will be

³⁶ Notice that this is related to the claim that personal commitments bind us because they are the exercise of a normative power, not in virtue of their content. As an agent, I have the capacity to create new reasons by committing; my personal commitment is normative because I have willed myself to be under its constraint, not just because acting on the commitment is supported by the extant values at play.

happier, or somehow concretely better off if the theater continues to operate.³⁷ Again, when resolute intentions are binding across time this is because reasonable habits of reconsideration keep them in place. Resolution is a way of showing how rational agents who exist through time make responsible choices. The binding force of personal commitments, however, does not seem to be this sort of inertia. Rather, when our personal commitments continue to be binding, this seems to be simply because we have willed it to be so.

We will continue to explore and defend the way in which personal commitments bind us throughout this chapter, and indeed the rest of the dissertation. For now, let's proceed with this suggestion: like resolute intentions, personal commitments need to be assessed against a broader window of time than the narrow present. Unlike resolution, however, personal commitments seem to aim at something more than mere prudential benefit, something like importance, significance, or meaning.

A Planning Theory of Intention

Moving on, let's consider just one more picture of intention, Michael Bratman's theory of intentions as a sort of *plan*. In contrast to models that define rational intention just in terms of rational belief and desire (whether on a narrower or broader window of time), Bratman argues that what it is to rationally hold (and act on) an intention cannot simply be reduced to what it is to rationally entertain other sorts of psychological

³⁷ Of course, having this sort of meaning may make one's life go better. Nonetheless, we can distinguish between making a decision directly to achieve some instrumental benefit, and pursuing a source of meaning that indirectly achieves some instrumental benefit.

attitudes. Unlike classic belief-desire models, Bratman construes intention as a primitive, unanalyzable stance, and thus our interest in this models stems from the claim that this stance is in part defined by a proper set of guiding norms. Let's consider whether these richer norms of intention might be able to recapture the sort of commitment involved in personal commitment.

Bratman's model builds largely on three foundational claims: first, that the paradigm case of intention is a future-directed intention, second, that one-off intentions are typically embedded in larger, more complex plans of action, and third, that these plans are typically only partially filled-in. Let's start with the claim that we ought to take future-directed intention as the paradigm case. For the belief-desire theorist, intentions directed at immediate action have to be central: intentions guide action precisely because they are in part composed of presently occurrent desires. (And then, as we have seen, various accounts have been constructed to extend the force of these intentions into the future.)³⁸ But as Bratman points out, although I can say, "I intend to type this sentence" as I type this sentence, that is an odd way of expressing myself.³⁹ It is much more natural to say before I have started typing, that I intend to do so, and once I have started, that I simply *am* typing. And the relative naturalness of speaking of intention in the future case seems, at least initially, to correspond to something intrinsically future-oriented in intention: intentions are about mapping out what it is we will go on to do, rather than acting primarily as reports of what it is we are doing.

³⁸ Though skepticism has also been raised as to whether such a project is feasible – see Ferrero.

³⁹ *Intentions, Plans, and Practical Reason*, 312.

Second, Bratman points out that one-off intentions are typically embedded in larger plans of action.⁴⁰ For example, an intention to revise this chapter over the coming week is one part of (and really only makes sense in terms of) a larger plan to write a dissertation. More specifically, plans are typically ordered in hierarchies. The goal of writing a dissertation is a higher-order plan; as part of the plan to write a dissertation, I have sub-plans about reading relevant material, writing, conferring with committee members, and so on. Because plans are hierarchically structured, higher-order parts of the plan will control lower-order parts: if it turns out that my sub-plan to read Davidson's *Actions and Events* will not actually serve the higher-order plan of writing this dissertation, then I should abandon that particular sub-plan. And even though I may not particularly feel like reading *Actions and Events* right now, if this sub-plan will serve the end of writing a dissertation, then the higher-order plan to write this dissertation may make it reasonable to read Davidson despite my current preferences. And even when I form a one-off intention that is genuinely not part of a broader, hierarchically-organized plan, (for example, I form an intention to scratch the tip of my nose), this intention will still have to be incorporated into my plans, at least in the minimal sense that it is made consistent with them.

The third and final characteristic is that our plans are typically only partially filled in, in the sense that I may not ahead of time know exactly which steps will be required to meet my eventual goal, or just when I should take the steps that are needed. Limited

⁴⁰ We have already seen this idea, that intentions are paradigmatically embedded in larger plans, in McClennen and Gauthier. Bratman's model differs, however, in that he takes the hierarchical nesting of intentions in plans to have more central significance for the norms of rational intending.

deliberative resources mean that we cannot, in principle, always have all of our plans fully fleshed out. Even if I had the time, energy, and prescience to map out my plans for finishing this dissertation in complete detail (e.g., on May 8 at 1:30 PM, I will..., and at 2:00 PM, I will..., etc), this would not be an efficient use of my time. After all, the nature of action (in particular, the fact that we can plan effectively for the future even on the basis of relatively incomplete plans) means that we do not need to have every detail settled in advance. Rather, by and large, I arrive at particular intentions through a process of filling in more or less broad plans of action as needed.

With this picture in hand, we can see how broader plans control lower-order intentions (and by extension, actions). To start, plans, like intentions, are subject to consistency and coherence constraints. Familiarly, a plan must be at least minimally internally consistent, so that the entire plan can be successfully completed; strongly consistent relative to the planner's beliefs, so that successful completion is consistent with those beliefs; and means-end coherent, meaning that the way in which the agent intends to proceed ought actually to lead toward the goal he intends to achieve.⁴¹ (For example, it would be means-end incoherent for me to plan to spend the weekend in Philadelphia, and yet make no sub-plans about how to get there from my current location in Washington, DC.) These constraints impose two conflicting kinds of pressures on deliberation. The consistency requirements mean that my standing plans act as a filter, ruling out candidate intentions and plans that would be inconsistent with what I currently

⁴¹ *Intentions, Plans, and Practical Reasoning*, Ch.3.2

plan to do.⁴² If I intend now to spend the weekend in Philadelphia, then, absent reconsideration of that plan, intentions to spend the weekend elsewhere, or working at home, are ruled out. In opposition to the filtering function of consistency constraints, means-end coherence will push me to form new intentions, as I fill in the sub-plans needed to achieve my higher-order plans. As we have seen, I don't need to have every component of my plan worked out in advance, (a week ahead of time, I needn't have decided whether to take public transportation or a taxi to the train station), but means-end coherence will require that I fill in these steps as needed.

Now, let's return to Bratman's claim that future-directed intentions should be taken as paradigmatic. As Bratman notes, this fits more naturally with our common ways of talking about intention. Starting with future-directed intentions, however, also seems to lead to a puzzle – how can an intention I form *now* continue to be normatively binding *in the future*? For Bratman, the fact that intentions are typically embedded in larger plans is part of the answer to that question. Let's say that, rather than intending to read *Actions and Events* right now, I form an intention to read it tomorrow afternoon. While both sophistication and resolution try to extend intention based in my occurrent desires to future cases (either by rigging the choice situation, or by extending the window for assessing rational action), Bratman starts by considering an intention in terms of its role in broader, hierarchically structured plans. If my current intention to read *Action and*

⁴² Of course, abandoning my original plan in favor of the conflicting plan under consideration would also maintain consistency. We will shortly see, however, why my current plans have a presumptive weight in choice situations.

Events is still normative tomorrow morning, that is because it is part of my higher-order plan to write this dissertation. As long as the higher-order plan stays in place, I don't need to compare the component intention directly against my current preferences.

As with resolution, part of the story about when we ought to revise or abandon our intentions is in terms of instrumental benefit. Both plans and intentions are often held in place by purely pragmatic considerations; after all, re-consideration and revision of one's plans can be costly. If I form an intention to walk to school today as part of a plan to do work at the library, second-order pragmatic considerations may support simply following through on this intention. Spending any significant amount of time reconsidering a relatively trivial intention would likely be a waste of time. (If I stop to reconsider whether the bus is a better option, I'd now have to spend time checking the bus schedule, seeing whether I have correct fare, etc.) In addition, plans need to be fairly stable if agents are to be able to coordinate their own action over time. If my transportation plans, once made, aren't relatively stable, this will complicate coordinating those plans with other intentions for my day. So, if I may or may not follow through on my plans to walk to school this week, I will have a hard time knowing whether or not to buy a weekly bus pass, whether it's important to carry an umbrella in my purse, and so on.

Besides instrumental benefit, Bratman appeals to another, distinct factor in the stability of plans, what he calls 'agential authority'. As we briefly saw above, plans that span across time can help to coordinate an agent with herself. In doing so, Bratman claims, these diachronic plans also help to constitute who I am – as Bratman sometimes

puts it, they determine *where I stand*.⁴³ Here, Bratman's claim is that plans figure among the sorts of overlapping psychological states that unify an agent across time on a broadly Lockean view.⁴⁴ Plans can help to unify a temporally extended agent through their own characteristic stability, and also by helping to organize and maintain connections among hierarchically lower items, like desires and intentions.⁴⁵ Because of this role, Bratman claims that these plans have agential authority – they can determine where *I* really stand, as distinct from merely being urges that I experience. So, for example, why stick to a plan to lead a vegetarian lifestyle when I'm ravenous and all that is available is some tasty but meat-filled bit of food? One answer is that if I have been operating with a standing plan to be a vegetarian, that plan is in some sense a part of me. Unlike a one-off desire to eat a delicious piece of sushi, my plan to be a vegetarian *speaks for me*; unlike mere preferences or desires, it determines what I “really” want.

Of course, even though their persistence helps to maintain my personal identity over time, I might still feel alienated from some of these plans: for example, I might regret, or even resent that circumstances force me to maintain a certain plan of action. Thus, Bratman adds one more condition: a plan has authority if it plays a role in maintaining the agent's identity over time, *and* the agent is satisfied with that plan.

While Bratman is explicitly building here on the notion of ‘satisfaction’ raised by Harry

⁴³ Though the metaphor of an agent's “location” is not a totally new one, Bratman appears to take the metaphor more seriously (and perhaps more literally) in his most recent work (“Anchors for Deliberation”).

⁴⁴ *Structures of Agency*, 29. To be precise, a sub-class of relatively general plans that Bratman calls ‘policies’ are more likely to play this role. Since nothing in our discussion trades on this distinction, I'll stick to the language of plans for simplicity.

⁴⁵ *Structures of Agency*, 268.

Frankfurt, Bratman's own rendering is a little different.⁴⁶ To be satisfied with a plan is for that plan not to be challenged by any other plans; it is for me to be able to follow through on the plan in question without having to compromise other plans I may have. While this allows for some internal conflict between plans and lower-order intentions or desires, to be satisfied entails that no other plan is interfering with the satisfactory plan's functioning in supporting coordinating Lockean ties. To sum up, temporally extended plans (with which I'm satisfied) have a certain authority in determining where *I* stand at a given time, given their role in maintaining my personal identity. Furthermore, this authority to determine where I really stand means that a plan may survive reconsideration, even when it isn't strictly supported by my preferences at that moment – the way in which my extended and overlapping plans make me, *me*, means that they may continue to be rational even when they aren't actually supported by my current preferences. Thus, on Bratman's picture, intentions that I formed at some time in the past can continue to be authoritative (in the sense that they continue to determine what it is I ought to do) through some combination of second-order pragmatic reasons against reconsideration, and the role the plans in which these intentions are embedded play in constituting my personal identity.

Bratman's model of intention does seem closer yet to our picture of personal commitment than any of the other theories we have considered so far. Like Bratman's

⁴⁶ In particular, in "Identification, Decision, and Treating as a Reason", (in *Faces of Intention*), Bratman argues that depression and other states show satisfaction is insufficient. Furthermore, he claims, satisfaction with mere desires lacks the right connection to our temporal persistence.

intentions, personal commitments are paradigmatically future-oriented. Though perhaps I can really commit to do something right now, (“That’s it! I’m committing to getting this house clean, starting right now!”), personal commitment is most often a stance toward one’s own *future*. (Compare this to sophistication, which seemed to tie rational intention too closely to one’s strongest occurrent desire to be an account of personal commitment).

In addition, Bratman argues that an agent’s plans figure among the elements that help to unify the different temporal slices of an agent into one person. Our plans aren’t just *useful*, they seem to match what Bratman calls “agential authority” in determining where we stand. Though we will have to return to this question in subsequent chapters, it at least seems immediately plausible that personal commitments could also help to determine where an agent stands – if I were to wake up tomorrow without caring at all about any of the personal commitments that had mattered deeply to me the day before, we might conceivably respond by saying I was no longer quite the same person. (Compare this to resolution, which appealed only to instrumental benefit in explaining how our intentions can have diachronic reach.) On this model, persisting resolutely in my intentions can help me to get more of what I really want by resisting temptation and fluctuating preferences. For Bratman, plans can do this, and also can help to make it the case that my future self who gets what it wants is part of the same person as my past self). Thus, in drawing on normatively richer resources, Bratman’s intentions and plans do seem to come closer to our pre-theoretical picture of personal commitments. In what follows, however, I want to argue that even Bratman’s model of intention falls short of

personal commitment in two important ways: even Bratman's intentions give rise to normative requirements and second-order reasons of efficiency, rather than first-order reasons to do as one intends, and they seem to fall short of the sort of meaning attached to typical personal commitments. And if even Bratman's relatively rich model of intention can't recapture personal commitment, then it seems that ultimately any theory of intention will fail to recapture personal commitment.

First, unlike intentions, personal commitments seem to give rise to new reasons for action. When Shannon commits to developing patience, this at least seems to add to the reasons that she previously had to changing her temper. Of course, forming an intention clearly changes what it is an agent ought to do. In particular, most theories recognize that intentions (and plans) change what it is we ought to do in two broad ways: first, intentions specify standing second-order pragmatic reasons of efficient deliberation, and second, they trigger certain sorts of consistency constraints. We've already seen how intentions can specify pragmatic reasons: for example, when Jason forms a plan to become a vegetarian, his standing reason to deliberate efficiently (and hence to abide by reasonable habits of reconsideration) will now include a more specific demand only to reconsider his vegetarianism if doing so seems relatively likely to change his mind. In addition, the demands of consistency mean that Jason should either follow through on his intention (when appropriately situated), or else revise or reject it. So long as Jason is being rational and retains his plan to become a vegetarian, Jason ought to go on to form new intentions that will help him to lead a vegetarian lifestyle, and to reject those

intentions that involve consuming meat. If Jason is not going to follow through on his plan, then consistency requires that he give up his intention to become a vegetarian.

These consistency constraints line up with what John Broome has called ‘normative requirements’. To put the point in terms of deontic logic, normative requirements are “wide-scope oughts”. So, it is the case that you ought [if you intend to be a vegetarian, not to eat meat]. Given the placement of the brackets, the consequent “you ought not eat meat” is not detachable. In less formal words, this means you don’t actually have a reason not to eat meat. What you do have is a disjunct: either you must drop the antecedent, the intention to be a vegetarian, or else you must follow through on the consequent, you eschew eating meat. To emphasize the point once more, these options are symmetrical: rejecting the antecedent or fulfilling the consequent are equally good ways of doing as one ought.

The demands of personal commitment, by contrast, don’t seem to be symmetrical in this way. Once Annie has truly made a personal commitment to raising money for the community theater, it is simply not the case that abandoning her commitment is just as good of an option as following through on it. If, after a sufficient period of time, her friend Elena draws attention to the fact that Annie never raised the \$5,000, responding that she has simply changed her mind won’t fully exonerate Annie. Rather, it will reflect (at least somewhat) poorly on Annie’s character, or perhaps implicate her self-respect, if she sees no need even to explain her decision not to follow through on her personal commitment. Obviously, this is not to say that personal commitments are irrevocable.

Rather, the claim here is that getting out of a personal commitment seems to involve some work; it requires that the agent at least excuse herself in some way.

The reason that Annie must excuse herself, I want to suggest, is that making a personal commitment gave rise to new reasons for action. Now, as I noted above, even bare intentions can trigger both normative requirements and reasons – these reasons, however, are specified second-order considerations of reasonable, efficient deliberation. It is inefficient, for example, to reconsider intentions unless one has good reason to do so. Without good evidence to do so, one ought not reconsider one’s intention to go to the library today, as that would be a waste of time. Similarly, personal commitments seem to trigger considerations of consistency, coherence, and pragmatic efficiency. Here, however, I am starting to make the additional claim that personal commitments provide *novel* reasons for action, both positive first-order reasons, and what Joseph Raz calls ‘exclusionary reasons’, second-order reasons that exclude certain first order reasons from consideration.⁴⁷ This would mean that my personal commitments can change what it is I have most reason to do, both by adding reasons to act upon my commitment and by silencing considerations that would have been weighed against a mere intention, or plan. So when Jason makes a personal commitment to become a vegetarian, in addition to various normative requirements and second-order pragmatic reasons, he gives himself two new sorts of reasons. Among Jason’s new first-order reasons for action will be a reason not to eat meat or meat products, a reason to take dietary supplements as necessary, etc. And in virtue of the new exclusionary reasons that apply to him,

⁴⁷ As we’ve noted before, Raz calls this combination a ‘protected reason’, (*The Authority of Law*, 18).

considerations that would otherwise figure in deliberation are irrelevant: unless Jason reconsiders his personal commitment, the fact that he finds meat delicious, that he will now have to learn new recipes, etc., are all silenced. These values are not outweighed; they are simply no longer relevant to Jason's deliberation about what to do.

That personal commitments seem to change our first-order reasons for action is perhaps the most clear when I commit to one end out of a group of equally eligible but incommensurable ends. Before making a personal commitment to working for D.C. voting rights, for example, volunteering to work for Congressional representation would be a valuable activity for me, one worth pursuing, but not something that I am *required* to do in any sense. I could have equally well chosen to spend my free time volunteering for a different organization, or pursuing some hobby. Once I have committed, however, I will be violating my personal commitment if I consistently miss DC Vote meetings, or fail to find some other organization with which to work. I will be worth blaming if I do so. And the fact that I feel like going out to dinner with a friend, which may have weighed (albeit slightly) against a mere intention, will be irrelevant if I have committed to volunteer at one of their events. Other times, it may be the case that I already had sufficient reason to pursue the end to which I commit. In the case of a resolution to exercise more regularly, I no doubt had prudential reason to exercise more often anyway; but by making a personal commitment, I add to the reasons that I have for exercising, making it the case that I am *required* to exercise as planned.

And in addition giving rise to new reasons for action, personal commitments are typically much more *significant* than a mere plan or intention. This is a theme that we have continued to see throughout this chapter: a personal commitment is a way for an agent to take a stand on some goal or end, to *stake herself*, at least in part, on its value. Because basic Humean models can only recognize one sort of practical weight (being desired or preferred), they can only attribute *more* weight to personal commitments: a commitment would have to involve having a stronger desire to do *x* or *y*. Personal commitments, however, seem to be different in kind – it’s not just that they happen to be relatively strong, they also seem to be *important*.⁴⁸ When we think of the way in which Jason cares about vegetarianism, or Shannon about improving her temper, it seems that personal commitments *matter* in a particular way.

One might be tempted to think that we can recapture this importance by appealing to the authority Bratman attributes to some of our plans. (Remember that Bratman argues that plans can take on a sort of authority by helping to constitute where an agent stands on some question.) And indeed, the fact that we have to make personal commitments in a deliberate way, putting our future selves under normative constraint, also sounds like personal commitments generally speak for what an agent genuinely wants, rather than simply reflecting our passing preferences.

Nonetheless, I don’t think personal commitments are normative for the sorts of reasons that Bratman describes. As he describes them, plans do reflect what agents really

⁴⁸ Of course, this is not yet to say that personal commitments always outweigh competing desires, however strong. The point here, to repeat, is merely that committing gives us reasons that are different *in kind*, not just stronger.

care about. Their special authority, however, comes from the fact that they contribute to my *numerical* personal identity – these plans can speak for who I am because they help to make me the same “me” over time. And while some (perhaps even all) personal commitments may play this same role, I don’t think they derive their normativity in exactly the same way. While authoritative plans, on Bratman’s suggestion, are normative because they help to maintain personal identity over time, personal commitments seem to be important primarily because they contribute to my identity in a thicker, narrative sense. Let me explain. Our personal commitments help to direct our lives; they are one way of constituting what we find meaningful in our own lives. Certain ends and values tend to be meaningful in almost all human lives: family, creative activity (whether that is found in work or leisure), and so on. Other sorts of values just happen to be meaningful for particular individuals. I personally find philosophy to be a very meaningful part of my life, and though I can tell a story about how this came to be, relatively little of that story has to do with my deliberate choices. By contrast, my sister (who in many ways is very similar to me) values philosophy very little; it is certainly not a meaningful part of her life. Again, this is largely just a factor of the person she happens to be.

Finally, there seem to be some goals, or pursuits, which make our lives meaningful, and which are idiosyncratically and deliberately chosen. As Cicero points out, such pursuits are likely to be relatively rare.⁴⁹ Nonetheless, this is an experience with which at least some of us are familiar. Before committing, the ends of becoming a more patient person, embracing a way of life that promotes certain ethical goods, and

⁴⁹ I.107-115

contributing in a particular way to one's community, are all equally good ways of spending my free time. Indeed, in the absence of a personal commitment to any one of these ends, I might devote some of my free time to each, or to other worthwhile ends. Once I have made a personal commitment to one of these ends, however, that end becomes meaningful for me in a way that it was not before. Whereas initially, I was faced with a variety of incommensurable pursuits, any one of which would have been a good use of my time, after committing to becoming a more patient person, that end has a special importance for me. Part, though not all of that importance, is reflected in the fact that I now have reason to practice meditation, or go to therapy sessions, that defeats (at least on some occasions) the reasons supporting other worthwhile goals.⁵⁰

In addition, the effort I expend in working on my personal commitment, and the successes and failures I experience in its pursuit, reflect on *me* in a particular way. Of course, there are different ways in which one might deliberately and freely choose to make some end meaningful. For example, an agent might make a religious vow, or a promise to a loved one. My claim is that, if we can ever make these sorts of choices in the privacy of our own hearts, then we can also create this sort of idiosyncratic meaning through a personal commitment.

⁵⁰ To be clear, I am certainly not arguing that the reasons of commitment defeat all other considerations. My commitment to becoming more patient will be trumped by most if not all of morality, and no doubt can be trumped by strong prudential considerations as well. My point is only that, in committing to cultivate patience, I seem to have bound myself to this end such that it now gives rise to requirements (which may in turn be outweighed), rather than a reasonable option. For interesting work on the distinction between these two sorts of normative standings, see Joseph Raz on reasons of respect and engagement, e.g. in *Value, Respect, and Attachment* (CUP, 2001), and Joshua Gert on justifying and requiring reasons, e.g. in "Two Concepts of Rationality", *Southern Journal of Philosophy*, 41(3): 367-398.

Personal commitments, like vows and certain promises, are able to play this meaning-giving role because they have an aspirational aspect⁵¹: in making one of these commitments, I have, to some extent, *staked myself* on the end in question. Indeed, some of my personal commitments – the deepest and strongest among them – may form part of my self-understanding. For example, if someone makes a very solemn personal commitment to environmentalism, and yet finds herself persistently yearning after large, shiny SUVs, she may perceive this yearning not just as impermissible, but also as (perhaps weakly) threatening her very sense of self. Indeed, though I won't go into any more detail here, I will argue in the final chapter that our power to make personal commitments seems to be grounded in something like an agent's authority to be self-constituting, and the role of aspirations in a meaningful human life.

The preceding discussion suggests that intention, as such, can't underwrite personal commitment, at least not on our everyday experience of commitment. Pre-commitment, resolution, and plans give a greater *degree of stability* to our intentions, but none can give intentions the different *kind of robust normative weight* that personal commitments seem to have. In considering the ways in which mere intention falls short, however, we have gained some lessons as to what desiderata we should look for as we consider alternative models of personal commitment. For one, we need to explain the way in which personal commitments seem to give rise to new reasons for action. If

⁵¹ I owe thanks to Alec Walen for this language. Of course, not all personal commitments will be aspirational in a straightforward sense: I may make a personal commitment to lounge about more regularly. Nonetheless, it will generally be the case that I make that commitment in service of some broader aspiration, such as to become a more laidback person, or perhaps just to exercise my "commitment muscle", so to speak.

personal commitments do give rise to requirements on action, how and why do they do so? Furthermore, we might take with us something like Bratman's concept of "agential authority". Though I have suggested here that the authority of personal commitment isn't just grounded in the need for continuity in personal identity, we may be able to use a related account, attributing some authority to personal commitments in virtue of their role in giving rise to shape and meaning across an agent's life. In the next two chapters, I want to continue to build on these lessons while turning our attention to alternative accounts of personal commitment. In particular, we will examine whether a full theory of personal commitment can be provided by either a virtue, in particular that of integrity, or else a transactional normative power such as promising.

Personal Commitments as a Matter of Integrity

In Chapter 2, we proposed that personal commitments give rise to new reasons for action. Unlike forming an intention, which merely makes us subject to certain consistency constraints, (i.e., I must either act on my intention or else change what it is I intend to do), making a personal commitment seems able to change what it is I in fact have most reason to do. In the present chapter, I want to consider whether we can do a better job of recapturing this aspect of personal commitment by appealing to integrity. That is to say, are commitment-based reasons essentially reasons to avoid threatening or damaging my integrity?

Intuitively, there is a connection between personal commitments, as we have described them so far, and integrity. (Indeed, I will have more to say about this connection in the final chapter.) Both personal commitments and the demands of integrity seem connected to what agents, as particular individuals, idiosyncratically take to be important. Unlike the virtue of honesty, for example, which at a general level makes the same demands of everyone, integrity involves staying true to what *you* in particular deem important. Furthermore, appealing to integrity might be able to help us to unpack the language of “taking a stand” which we have occasionally used so far. “[I]ntegrity is essentially *moral courage*, the will and willingness to do what one knows one ought to do.”⁵² Intuitively, to say that someone has integrity is to say that she is

⁵² Solomon, 168

faithful to *her* own values, that she stands by them when called upon to do so. In making these demands, integrity seems both to give rise to a notion of commitment, by pointing to the importance of standing by one's beliefs, and also to recapture the way in which our commitments can be *personal*. Whether or not they are widely shared, maintaining integrity seems to demand fidelity to my own particular commitments.

Our common understanding of integrity, however, is actually a bit more complicated than this, containing two distinct strands. The first strand, on which we have been focusing so far, casts integrity as sticking to one's own beliefs and commitments, whatever those beliefs happen to be. In this sense, personal integrity is sometimes analogized to the integrity of an object: a person has integrity if her "pieces" (like her commitments, or perhaps her psychological states) are more or less internally well-ordered, and remain constant in the face of external opposition or flux.

The second strand, however, casts integrity primarily as a *moral virtue*. On this understanding, integrity means staying true to one's moral convictions. Moral integrity is the virtue of prioritizing one's moral beliefs over other values, (both one's own extra-moral priorities and the moral values of others), and remaining faithful to those beliefs. Even on this conception, integrity seems to be compatible with a range of personal values, since we can attribute integrity to people with whom we have at least some moral disagreement. I, for one, admire what I take to be the significant integrity of several Roman Catholic friends, although I strongly disagree with them on a number of moral questions. Nonetheless, a moralized reading of integrity may rule out some substantive commitments. We might, for example, call Unabomber Ted Kaczynski a man who was

quite seriously committed to slowing the spread of technology, while being uncomfortable with calling him someone with a moral virtue of integrity. The clearly immoral nature of a commitment that involves murdering people as a means seems to be incompatible with a moral virtue, however much it fits with the other parts of that virtue. We are also unlikely to attribute integrity to someone like a committed sycophant.⁵³ However thorough-going one's servility and hypocrisy may be, and even if these attributes contingently end up serving a greater good, they seem incompatible with a moral virtue of integrity.

In any case, for our purposes I will put the moral conception of integrity to the side. As I will suggest in the next chapter (where we consider whether a personal commitment is really just a promise made to oneself), I think it is a mistake to take personal commitments as having characteristically *moral* significance. Though personal commitments, if we can make them, are in part defined by their weighty normativity, I think calling that normativity "moral" is a mistake. And if that suggestion turns out to be plausible, then personal commitments can't be best understood in moral terms, whether as a moral virtue or some other sort of moral commitment. Thus, in this chapter, I will focus on whether personal commitments look like the demands of integrity understood broadly as something like the character trait of staying true to one's values, or as an internal sort of wholeness. Of course, even if we focus on integrity as a sort of wholeness there are a few different ways in which we might understand what it means to have integrity, so we will consider three broad proposals: that integrity is a matter of

⁵³ Cox, LaCaze, & Levine

maintaining one's fundamental identity, of fidelity to some central set of values, or of creating and sustaining a coherent life story.⁵⁴

Integrity as Identity

The discussion of integrity as maintaining one's identity has its roots in Bernard Williams's rejection of impartial theories of morality, including both utilitarianism and Kantian deontology. Such moral theories, Williams familiarly argues, wrongly subordinate the personal point of view and choke out the important particularity of individual human lives. To resist this tendency, Williams puts forth an "identity" notion of integrity, on which integrity involves staying true to one's most basic self, and then argues that the value of maintaining this sort of integrity is important enough to override even the demands of impartial morality.

To review his proposal briefly, Williams begins by distinguishing between two sorts of desires. Some desires are conditioned on my existence: if I happen to live to see tomorrow, I want to eat lunch and to prepare for class next week; if I live long enough, I would like to buy a house, travel to Asia, learn German, and so on. When I claim to have this sort of desire, in essence I am saying "if I happen to have the opportunity, I'd like to *phi*". In other cases, the order is reversed, and it is the value of my existence that is conditioned on a desire. That is to say, I am sometimes moved to make it the case that I live long enough to satisfy a given desire. Williams calls this sort of desire a 'ground project'. Finishing this dissertation, for example, isn't just something I want to do if I

⁵⁴ In what follows, I rely heavily on the distinctions put forward by Calhoun (1995) and Cox, La Caze, & Levine (2005).

happen to continue living; rather, it provides me with a reason to continue living – while eating lunch is something I want to do if I happen to get out of bed, editing this dissertation is something I care about enough to get me out of bed in the first place.⁵⁵ Ground projects pull us forward into our own futures, and they help us to make sense of our own lives.

For Williams, integrity is a matter of staying true to one’s ground projects, that set of commitments and ends that happen to make one’s life meaningful and valuable from one’s own point of view. The phrase “that *happen to*” is important here; which commitments and ends figure among my ground projects is a deeply contingent and idiosyncratic matter. As a statistical matter, some sorts of ends are common ground projects: many people happen to be devoted to their families, to their religions, and so on. But no one need have these particular ground projects. And despite this contingency, to say that something is someone’s ground project is a very strong claim. At least on Williams’s understanding, if being a vegetarian is one of Jason’s ground projects, then integrity requires maintaining that vegetarianism not just in the face of a conflicting desire to eat meat, but even in the face of moral demands to do so. (To extend a case from one of his own examples, Williams would claim that a ground project of vegetarianism could militate against eating meat, even if an evil tyrant threatens to kill others if one does not.)

⁵⁵ I do recognize that there is some humor in claiming that a dissertation provides someone with a reason to live, as opposed to high levels of anxiety and stress. But this only underlines the fact that our ground projects needn’t be pleasant, at least not constantly, or in obvious ways. Ground projects give *meaning* to our lives, not pleasure.

How do ground projects get this very strong normative force? Clearly the value of this sort of integrity isn't moral, since Williams sets it up precisely as a limit to the reasonable demands of morality. Rather, a person's ground projects matter simply because they are what make his life valuable *for him*.⁵⁶ If Jason really has a ground project of vegetarianism, this is to say that being a vegetarian is one of the pursuits that make Jason's life worth living, *for him*. This needn't imply that Jason thinks human lives as such aren't worth living if omnivorous, but simply that Jason might no longer take his own life to be worth living if he were forced to abandon this project.

To be clear, that our ground projects are what give us reason to live is not to say that, if someone's projects were thwarted, that person would necessarily kill himself, or somehow die of a broken heart. First of all, agents may have a set of ground projects. In this case, while the loss of one may be damaging, there are other ground projects to give meaning to their lives. If Jason is forced to give up vegetarianism, for example, his love of his family may be enough to impel him to muster on. And even if most or all of his projects were frustrated or destroyed, an agent might happen to find a way to keep going. "Other things, or the mere hope of other things, may keep him going. But he may feel in those circumstances that he might as well have died."⁵⁷ This point is crucial: it is not just that it would be painful to abandon my projects, or unfair to ask it of me. The claim is

⁵⁶ To come back to an earlier point, ground projects are the sort of thing on which an agent's continuing to care about her own existence is conditioned. And if she is to be motivated to act on any reasons at all, she must be motivated to care about her own continued existence. Of course, Williams makes the connection from this claim about an agent's subjective conditions for valuing of her own existence to the objective value of her ground projects by being an internalist about reasons. But an externalist might rephrase the point to say, I would lose all motivation to follow the reasons I objectively have. Though that's a weaker claim, it still seems to invest ground projects with a pretty significant weight.

⁵⁷ Williams, 1981, 13

that the reasons given by my ground projects cannot be trumped by other sorts of considerations because, without my ground projects giving me reason to care about my own future and what I am to do, I have no reason to continue acting at all. If I am unable to continue to pursue my ground projects, I *may* somehow find a reason to go on, but it will be perfectly understandable if I don't.

Given this model, can we understand personal commitments in terms of ground projects? Ground projects are indeed similar to the picture we have been drawing so far of personal commitments in that each are a sort of weighty, but idiosyncratic, project. In both cases, what matters is not the "objective" importance of a given end, but the role that end takes in my life. That is to say, ground projects and personal commitments both seem to give rise to a particular sort of meaning in my life. Furthermore, this meaning seems to have normative weight, at least sometimes changing what it makes most sense for an agent to do. Apart from her personal commitment to becoming more patient, for example, it might make most sense for Shannon to castigate a subordinate at work who seriously messed up an important account, (assuming she doesn't take the dressing-down to disrespectful extremes). Giving vent to her temper would motivate the subordinate to be more careful in the future, and would make Shannon feel much better. Nonetheless, a personal commitment or a ground project to improving her temper seems able to make it the case that, *for Shannon*, what matters most here is keeping her cool. And as with ground projects, I imagine making sense of this claim will involve appealing to the sort of *importance* that an agent can idiosyncratically give to an end. Like ground projects,

personal commitments seem to give rise to weighty reasons because they have a special sort of *significance* in our lives.

If personal commitments can really vary in content and significance in the way we've described, however, they clearly cannot be simply identified with ground projects. Though integrity-as-identity can explain one way in which an agent's commitments can give rise to demands, the particular way in which it grounds those commitments is too strong for our purposes. Ground projects are normative because they serve as a condition for caring about our own lives; hence, only our most central, most closely held ideals can serve as ground projects. While some personal commitments may be of this sort, many will not. A personal commitment to one's religion, for example, might serve as a ground project, but at least in our day-to-day experience, we also seem able to make personal commitments to raising money for a charity or even to learning a foreign language, commitments that are unlikely to be crucial to that person's sense of self.⁵⁸ If we are able to recapture the range of personal commitments that we take ourselves to make, it won't be in terms of something as weighty as a ground project.

In addition, agents seem able just to find themselves with ground projects, whereas they must actively create personal commitments; as we said above, one can "just happen" to have a given ground project. Indeed, it would seem odd to say that someone could choose those ends on which his interest in his own life hangs. It seems much more likely to be the case that he would simply find herself with a ground project, (as when a new

⁵⁸ Of course, an agent *might* have a second-order ground project of remaining true to his personal commitments, giving less weighty, first-order personal commitments a derivative force. Which ground projects a person has, however, is a contingent matter, and there are (at least according to Williams) no necessary ground projects that all agents must have.

father immediately feels that he would sacrifice his own life for that of his child), or perhaps that a ground project would develop over time, (as when, after years of practicing a religion, one finds that one can no longer imagine oneself without it). And even if we could, at least sometimes, simply will a ground project into being, it is never the case that we can simply find ourselves with a personal commitment, at least in the sense in which we have been using this term in this dissertation. A personal commitment results from the exercise of a normative power. *I* commit *myself*. As the exercise of a normative power, this is only something I can do self-consciously and deliberately. A personal commitment is an active “self-binding”.

Finally, while Williams’s argument can show how my ground projects are protected against external threats, it says nothing about the threat *I myself* might pose to my own commitments. Because ground projects are defined by a psychological sort of caring, they are generally accompanied by a fairly strong motivation in their favor.⁵⁹ Thus, for integrity-as-identity, threats to integrity are primarily external, when a third party or the norms of impartial morality demand that one sacrifice one’s projects. Personal commitments, by contrast, seem to be essentially normative creatures, and the motivation to abide by them can wax and wane. This means that my personal

⁵⁹ As Henry Richardson pointed out, ground projects might come apart from actual motivation in cases of serious depression or *accidie* (see Aquinas, Stocker). When deeply depressed, agents can come to the point where they no longer care about even their most meaningful projects. Nonetheless, I take “ground projects” to be a fundamentally motivational category, defined in terms of its role in an agent’s psychology, whereas “personal commitment” is essentially a normative item.

commitments can be threatened both from the outside, by people trying to tempt or force me to violate my commitments, and also from the inside, by weak will and apathy.⁶⁰

Of course, an agent might be ambivalent or conflicted about her ground projects, or deceived about what her ground projects really are or what they entail. Such conflict or deception might threaten her sense of self and her sense that her life is worth living, but this alone fails to address the range of threats that we can pose to our own personal commitments. In the case of a ground project, if it turns out that I simply no longer care about the end in question, that alone is enough to show that I am no longer bound by that end. My ground projects are normative only insofar as they are the precondition for me finding my own life meaningful. My personal commitments, by contrast, seem able to bind me even when my strongest inclinations point in another direction. Jason cannot simply excuse himself from his personal commitment by saying that he *really* feels like eating meat today, or no longer feels particularly moved by the considerations that caused him to become a vegetarian: eating meat just because he had a strong desire to do so would be *violating* his commitment, not rescinding it. Of course, as I'll stress throughout this dissertation, I am not claiming that we can never escape our personal commitments; at least in the right circumstances we clearly can. Jason's personal commitment to vegetarianism is not irrevocable: he might decide to abandon his commitment because he discovers it's causing anemia, or finds non-factory farmed sources of meat with which

⁶⁰ Of course, as we have suggested above, and will claim in more detail in the final chapter, a personal commitment may only be able to survive a certain degree of contrary motivation. The fact remains, however, that personal commitments are completely compatible with some disinclination, and in fact, it may be in moments of temptation or apathy that the fact of commitment becomes most apparent.

he's comfortable. But as these cases illustrate, there are constraints on exiting our personal commitments, and these exit constraints outstrip psychological shifts alone.

To sum up the difference, imagine that despite a commitment to vegetarianism, Jason wakes up one morning with total apathy towards that end. If Jason's commitment were just a ground project, he is no longer bound by it. The normative force of a ground project is grounded in its psychological force in his life – it is important because it happens to be something about which the agent cares so deeply. By contrast, if that commitment was what we have been calling a personal commitment, apathy alone does not seem enough to let Jason off the hook. He must at least find a way to excuse himself from his personal commitment, or else take on some normative remainder, a kind of blameworthiness. The personal commitments that we have been describing are not just objects of deep care; they can bind us even when we don't particularly care.

Integrity as Fidelity

The “life or death” importance that characterizes ground projects seems too monumental to recapture personal commitments as a class. Though our personal commitments feel important, we have also characterized them as the sort of thing we can revise, or even revoke at will. We might, instead, turn to a picture of integrity as fidelity to a slightly broader range of commitments, commitments which may not play a part in one's essential identity, but about which one nonetheless *cares*. On this sort of approach, integrity is understood simply as being true to those ends in which one is significantly invested.

One familiar model of caring is that put forward by Harry Frankfurt. As Frankfurt says, “A person who cares about something is, as it were, invested in it. (...) Insofar as the person’s life is in whole or in part *devoted* to anything, rather than being merely a sequence of events whose themes and structures he makes no effort to fashion, it is devoted to this.”⁶¹ Less metaphorically, Frankfurt analyzes caring as a complex of first and second-order desires. When I care about *phi*-ing, I have a first-order desire to *phi*, a second-order desire that my desire to *phi* actually move me to act, and a second-order desire to continue to want to *phi*. Because of this last component, Frankfurtian “carings” help to unify agents across time, maintaining a certain degree of continuity in what they want and how they act. In this respect, they are reminiscent of Bratman’s plans.

Like Bratman’s plans, caring nicely focuses on the diachronic aspect of attitudes like personal commitment: part of what it is to care is for one’s desire to persist across time, but also, for one to *want* that desire to persist across time. Similarly, personal commitments seem in part to be defined by diachronic persistence. It would seem odd for Jason to claim that he had made a personal commitment to be a vegetarian for precisely the next two meals, or even just until he changed his mind. Again, this is not to say that either caring or personal commitment are irrevocable or unalterable. Jason may eventually stop caring about vegetarianism, or may revoke a personal commitment, but both my personal commitments and the objects of my care characteristically have a persistent and robust role in my life. More than planning, however, the language of

⁶¹ *The Importance of What We Care About*, 83

“caring” also intuitively sounds like the way in which we might talk about our personal commitments. When trying to convey that Annie has committed herself to raising the needed funds, I may say that she “really cares” about helping the community theater in this way, or that she cares about following through on her personal commitment.

Speaking in terms of caring seems to convey the sort of attachment that is generally associated with our personal commitments. Indeed, it would be unlikely (though perhaps not unimaginable) for Annie to say that she had made a personal commitment to an end about which she didn’t really care.

Yet while we can very naturally talk about personal commitment in the language of caring, and while there are areas of overlap between the two, I don’t think personal commitment is equivalent to caring in Frankfurt’s sense. If we find our working picture of personal commitment persuasive, then personal commitments do not seem to be solely made up of desires, however complexly structured. In the first place, as we claimed in Chapter 2, the normative force of personal commitment seems to outstrip that of bare desire.⁶² Secondly, (and relatedly), personal commitments seem to involve a richer set of psychological states than desire alone. As Shannon is working on her personal commitment to developing patience, she may feel satisfaction as she makes progress, pride when she resists the temptation to yell or curse, disappointment or even shame when she fails, stress when the demands of her personal commitment come into conflict

⁶² In fact, we actually claimed that the normativity characteristic of personal commitment outstrips that of belief and desire joined together. But, of course, if the norms of the two states together aren’t sufficient to recapture the norms of personal commitment then that implies that the norms of desire alone won’t be sufficient either.

with the time and energy needed for other projects, and so on. Of course, these richer sorts of emotional elements could just happen to be associated with a Frankfurtian caring. Intuitively, however, the sort of picture that we described in Chapter 1 seems to underwrite a slightly stronger claim, that these sorts of emotions are part of what characteristically defines a personal commitment.

So let's turn, instead, to a notion of 'caring' that incorporates these richer psychological states. Agnieszka Jaworska, for example, offers one such account. It is difficult to give a simple definition of her sense of caring, but Jaworska offers a few illustrations. For one, we can think of the way in which toddlers often care about independence – they want to do things for themselves, and this desire is associated with certain characteristic emotions. They are engrossed by their efforts to learn new skills, like feeding and dressing themselves, frustrated by parental attempts to intervene, and proud of their successes. Alternatively, we can think of what it is like to care for another person. Here again, Jaworska points to how early on a child develops what we might call the capacity to care for another human being – as a young toddler, he can already be deeply attuned to physical cues about a parent's emotional state, and impelled to offer comfort or some sort of solution when that parent is upset.⁶³ What these examples of caring share in common is a “sustained and caring interest” in the object of care, a “pattern of emotional investment retained over time”.⁶⁴

Interestingly, Jaworska's account of “caring” seems both simpler and richer than Frankfurt's in different ways. It is simpler in that Jaworska takes caring to be a capacity

⁶³ 485

⁶⁴ 487

that humans develop quite early in their lives, before they have fully mastered epistemic norms and other sophisticated forms of evaluation.⁶⁵ Unlike Frankfurt, Jaworska does not think that caring is essentially related to reflexivity, the ability to evaluate the desires with which we find ourselves. Her account, however, is more complex in appealing to thicker psychological states than mere desire. “[W]hat I mean by an emotion is a more enduring state, an ongoing psychic orientation, compromised of various interrelated emotional episodes and dispositions to experience subsequent concurring emotional episodes, unfolding intermittently over time, waxing and waning, at least partly in response to the context at hand.”⁶⁶ To be able to care, on this view, seems more related to being a “*valuer*”, than being just a “*desirer*”.

Jaworska’s analysis of caring does seem to recapture the affect that is characteristic of personal commitment. If I am moved to make a personal commitment, this is generally because the object of my commitment feels important to me, in an emotionally laden way. Though this feel may not be a necessary component of personal commitment – as we have said, Jason may need to do more to revoke his commitment than just pointing out he no longer cares about being a vegetarian – we might nonetheless call this a seriously defective case of personal commitment. Having the appropriate affect seems both like the more common case in personal commitment, and also what we could consider preferable from the point of view of the committed agent’s flourishing, or well-being.

⁶⁵ 482

⁶⁶ 483

Even on this richer picture, however, caring seems to come apart from personal commitment in two ways. First, caring doesn't seem to have the same normative upshot as committing to something. For both Frankfurt and Jaworska, caring is primarily a psychological state. For Frankfurt, caring about running a marathon, for example, would involve both being motivated to pursue that goal in various ways, and also to want to protect that motivation. For Jaworska, this caring would be more centrally defined by having the right sort of sustained affect toward that end. In both cases, however, these are primarily descriptive claims about my motivation. To be fair, caring is a psychological phenomenon that is *deep*, as well as weighty. It has structural complexity, and particularly on Jaworska's conception, what I care about affects the way in which I understand myself and likely also how well my life is going. Though I take it that personal commitments are also generally characterized by psychological depth, however, this isn't equivalent to *normative binding*.

Compare the case of caring about the environment to making a personal commitment to environmentalism. If we take our experience of personal commitment at face value, our commitments seem to give rise to reasons for action. Thus, if I fail to follow through on my personal commitment, I should hold myself responsible. I will be *required* to act on this commitment in various ways, and it will be appropriate for me to reprimand myself in the case of failure and perhaps to try to make some amends. If, by contrast, I claim to care (in a Frankfurtian sense) about the environment and yet never bother to recycle, or to conserve energy, then what I have done is show that I actually don't care, or at least that I don't care that much. I lack the appropriate desires to count

as caring, or they have been outweighed by something I care about more. Our response may be slightly more complicated if we turn to Jaworska's notion of caring since my emotions may not always line up quite as neatly with what motivates me. Nonetheless, the sorts of criticisms associated with this sort of caring also seem to diverge from those associated with personal commitment. We can, of course, say that it would be better if someone didn't have the emotional attachments he has, or criticize the weak will that prevents him from honoring those attachments in the right way. At some point, however, a failure to act on one's caring will be evidence that one lacks the appropriate emotions, at least in sufficient strength, and hence that one does not actually care. By contrast, again, simple failure to follow through may be evidence that one ought to revoke my personal commitment, but it alone will not show that I am not actually committed in the sense we are using. When we fail to follow through on our personal commitments, we seem to be doing something blameworthy, not to be demonstrating that we are not actually bound.

One response here would be to turn to yet a further conception of caring, one that can share the sort of normativity we have been associating with personal commitment. This, however, brings us to the second disanalogy between personal commitment and caring, which I suspect would hold even on a different conception of caring: whether I care does not seem to be under my direct control. Whereas we seem to make personal commitments through a volitional act, by willing it to be the case, actually caring about some end is not simply a matter of deciding to care. This point seems fairly clear for both Frankfurt and Jaworska. I can work to cultivate some desire or emotion indirectly:

if I want to care about recycling, for example, I might form certain precommitments that will trigger the desire to recycle, as well as the desire to keep desiring to recycle. At some point, I may even be able to cultivate the appropriate desires or emotions so that eventually I “really” care about recycling, as we might say. I cannot, however, directly will these desires or emotions into being. Even on a different conception of caring, a certain independence from the will seems like part of our intuitive notion of what it is to care. I may really want to care about some end, or person, without being able to bring myself to do so; conversely, I may find myself unable to stop caring when I wish that I would. Again, sometimes I can successfully work my way into caring through indirect means. Nonetheless, I do not seem able to make it the case that I care about something, simply by willing it to be the case.

Of course, it also true that some aspects of personal commitment may be out of my control: it will not, for example, always be in my power to meet the *demands* of my personal commitments. We sometimes lack the physical power to do what is needed, (a robber has tied Annie up so that she can’t go fundraising today), and sometimes we may not be able to bring ourselves to act, (overwhelming depression or anxiety make it impossible for Annie to get out of the house and go to the fundraiser). And of course, we may lack the appropriate emotions towards our personal commitments. As I suggested above, if Annie lacks the right sort of emotional attachment to her personal commitment, then we may want to call it a defective case of a commitment. When it comes down to making a personal commitment, however, that seems to be something that we do at will.

Narrative Integrity

A final way of understanding an agent as a whole, and her integrity as maintaining that wholeness, is in terms of narrative identity. Championed by Margaret Urban Walker and Hilde Lindemann, among others, narrative ethics draws an analogy between a human life and a story. The integrity of a story is a sort of wholeness, but it is one that centrally hangs upon questions of intelligibility and responsiveness to value, rather than simple consistency or internal integration. Similarly, the idea goes, we can assess the integrity of a human life in terms of its narrative arc, and whether the values, priorities, and goals that characterize that life hang together in the appropriate way.

These views reflect the idea of moral responsibility (in prospect or retrospect) as attaching to persons, a conception of a person as identified at least in part by a history... It is not only for moral purposes but also for purposes of intelligibility over time that we read and reread actions and other events backward and forward, weaving them into lives that are anything more than one damned thing after another.⁶⁷

To have narrative integrity is to act in ways that maintain an intelligible life story and to reject those intentions that fail to cohere with that story so far.

Let me return to an example. Once she has made her personal commitment, narrative integrity would demand either that Annie raise the money needed by the community theater, or that she be able to come up with a coherent story as to why she doesn't actually need to do so. Narrative integrity may also demand more than mere coherence: a stronger sense of narrative integrity may involve making choices that underwrite running themes in one's life (for example, if Annie's current personal

⁶⁷ Walker, 67

commitment is part of a larger role as a patron of the arts), that play into recognizable literary tropes (where Annie's personal commitment can be part of a sub-narrative in her life, such as an episode in which she helps to save the community theater from impending financial disaster), or that incorporate other narrative elements. On this understanding, Annie's personal commitment would give rise to demands as a way of maintaining its role in her life story, and figuring out just what is demanded by her commitment will require looking more concretely at how it fits into her narrative.

So far, we have been gesturing at the norms of narrative integrity with the word, 'intelligibility,' but let's pause now to take a closer look at just what intelligibility entails. In the first place, we have to be able (at least in principle) to make sense of our personal commitments to others. I can't be committed to reversing the flow of time, or committed to '3', because these are nonsensical commitments. This requirement also seems to make a personal commitment to counting blades of grass difficult, if not impossible, to maintain. Although it is not literal nonsense to say I am committed to such a thing, it is stretching the limits of intelligibility to say a sane human agent is genuinely committed to this end. Intelligibility requires that the content of my commitment be the sort of thing that my community can recognize as potentially valuable. Though this is not yet to say that my personal commitments need to be impersonally valuable, they should be the sort of thing to which others can imagine attaching a special value.

Maintaining a narrative, however, can require more than mere intelligibility – it also has a diachronic aspect, which requires honoring one's commitments across time. In order to maintain the thread of my narrative, I will have to take my extant commitments

into consideration as I deliberate. Moreover, unlike the psychological persistence characteristic of caring, appealing to intelligibility can also help to explain why the normative force of my personal commitment can come apart from my descriptive motivation. In order to maintain intelligibility when dropping a commitment, I will need to excuse myself, to show that this commitment shouldn't bind me in the way it once did. As Walker puts it, integrity entails reliability *ab initio*, in having and following through on our firm convictions, and also *post facto*, in being able to respond to what we have done. If Jason wants out of his personal commitment to vegetarianism, he will need to be able to show why that commitment should no longer have the same weight in his life that it once did, or else face up to the fact that he has failed by the lights of his initial commitment. In the case of something like an intention, the mere fact that Jason is no longer motivated may be enough to provide this explanation. In the case of something more weighty and significant, like a personal commitment, providing an appropriate story seems like it will require more.

On this proposal, our personal commitments would bind us because of the characteristic role they play in giving our lives diachronic consistency and intelligibility. And indeed, this sounds both like our intuitive description of personal commitment and like common threads we've seen so far in other proposals. In making a personal commitment to some end, the agent would be giving that end a specific sort of authority in her life story. Consider, again, Shannon's personal commitment to improving her temper. This personal commitment confers a particular sort of thematic importance on

developing patience, making it a leading “storyline” so to speak. In order for this end to be a candidate for personal commitment, narrative integrity will require some sort of back story that can make the act of commitment intelligible. Maintaining integrity will require that Shannon spend sufficient time and energy on her temper to bear out the importance she has given this commitment in her life’s story. It would be a significant breach of coherence, for example, for Shannon to make this personal commitment and then to wholeheartedly and gleefully curse the drivers around her during rush hour each day. And if Shannon wants to revoke her personal commitment, she will have to be able to make that choice coherent with her past act of commitment and her future choices.

This idea of narrative integrity picks up on themes we have already seen. Like Bratman’s metaphor of location, appealing to narrative integrity involves the idea of taking a stand on something. To make a personal commitment is quite literally to stake oneself on an end, since it is to take that end up into one’s life story. And like Jaworska’s notion of caring, narrative integrity attributes importance to those ends which are idiosyncratic, and yet also a deep source of meaning and concern in my life.

In addition, however, a narrative notion of integrity seems to come closer to the normative contours of personal commitment than either Bratman’s account of intention or Jaworska’s account of caring. Narrative integrity’s demand for ‘intelligibility’, as we have just seen, is already a normative rather than a merely psychological demand for consistency and coherence. For both Bratman and Jaworska, our commitments seem to be binding because we are in fact deeply motivated by them, or because abandoning them

would incur serious psychological costs. Rather than mere consistency, or fidelity to some true self, narrative integrity requires being able to present a coherent account of one's choices and actions. "[Integrity's] point is not for us to will one thing nor to be it, but to maintain—or reestablish—our reliability in matters involving important commitments and goods."⁶⁸ Reliability with respect to our personal commitments involves following through on their demands, or at least being ready to provide a good excuse when we do not.

Given the description so far, however, one might worry that this appeal to an intelligible narrative rules out genuinely *personal* commitments, generated by an agent's (in-principle private) volitional acts. Indeed, the notions of intelligibility and responsibility on which Walker relies are inherently social ones. As she says:

But these narratives, even if individuating, cannot be private or idiosyncratic. They serve purposes of *shared* understanding, not only of self-guidance but of justification and criticism. We are neither unfortunate enough to have to go it all alone in trying to find and keep an acceptable and vital moral order in our lives nor lucky enough to have the last word on whether we have succeeded.⁶⁹

Narrative integrity is in part a matter of being intelligible to, and reliable for others. It is not the case that I can simply stipulate, first-personally, what my narrative is, and how we ought to understand it. As Lindemann points out, the impression that I can just “make up” my story in this way leads to skepticism about any given narrative – why accept my version, rather than yours, or his, or hers? Lindemann responds, “...the answer to that question is, Wait – we're not finished yet. When the case has been retold often enough to

⁶⁸ Walker, 64

⁶⁹ 63

get a sense of who the participants are and what moral considerations ought to be brought to bear on them ... it is time to tell the story forward.”⁷⁰ In other words, getting a sense of what my narrative is – and hence what narrative integrity demands – may often require coherence with other takes on my story, in addition to the story’s internal coherence.

I don’t think, however, that this point rules out genuinely personal commitment. Though sometimes individual narratives will overlap and incorporate each other in a way that requires reaching consensus on just how the story goes, other times I really may have sole authority to “author” my narrative. Presumably, for example, how I conduct myself in private is relevant to my integrity, even if no one is actually around to observe my behavior, or to witness my moods. In these cases, this social intelligibility must be a counterfactual thing: if I had to explain myself to someone else, *could* I do so? To gesture again at Wittgenstein, a version of the private language argument seems to establish the necessarily social nature of language, but this doesn’t show (nor is it meant to show) that given my facility with this social system, I cannot now have private thoughts which I don’t share with anyone else. Similarly, narrative constraints on intelligibility are primarily social, but as part of a society with a substantive set of such constraints, I am able to apply them privately to my own life. Indeed, it seems that I *ought* to apply them to my own life, even when no one else is ‘watching,’ so to speak. After all, while narrative integrity is a matter of reliability, it is just as important that I can rely on myself as that others can rely on me. So, according to narrative integrity a

⁷⁰ 24

personal commitment binds me as a particularly significant element of a coherent story, one on which both others and I myself can rely.

Out of all of the understandings of integrity as “wholeness” that we have considered, narrative integrity seems to come the closest to the sort of normativity we described in Chapter 1, with personal commitments being genuinely binding and at least in principle, private. At the end of the day, however, I suspect that any theory of integrity will fail to recapture our sense of personal commitment fully for two reasons: first, integrity gives rise to essentially holistic demands, and second, integrity by itself seems unable to explain the characteristic contours of personal commitment.

First, that integrity is a holistic concept is a common thread running throughout all of the theories we have considered. That is to say, integrity demands that we abide by our commitments enough of the time, and to a sufficient degree: we may violate our commitments on a certain number of occasions, and still be said to have integrity. Shannon, for example, can be said to have integrity even if she occasionally violates her personal commitment to patience by giving in to a fit of pique. This is particularly true if she is open about her failures, and feels the appropriate regret for them. But while this sounds plausible with respect to integrity, it does not fit with the picture that we have been developing of personal commitments. Shannon’s personal commitment, rather, seems to demand that she work to curb her temper on each relevant occasion, not that she do so some minimum number of times. If she is truly committed to improving her temper, then merely pointing out that she was very patient last week doesn’t seem to

excuse her from trying to be patient today. Even if this failure is excused, Shannon will be to blame for failing in her commitment.

Of course, some personal commitments may impose only holistic or threshold constraints. A personal commitment to learn German will require just that I study “enough” to get the job done. This commitment, however, is holistic because of the nature of the end – “learning German” precisely means studying German for a sufficient amount of time to achieve some level of fluency – not because of the nature of personal commitment. Furthermore, that personal commitments are normative in each (relevant) instance, rather than imposing holistic constraints, is obviously compatible with recognizing mitigating circumstances and excuses for violating our commitments on occasion. There will clearly be times when the demands of our personal commitments are outweighed by other, more pressing demands: Annie obviously may (and perhaps really must) abandon her fundraising efforts if a chronically ill family member needs her help. Even when one is not excused from one’s commitment, we may recognize mitigating circumstances: for instance, we may be more inclined to forgive Shannon’s occasional bad temper when she’s having a rough day. Whatever sort of reactive attitude is appropriate, however, the point is that her personal commitment seems to have normative force each time Shannon’s temper is at issue, albeit a force that can be mitigated or outweighed. By contrast, the requirements imposed by integrity often don’t “distribute,” so to speak, to each relevant instance.

Secondly, appealing to any theory of integrity simply lumps personal commitments in with all of the other material that is relevant to our integrity. For

Jaworska and Frankfurt, for example, personal commitments are significant simply *qua* member of the class of our carings; for Walker, they would have be just another element in our narrative. If we want to retain our initial picture of personal commitment as a somewhat distinctive phenomenon, however, then we need a story as to what distinguishes it from other things (whether those are carings, desires, ground projects, or sub-plots) that figure into integrity. Why do personal commitments, unlike mere desires, seem to give rise to structurally complex reasons for action? Why does the normative force of personal commitment, unlike some other elements of our life stories, characteristically weigh in on each relevant circumstance and not holistically? To the extent that we find the phenomenology described in Chapter 1 persuasive, integrity cannot be the whole story. We need an account of personal commitment as such. In the next chapter, we will see if we can arrive at such an account by analyzing personal commitments as promises to the self.

Personal Commitments as Promises to Oneself

Thus far, I've tried to make it plausible that personal commitments are more robustly normative than intentions (in that they give rise to reasons for action), and that personal commitments are not just a way of re-describing the demands of a virtue like integrity (in that their reasons are not holistic, but rather "distribute" to each relevant choice situation). A third alternative we might consider is whether we could analyze personal commitments as a promise to oneself.

There are at least some initial similarities between promises and the picture we have been drawing of personal commitments. Like promises, personal commitments seem to give rise to new reasons for action: they can change what it is I actually ought to do. Whether Carrie promises her mother to get her undergraduate degree, or makes a personal commitment to do so, it seems that Carrie now has a new reason to finish college. And promises give rise to the same sort of structurally complex reasons which personal commitments seem to generate, what Raz calls "protected reasons", a combination of first-order reasons to act on one's commitment, and second-order reasons not to act on competing considerations. When I promise to meet you for lunch, I now have a positive reason to show up for our lunch date, and a second-order reason not to act on a wide range of competing considerations, such as a preference to stay home and eat alone. Because of these second-order exclusionary reasons, a promise turns an act that is antecedently optional (although good to do) into a matter of requirement. When I make a promise to meet you for lunch, it is not just the case that meeting you for lunch has

become more valuable, or more choiceworthy than it would have been beforehand. Rather, I now have an obligation to meet you for lunch; I am required to do so, (though that requirement can be outweighed by a heftier moral duty, or dissolved with your consent).

Like the demands of promises, we have noted, a personal commitment to spend more time with friends seems to make it the case that I am *required* to spend more time with friends, not that spending more time with my friends has simply become more valuable, and hence a better option. There are many sorts of considerations that can make an action more choiceworthy: the fact that some of my friends are philosophers, and so can help me work through some thorny issues in this dissertation, could make spending time with them more valuable than if it were just a matter of socializing. But if this is only a matter of added value, I can nonetheless permissibly decide that I prefer to stay home and relax. I may decide to forego the benefit of their insight in order to pursue some other good. If I am truly *committed* to spending time with my friends, however, then it seems that there will be at least some times when I cannot permissibly engage in some other activity. Like promises, my personal commitments seem to generate *requirements*.

Promises and personal commitments also seem alike in that they each give rise to these reasons through an act of *commitment*. In part, this is to say that both promises and personal commitments result from the exercise of a normative power, a power to change one's reasons for action through normative means. Unlike cases of pre-commitment, promises create normative change directly, not by piggybacking on causal changes.

Similarly, the ability to make personal commitments seems to be a normative power: if I can make a personal commitment, then I can privately bind myself simply by willing it to be the case. But in addition, claiming that promises and personal commitments are both acts of commitment is to call them normative powers of a certain kind. They are normative powers that essentially involve an agent binding his own future self – committing to do some future action, eliminating other future actions as valid choices, and at least in some way staking his character on whether he follows through. So, for example, personal commitment and promising are to be contrasted with normative powers like the authority to issue commands, or to perform marriages, which license putting *other* people under obligation. Though this language is rather metaphorical, both promising and personal commitments seem to be ways of “putting oneself on the line”.

Now promises, of course, are standardly between two or more people, whereas we have stipulated that personal commitments, if they do exist, involve only one person and indeed can be made privately. Notice, though, that we do sometimes talk in terms of promises to ourselves. It is perfectly natural to say, for example, “I promised myself that I would clean my apartment this weekend,” or “She promised herself that she would support the community theater.” Furthermore, we sometimes talk and act as if these self-promises are reason-giving. Compare this to the case of intention. If Annie tells me that she is planning to raise money for the community theater, and yet she never does, I may genuinely criticize her for weak will, or for backing out of what would have been a generous act. More, however, seems to be implied by the language of a “self-promise”. If Annie tells me that she has *promised herself* to raise the money in question, and then

never does so, I seem able to criticize her in a new way. As in the case of a mere intention, it will still be appropriate to raise the worry of weak will, but something more seems to be at stake: I may also think she is exhibiting a sort of bad faith, or a lack of respect for her promise.⁷¹ Having made a promise to herself seems to provide an extra “glue” in binding Annie to this end, in a way that mere intention would not.⁷²

I think these initial similarities are illuminating: looking at the way in which personal commitments are like a promise highlights their dependence on the committor’s will, and the sort of normative upshot they produce. Nonetheless, I think that analyzing personal commitments as promises to oneself (or as I will sometimes say, “self-promises”) is a mistake. At the first order level, assimilating personal commitments to promising fails to mesh with our pre-theoretical experience of the ways in which our personal commitments bind us, in terms of how we make each, break each, revoke each, and are guided by each. In addition, most theories of what grounds the normative force of promises don’t extend well to the case of personal commitments: they either make personal commitments look mysterious, or else as if they can’t possibly have normative force. To get a clearer picture of these differences, let’s take each in turn, starting with the first-order picture.

⁷¹ Of course, a pattern of going back and forth on one’s intentions may raise questions about the firmness of the agent’s character, in addition to worries about our ability to rely on her. If Annie regularly waffles as to what she is going to do, we will be less likely to trust her, or even just to take her seriously. But, as we saw in the last chapter, the force of personal commitments doesn’t seem to be holistic in this way.

⁷² Again, as I noted in the introduction, our intuitions are likely to be complicated when the end to which someone commits herself is not clearly valuable on its own, in the way that Annie’s commitment to the local theater is. Nonetheless, I’d like to focus here on the more paradigmatic cases of personal commitment, which I take to be to worthwhile ends, and return to the question of less valuable (or completely valueless) personal commitments in the final chapter.

First-Order Normative Contours

Though there are many different theories of promising, (some of which we will consider in the next section), most of these agree on what promises look like at the first-order, disagreeing only as to what grounds this first-order normativity. The standard first-order picture of promising has broadly three components: how one makes a promise, the kind of normative force exerted by a successful promise, and the ways in which a promise can be dissolved or outweighed. On this common story, a promise is made when the promisor (the person making the promise) attempts to obligate himself to the promisee (the recipient of the promise) to do some act(s) *phi*. In other words, the promisor hands the promisee a particular sort of right over this bit of his future behavior. For the promise to be successful (such that the promisor is now in fact bound to the promisee), there must be what J.L Austin calls “uptake”: the promisee must recognize that the promisor is trying to bind himself, and accept (whether explicitly or tacitly) the right that the promisor is offering her.⁷³ Once a promise has been successfully made, it issues a requirement on the agent, one which exerts moral force: unless a promise is outweighed by a stronger moral obligation or otherwise dissolved, to break a promise is to commit a moral wrong. One’s promises continue to bind one unless one either discharges the promise by giving the promised good to the promisee, or else the promisee consents to dissolve the promise in question.

⁷³ *How to Do Things with Words*, 117

To begin with the first of these distinctions, we seem to create personal commitments and promises in different ways. Promises are generally created by employing a conventional act: simply saying “I promise to *phi*” to another person is enough to obligate oneself to *phi* (assuming that one is clearly serious, and that there is uptake on the part of the promisee). Whether or not one sincerely (in the privacy of one’s heart) intends to follow through on the promise is irrelevant – one is bound just so long as the promisee reasonably takes one to be making the promise, and accepts it.

In the case of personal commitment, there is clearly not an analogous formula that we use privately to commit ourselves. When I make a personal commitment, I may literally say “I promise myself that I will...”, or I might say “From now on, I’m really going to...”, or employ a number of other phrases, or perhaps even come to a moment of commitment in a way that doesn’t involve propositional thought. Given that personal commitments are at least in principle private, there certainly doesn’t need to be an outward act, and it feels overly legalistic (and failing to conform to our actual experiences) to insist on some formally defined inner act.

Of course, even in the case of promises to others, we needn’t always employ the conventional formula; I can bind myself simply by deliberately implying that I am making a promise. If I inform Dan of my intention to babysit for him this weekend, adding “you can count on me!”, that counts as a promise. Nonetheless, this interaction constitutes an implicit promise because a normative transfer of the kind we have already mentioned has taken place: I have to hand Dan a claim, such that he has the right to me babysitting for him this weekend, and can demand performance. This sort of transfer,

however, seems odd in the intrapersonal case – however our personal commitments bind us, it does not seem to be through a transfer of rights. To compare, I *could* hand myself an IOU for ten dollars, giving myself a right to be paid ten dollars by myself, but it is unclear why I would want to do so, or what sort of normative effect that would have. Unlike the case in which I hand you an IOU, which changes a normative property (however much money I have, ten dollars of it ought to be yours), an IOU to myself doesn't affect who owns that money, or who ought to have it. This point requires more elaboration, but we will say more – addressing whether this sort of intrapersonal transfer of rights is possible, and even if it is, whether that's what's at work in personal commitment – as we discuss different grounding theories of promising in the next section.

Next, let's compare the way in which promises and personal commitments can, respectively, be dissolved or outweighed. Let's start with a few similarities. First, it is obvious that both promises and personal commitments can be outweighed by a more significant demand: my promise to meet you for lunch will be outweighed if I pass a traffic accident on the way to meet you and am able to provide first aid to the victims; likewise, Jason's personal commitment to vegetarianism will be outweighed if it threatens his health by causing him to develop anemia. Under those circumstances, I ought to skip lunch and stay to help the accident victims, and Jason ought to go back to eating at least some meat.

And in both cases, when one of our promises or personal commitments is outweighed, that we don't have to act on the commitment in question is not to say that the promise or commitment no longer exists at all. Though its typical force is currently being silenced, the promise or personal commitment is still "lurking" in the background. Most obviously, if the outweighing force disappears for whatever reason, then the original requirements will re-exert themselves. If it turns out that no one was injured in the accident, I should try to make our lunch date, and if Jason's anemia recedes, he should go back to being a vegetarian. But even if their demands continue to be outweighed, a promise or personal commitment may continue to express itself in the form of some normative "residue".

The idea behind normative residue (generally referred to more narrowly as 'moral residue') is that even when one rightly ignores some obligation, that obligation may continue to exert *some* sort of force; one may be open to a sort of criticism, or required to do (or feel) something else in order to make up for ignoring that obligation (even though one was justified in doing so).⁷⁴ Outweighed promises generally leave behind some normative residue – for example, even though I am fully justified in skipping lunch to help with first aid, I at least ought to call you afterward to explain what happened, and perhaps also ought to meet you for lunch another time. Given the scenario of a serious accident, this doesn't just mean that I may permissibly skip lunch, arguably it means that I *must* do so. Even so, my promise to meet you for lunch means that I owe you something to make up for missing our date.

⁷⁴ See, e.g., Williams (1965).

I suspect that something similar is true in the case of personal commitment. If our personal commitments really are genuinely binding, then even when they aren't the strongest consideration on the table it seems that their normative force should show up in some way or other. Even though Jason's health feels like a sufficient reason to ignore his personal commitment – indeed, if the threat is serious it may even require that he do so – the force of his personal commitment ought to show up in *some* way or another if he is truly committed to vegetarianism. Perhaps Jason ought to find more humane sources of meat, or at least feel regret that he can't be a full vegetarian. The general claim is that even when the demands of our personal commitments are outweighed, so long as we remain committed, we may be required to show respect for those commitments in alternative ways.

As opposed to being outweighed, other times a promise may be revoked. In this sort of case, a genuinely binding promise is dissolved, so that no normative force remains.⁷⁵ (We can call the conditions for successfully revoking a normative commitment “exit constraints”.) One of the defining characteristics of promising is that this revocation can be achieved only by the promisee agreeing to waive her rights. Familiarly, if John promises to go to an art exhibit with me on Saturday, only to realize when Saturday comes around that he is really no longer in the mood to spend the day in a

⁷⁵ To be clear, the promise *itself* exerts no more normative force. A revoked promise may, however, continue to leave behind some normative trace through its connection to virtues like generosity or fidelity. For example, even if a friend genuinely releases me from a promise to help her move to a new home, generosity may weigh heavily in favor of me insisting that I help anyhow, or at least finding some other way to give her a hand. Similarly, Jason's wife releasing him from a promise to be a vegetarian means that he is no longer bound by that promise, but fidelity to someone he cares about may demand that he find some other way to support and promote his wife's values. Thanks to Jeremy Snyder for conversation on this point.

gallery, I can release John from his obligation just by saying “never mind”. Indeed, I can release John from his promise for any reason at all, or even, as we sometimes say, for “no reason.”⁷⁶ Unlike being outweighed, when the promise continues to exist in the background, or a misfire or failed attempt, when the promise was never really binding to begin with, “revocation” means that there was a genuine obligation that simply no longer exists.

We also seem to be able to revoke our personal commitments, so that they are no longer binding on us. If Shannon comes to feel that a quick temper is really not that big of a deal, and that she ought instead to be focusing on other priorities, she may decide to revoke her personal commitment. She can “take it back”, as we say. There may be some lingering normative effects – having revoked one personal commitment, for example, perhaps she needs to be more careful not to let revocation become a habit, in order to maintain her integrity or to avoid laziness – but the personal commitment itself no longer seems to be in effect. Indeed, it would verge on the ridiculous to say that having once made a personal commitment to developing patience means that Shannon is stuck with this obligation for the rest of her life.

It does not seem, however, that personal commitments can be revoked in just the same way as promises. In particular, unlike the case in which someone else has made a promise to me, I do not seem able to let myself totally off the hook simply because I decide to take back the commitment. While Jason’s wife can release him from his

⁷⁶ While we occasionally say we have acted “for no reason,” what we generally mean is “for no good reason,” or that our reasons are trivial or uninteresting. A requirement on rational agency generally is that one’s actions be intelligible, and hence not literally for “no reason.” Additionally, we may think that an agent is foolish (or in some other way criticizable) for releasing a promisor in some circumstances, but the point here is that she nonetheless has the authority to void my obligation at will.

promise to be a vegetarian for any reason or “no reason at all”, more seems to be required in the case of a personal commitment. If an agent could genuinely release himself from a “personal commitment” at will, for no good reason and with no remainder, then this “commitment” would not seem to be robustly binding. After all, he could simply dissolve his obligation any time he didn’t want to do what he had made a “commitment” to doing. If Jason could revoke and reinstate his personal commitment to vegetarianism at will and with no normative consequences – for example, revoking his commitment just long enough to enjoy one evening out at the steak house – then his personal commitment hardly seems binding.⁷⁷ As we have noted, this very thin sort of normativity does characterize bare intention, but it is not able to recapture our intuitive picture of personal commitment.

Rather, I think that fully revoking a personal commitment has slightly more complicated conditions than releasing someone else from his or her promise.⁷⁸ Now, there is a sense in which we can revoke our personal commitments at will. If Jason genuinely and deliberately decides that he no longer wants to be bound by a personal commitment to vegetarianism, it seems either to be a sort of stubbornness to insist that he is still committed, or else fetishism about the binding power of our commitments. Jason can release himself. Nonetheless, even though Jason can release himself at will, there is still a difference between revoking a personal commitment, and breaking that personal

⁷⁷ One might suggest here that releasing myself will not be so easy if we take a more complicated view of the self – for example, if a personal commitment is really some sort of contract between my present self and some future self. I address this sort of suggestion below, in considering the kind of normative force that characterizes personal commitment.

⁷⁸ Indeed, I think exiting a personal commitment is even more complicated than we will see here, but for the moment I only want to focus on distinguishing personal commitment from promising. We will return to the question of exit conditions in the next and final chapter.

commitment.⁷⁹ Unlike the case of promising, in which the promisor can switch the toggle back from “bound” to “free”, getting out of a personal commitment is a more complicated and piecemeal matter.

To come back to the example of Jason, though he has the power to revoke his personal commitment to vegetarianism, doing so out of mere laziness or caprice seems to leave him open to criticism – abandoning his commitment out of pique to annoy a sanctimonious friend fails to reflect the seriousness of having committed himself. And in failing to respect the weight of his commitment, Jason’s revocation is incomplete; it leaves behind normative residue. He will be blameworthy for jettisoning his commitment in this manner, and perhaps owe some amends, whether by working towards the end to which he committed in some other way or being more vigilant in the rest of his commitments. Being *fully* released from his commitment requires that Jason put an appropriate amount of thought and care into its revocation. He needs sufficiently good reason to shed his commitment completely, and to come to that decision in a sufficiently thoughtful way. Again, this does not mean that we should say that we can never fully release ourselves from our personal commitments – this would be far too strong. Our experiences of our own personal commitments do not suggest that they are set in stone. For some commitments, time limits are inherent in their content, like if Carin makes a personal commitment to run a marathon. But while others, like Shannon’s commitment to increasing her patience, will be intrinsically more open-ended, we still seem able to

⁷⁹ Tom Hill argues that we can recapture the distinction between breaking and revoking a self-promise by stipulating that a self-promise can be revoked only for good reasons that are discovered in the absence of immediate temptation. The need to stipulate these conditions, however, only supports our point here, which is that personal commitments and promises are characterized by different first-order contours.

revise and even revoke our personal commitments with sufficient reflection.

Appropriately and definitively exiting a personal commitment requires respecting the significance of an act of commitment, in a way that will generally be more complicated than a promisee simply waiving her rights over the promisor.⁸⁰

One might worry that I have too quickly rejected the possibility that personal commitments and promises share their exit constraints – that perhaps, in both cases, waiving the commitment is all that is necessary and any remaining normative residue is really just a question of character in either case. As I will argue in the next chapter, I don't think that this is so. In any case, however, my goal here is to highlight the way in which the exit constraints are different in the case of promising and personal commitment. Though the commitment can be dissolved by an act of the will in either case, *whose* volitional act differs (i.e. the committor cannot release herself in the case of a promise), and whether the volitional act is grounded in good reasons matters only in the case of a personal commitment (a promisee may release a promisor for better or worse reasons, but these generally won't reflect on the promisor's obligations).

Finally, the normative force of personal commitments seems to be different in *kind*⁸¹ from that of promises. On the one hand, personal commitments can feel like

⁸⁰ Of course, I will go into more detail about the conditions for exiting a personal commitment in the next chapter.

⁸¹ I'm leaving what I mean by a "kind" of normativity somewhat vague here. In the first instance, I think there is a widely shared intuitive sense of what I mean – given a set of norms, we could often pick out which should be lumped with the law, which with morality, which with etiquette, and so on. But to say a bit more, I think a sort of normativity is distinguished both by what grounds its norms (the rulings of a legitimate government ground legal norms, while social mores ground the norms of etiquette, etc.), and also by a characteristic normative profile. (moral norms tends to be quite heavy and will sometimes require an

promises, in that they seem genuinely to bind us, to put us under requirement. As we said in the last chapter, personal commitments can feel binding in various ways, such as occasioning satisfaction and pride when we are doing well, and frustration, stress, and shame when we are not. For example, Annie will likely feel that she is doing something wrong if she shirks her personal commitment to raise money for the community theater, and if a third party is aware that she has made this commitment, that party may also judge Annie negatively for failing to live up to it. And when tempted to break her personal commitment, Annie may feel its normative force pulling her back toward following through. On the other hand, however, we can also recognize that we are bound by our personal commitments only because we have willed them into being. Even as she feels bound by her personal commitment, Annie can step back and recognize that she brought this obligation upon herself, and unless she has a particularly (perhaps pathologically) strong super-ego, Annie should also recognize that she has the power to revoke the commitment in question (under the appropriate circumstances). At least as we have been describing personal commitment so far, Annie is at the same time bound by her personal commitment and free to dissolve this bond.

To borrow language from Ruth Chang, personal commitments seem to partake in a dual nature.⁸² The same personal commitment (sometimes at the very same time) can feel both like an object of duty, and like a contingently adopted end that one can chose to

action while other times providing more justification for doing something, legal norms on the other hand come in a variety of weights but are almost entirely requirements and so on). In the final chapter, I'll say a bit more about what I take to be the distinctive "normative profile" of personal commitment, though I take it more could be said.

⁸² I borrow this description from an analogy Ruth Chang draws between personal reasons and stipulative definitions. In both cases, Chang claims, the norms we have created present themselves in this "Janus-faced" manner (34). Personal commitments seem to fall into the same category.

revoke (albeit under appropriate conditions, or else the threat of normative remainder). Particularly when faced with the choice to abide by one's personal commitment or not, the feelings of constraint may dominate one's perception of the commitment. Jason's feeling of temptation in the presence of a juicy steak will only highlight the normative force of his personal commitment. Nonetheless, it is always in principle possible to step back and notice the fact that the continued normative force of this personal commitment is contingent on his will. As we have said before, personal commitments are not irrevocable: one has the power to revoke one's commitments.

Compare this to the case of promising. Though my promises are created through my activity, once created, there is a sense in which they "belong" to my promisee.⁸³ More concretely, making a promise gives the promisee a claim over me. As long as I am a conscientious moral agent, I should recognize that having promised, this obligation is outside of my immediate control: my promisee has a right to what has been promised, and I can be released only at her consent, or excused by a more pressing moral obligation. If John promises to meet me at the art gallery, for example, this promise is more like one of his standing moral duties than his freely chosen plans and goals. (Indeed, though they are not standing obligations, promises are typically taken to have moral weight. We will return to this point in considering a Kantian theory of promising.) It is my promisee, not I, who has the power to dissolve this normative force. My personal commitments, by contrast, are in the funny position of being genuinely under my control (in the sense that I have the power to revoke and revise them) at the same time that they are genuinely

⁸³ This intuition is highlighted by transactional theories of promising, to which we will return below.

binding – they remain, in some sense, more a “part of me”. Whereas part of what characterizes promising is that transfer of a claim to someone else, it is precisely part of the puzzling nature of personal commitments that they can be, at the same time, a matter of obligation and an object of choice. The next question is how to make sense of the way in which this sort of act can have this sort of normative consequence.

Grounding Theories of Promising

Personal commitments and promises already seem quite different in how they operate at the first-order level. The disanalogies, however, become even more marked when we turn to the question of what *grounds* the normative force of personal commitments. As I have said, there are many different theories of promising, so for our purposes I’d like to pick out three accounts, and compare promises to personal commitments on each. Here we’ll consider Kantian, “raised expectation”, and transactional theories of promising.⁸⁴

A Kantian Theory of Promising

Let’s start by considering a broadly Kantian theory of promising. On the Formula of Humanity, our promises presumably bind us in virtue of a duty to respect humanity.⁸⁵ Roughly, once I have given my word to another member of humanity, I fail to treat that

⁸⁴ I should note that my division of the sorts of positions one might adopt on promising here does not closely track the way the terrain is generally divided in the philosophical literature. But since we are more interested in using promises as a tool to get clearer about personal commitments, rather than joining the conversations taking place in either moral or legal philosophy, I don’t take that as a worry for us.

⁸⁵ Arguably the Formula of the Universal Law is a more obvious choice for discussing something as conventional as promising, but I think that the Formula of Humanity can do a better job of pulling out the similarities (and of course the differences!) between promising and personal commitment.

member with respect if I break my promise. If, for example, I break my promise to meet you for lunch because I no longer feel like going, I'm treating my laziness as more significant than my commitment to you – I am treating you as a mere means to the satisfaction of my laziness. Showing respect for you demands that I at least get your consent if I want to do something else. And even if I just break my promise because I forget our agreement, my failure to find ways to keep my promise in mind may also fail to treat you as an end in yourself. Furthermore, appealing to respect for humanity can explain the normative residue that results when a promise is outweighed: if I have to break my promise because it conflicts with a more important moral obligation, then I will need some alternate way to demonstrate that I really do respect you. For a Kantian, the demand to show some alternative form of respect is the basis for the specific ways in which we are bound by our promises.

And in some ways, this theory of promising extends well to the case of a self-promise. As Kant himself repeatedly emphasizes, I too am a member of humanity, and so I owe myself the same respect that I owe to other agents. So if, from a Kantian perspective, I have a duty to keep my promises as a matter of respect to the promisee, then this duty will apply equally well whether my promise is to myself, or to another.⁸⁶ If it is disrespectful to abandon a promise to someone else, it is equally disrespectful to flout a promise to myself (again, without sufficiently good reason). Just as skipping a

⁸⁶ This may be a little imprecise. After all, Kant argues that I don't generally have a duty to promote my own happiness (though I do have an imperfect duty to promote the happiness of others) because I am naturally inclined to work towards my own happiness (MM 386). Similarly, I may be pre-disposed to fulfill self-promises more often than promises to others, particularly if I have promised myself to do something that furthers my own interest. Without diving too deeply into exegesis, however, I think we can see the application of the point about self-respect here.

promised lunch with you in favor of taking a nap is wrongfully subordinating you and your ends to my physical desire for sleep, so Jason cheating on his vegetarianism would be subordinating himself and his reflectively adopted ends to a fleeting desire for steak. And in the case where a promise to myself is outweighed, appealing to self-respect can show why there is some normative residue, some demand to show the breaking of the promise is not actually due to a lack of respect for myself. For example, though serious health problems may outweigh his personal commitment, maintaining self-respect will require Jason to find some other way to promote animal welfare (or whatever the basis of his commitment may be).

Indeed, I think the Kantian approach recaptures some of our basic intuitions about the importance of personal commitment: someone who is lackadaisical about her personal commitments does seem to express a sort of disrespect for her own will. Treating ends that one has to which one has given a special significance as easily ignored or disposed seems to disrespect one's autonomy, understood as the capacity to make decisions that guide one's own life. As Seana Shiffrin puts the point regarding interpersonal promises:

...The power to make binding promises, as well as a variety of other forms of commitment, is an integral part of the ability to engage in special relationships in a morally good way, under conditions of equal respect... Agents could not enjoy a meaningful form of autonomy, in relations with others, without the power to promise, a power that includes the ability to form a range of commitments that include but are not limited to the explicit, formal promise.⁸⁷

Similarly, one might claim, agents need to enjoy a meaningful form of autonomy in relation to themselves, which may include the ability not just to determine their own

⁸⁷ 4

action through easily revocable intentions, but perhaps also to bind their future selves more robustly through personal commitments.

But while I think something like this claim is right about personal commitments – namely, that the power to make these commitments is an integral part of what it is to be an autonomous agent – there is a way in which simply extending this observation to personal commitment is too strong. Namely personal commitments, unlike promises, do not typically seem to have moral force. Imagine Carin makes a personal commitment to running a marathon. She seems able to genuinely bind herself by making this personal commitment. It is now the case that Carin really *ought* to do her daily training, that she really *ought* to research what sorts of training and diet regimens can help her successfully complete this goal; that she really *ought* to register and run in a marathon once she is able to run the full distance. Indeed, Carin will be blameworthy if she doesn't take these necessary steps. It will be appropriate for her to feel that she is failing in some way. If Carin has told her friends about her commitment, it may be appropriate for them to hold her responsible for sticking to it, or to refuse to be complicit in any failures (for example, by declining to going out drinking with Carin the night before her marathon). These “oughts”, however, are simply not moral “oughts”. Carin is failing in some way if she doesn't train regularly for her marathon, but she is not being a morally bad person; her moral character is not implicated. A pattern of failing on her personal commitments may be indirectly immoral, in implicating her self-respect (because she doesn't care enough about her own reflective commitments) or her integrity (because she fails to follow through on what she has claimed really matters). Failing to follow through on this one

personal commitment may indeed contribute to an erosion in these virtues, but as I claimed in discussing integrity on its own terms, this cannot explain the sort of normativity that *this* particular commitment seems to have.

Of course, some personal commitments may be morally significant. If I make a personal commitment to a morally weighty end or in a way that intersects with a moral virtue, (both of which could be true in the case of a commitment to donating to healthcare in developing nations), then my personal commitment will also happen to have moral weight. And to go back to an earlier example, if Annie tells the community theater that she has made a personal commitment to raise the money they need, and they stop looking for other donors and so come to rely on that money, it would be morally wrong of her not to follow through. The moral force, however, comes from the fact that Annie would be knowingly harming the community theater, not from the fact that she made a personal commitment. What makes something a personal commitment, if our current picture is accurate, is the private sort of normative power that gets it off the ground, and we have seen many examples of personal commitment which just don't seem morally relevant. When personal commitments do have moral weight, it is indirectly in virtue of that particular commitment having moral content, or being a part of a larger pattern of questionable behavior, or somehow being connected to the rights or interests of others. The mere fact that one has made a personal commitment is not, by itself, doing the moral work.⁸⁸

⁸⁸ Indeed, Hill recognizes this fact – that the vast majority of what we might call “promises to oneself” are extra-moral – and therefore concludes that only those self-promises that have antecedently moral content are binding. As I suggest here, however, I think we can agree that only a limited set of personal commitments are *morally* relevant, while still attributing some sort of normativity to the whole class.

One might, however, worry that I have dismissed the idea that personal commitments are moral too quickly. Here again we can turn to Habib, who points out that apparent differences in normative force are sometimes the result of overlapping factors complicating our intuitions. For example, even if we stipulate that I have the same duty of respect to each member of humanity we may be less likely to notice, let alone blame, my failures in respect towards a mass murderer, or even just a universally hated neighbor. So while it may be the case, as I just suggested, that personal commitments don't have the same weighty significance as promises because they are not typically moral, it might just be the case that the private nature of personal commitment is complicating our intuitions. Perhaps personal commitments are actually just as morally demanding as promises, but we are more inclined to forgive non-performance because the commitment is private and hence the stakes are generally pretty low. After all, we sometimes see similar reactions in the case of promising, as when we are more inclined to let people "off the hook" for breaking relatively insignificant promises (like a promise to pick up some soda at the grocery store), even though technically this promise was genuinely binding unless and until the promisee said "never mind".

I think Habib raises a really important point here, and in fact I suspect that it is just these sorts of complicating factors surrounding their privacy that sometimes lead people to deny that personal commitments are normative *at all*. Being private does often mean lower practical stakes, and as Habib points out, we often have complicated feelings about the way agents treat themselves (particularly if that treatment can be construed as

self-rewarding). Nonetheless, I don't think that appealing to morality gives us the right account of personal commitment as such. In addition to the fact that calling them moral can overstate the significance of a good deal of our personal commitments, it also seems to trivialize the importance of the general claim that an obligation is a moral one.⁸⁹

This claim is complicated by the fact that there is clear and deep disagreement within philosophy on just what it means to call a consideration a "moral" one, but whether we gloss that claim as involving a flourishing human life, or the respect that agents owe one another, or the basic value of all sentient life, there is fairly widespread agreement that to call a consideration moral is to say that it is deeply important. It is, at least, generally understood as saying that the consideration has an extremely weighty, significant, universal sort of normativity. On the vast majority of views, morality is a very special, weighty kind of normativity, and there is good reason to reserve using this category for a limited class of reasons: to say that a consideration is moral is a very strong claim. While some of our personal commitments can be very weighty, and among those very weighty commitments, some will be moral, our examples so far have spanned a broader range in terms of their weight and significance. If we want to recapture our original picture of personal commitment, it will have to be with a normative force that is not a moral one.

⁸⁹ I don't intend to take on the question here of whether moral demands are always lexically prior to the demands of our personal commitments. (Indeed, the likelihood of this sort of ordering strikes me as implausible.) In any case, however, my claim here is simply that the sort of normative force exerted by morality and personal commitments, respective, seem to be different *in kind*.

Now, given my continued insistence that personal commitments are normatively binding, my current claim that they are not moral may sound odd. After all, on one common picture, there are really only two kinds of practical normativity: instrumental norms, and moral norms. If we accept this sort of picture, then an end can affect what I ought to do only via the hypothetical imperative (which means that I should take the means necessary to the ends I happen to have, but not that I am bound to maintain any of those goals) or else as a morally required end. As I argued in Chapter 2, instrumental rationality alone is too thin to recapture the sort of robust normativity that we intuitively want to attribute to personal commitment. But if instrumental rationality cannot underwrite our initial picture, then we may seem forced into either acknowledging personal commitments as moral, or else claiming that they are not normative in the way we thought they were.

Despite their lack of moral force, I want to hold on to the claim that personal commitments are genuinely normative, and in a way that goes beyond mere intention by creating new reasons for action. I suggest that we consider a third option: personal commitments are characterized by a distinctive sort of normativity – that is to say, while commitment makes use of the familiar deontic concepts of “ought”, “may”, “justifies”, and so on, I take personal commitment to be characterized by a distinctive *feel*, and by certain characteristic combinations of those deontic concepts.⁹⁰ But without going into the substantive nature of this normativity in this chapter, let me just try here to make room for a commitment-based normativity. To start, notice that many of us already at

⁹⁰ Here, ongoing work on the variety of deontic concepts by Margaret Olivia Little and Coleen Macnamara would be illuminating.

least seem to recognize a third kind of practical normativity, namely, prudence. In many (perhaps most) cases, the demands of prudence are not clearly moral: it is not really a moral wrong if I accidentally touch a hot pan and give myself a blister. But intuitively, we can say something about why I ought to avoid touching hot pans, and something which goes beyond merely referencing what I happen contingently to desire. Even if I genuinely just don't care about the pain of burning my fingers, (and, let's stipulate, I'm not doing permanent nerve damage or anything that serious), there still seems to be some sense in which I ought not touch a hot pan. We can at the least recognize this in our reaction to the person who doesn't care about painfully burning her fingers – there seems to be something wrong with her. Whatever this person happens to care about, she seems to violating a norm of prudence, a norm which is simply different in character from either morality or instrumental rationality.

For those of us who are willing to recognize prudence as a distinct normative kind, I suggest we at least consider taking on some sort of commitment-normativity as well. After all, just as we seem to recognize the demands of prudence as genuinely binding upon us, and yet not belong either to morality or instrumental reasoning, so we have seemed to recognize genuinely binding personal commitments. We feel their force guiding us when we deliberate, weighing on us when we fail them, and happily met when we obey them. The simplest explanation of this phenomenon, it may turn out, involves a distinctive kind of normativity, more robust than instrumental norms, less weighty than moral norms, and more agent-dependent than prudence.

Raised Expectation Theories

A Kantian theory of promising doesn't seem to recapture the normative force that characterizes personal commitment. Instead, let's consider a second possibility – a raised expectation theory of promising (sometimes developed as a rule utilitarian account). On this picture, promising generates obligations because we have a moral duty to satisfy any reasonable expectations that we have raised in others. This general account permits for different explanations as to why we ought to satisfy reasonable expectations. Adam Smith, for example, appeals to a natural disapproval triggered by failing those we lead to rely on us, whereas John Stuart Mill appeals to the utility of requiring fulfillment of those expectations. But whatever the underlying moral theory, raised expectation accounts share the claim that the obligation to keep my promise piggybacks on a basic moral duty to meet the reasonable expectations that I have raised in others. Once I have caused someone else to rely on my action by giving rise to expectations as to what I will do, I would be in some sense harming the promisee if I failed to follow through. In other words, in causing a promisee to rely on me, I have given him or her the right to expect the promised good.

Raised expectation theories take their cue from the fact that promises are typically a tool for providing reassurance to one another and for coordinating our actions.⁹¹

⁹¹ The role of promises in providing reassurance may seem significant only as a means to aiding in coordination. T.M. Scanlon has argued, however (and persuasively I think), that sometimes reassurance can be a good in itself, and one which promises can provide. A promise not to share your secret with others, for example, is a fairly common sort of a promise. Generally, however, we seek each others' reassurance that a secret will be kept in confidence, not in order to be able to plan our future action around that confidence, but simply because we care about our secrets not being told. There seems to be a value to reassurance itself.

So, for example, when I make a promise to Dan to babysit for him this weekend, it is to reassure him that he will have my help, and to allow him to coordinate the various other plans he has to make (such as the time for which he should make dinner reservations, when he has to meet me at his house, and so on). In order to be able to underwrite this coordination, promises need to be more than just predictions of our future behavior: they need to control our future behavior. In other words, promises need to be normative.

A raised expectations theory does point to important aspects of both promising and personal commitment. After all, in addition to needing coordination with one another, agents also clearly need to coordinate themselves, both in order to undertake temporally extended activities, and as a condition of being able to coordinate themselves with others. From earning a doctoral degree to meeting friends for dinner, our plans often require being able to orchestrate our own plans and actions across time, and personal commitments do seem to be among the sorts of volitional acts that serve this coordinative role. In committing to working on her temper, Shannon gives herself reason to pursue this goal through (at least some) fluctuating preferences. Her personal commitment can outweigh one-off desires to do something else with her time, the relief that blowing up at a colleague would give her, and so on. Such stability is a necessary precondition for Shannon to develop the requisite skills and character traits to achieve her goal.

Shannon's personal commitment, however, does not achieve coordination in quite the same way as a promise. To see why, consider the other typical function of promising, reassuring the promisee that the promisor intends to follow through. When I make a

promise to you, you can coordinate your action around mine because I have given you a special sort of reassurance about how I will go on to act, and as a result, you are licensed in forming certain expectations. Promises offer this sort of reassurance in two ways. In the first sense, promises serve an epistemic function: they reveal the promisor's true intent to the promisee. Since I am unsure whether you really mean to meet me for coffee this afternoon, you say, "I *promise* to be there." Making a promise is a way of underlining your claim about your own intentions. On this first understanding, however, reassurance does not seem very useful in the intrapersonal case, neither for its own sake nor as a means to coordination, given the special relationship I have to my own mind. If I do intend to do a certain thing, chances are I know it, and don't need reassurance. And if I am not sure whether I really intend to do that thing, then I don't have the necessary knowledge to be able to provide reassurance.⁹²

As David Owens points out, however, more often promises aren't about informing one another of our future plans: rather, they are ways of giving each other *authority* over those future plans. On these occasions, promises are a means of reassurance because they give the promisor additional reasons to perform the promised action, namely, that the promisee has the authority to *demand* performance. When I promise to meet a student for office hours tomorrow, for example, the student can assume that I will be there because I have obligated myself to be present. He now has a right that I be there tomorrow, and so (limited) authority over me. Furthermore, he has standing to

⁹² Of course, whether or not I really intend to follow through on my commitment may change over time. The point is that if I am aware of any fluctuations in my motivational commitment, I will be able to change my plans in response, in a way that I cannot in the interpersonal case. If I am not aware of my changing intentions, then I am not able to offer reassurance anyway.

insist upon this right, and to hold me accountable if I don't show up. In this sense, promises offer reassurance because they bind us, rendering actions that before may merely have been choiceworthy as *required*.

At a relatively general level, this claim is also true of personal commitments. As I have repeatedly claimed, personal commitments seem to give rise to new reasons for action by binding us, and in binding us, they could play a role in coordinating ourselves over time. Again, Shannon may make a personal commitment to improving her temper precisely because the constraints put in place by this personal commitment will underwrite the stability and persistent effort required to reach this sort of demanding and complex goal. And like promises, this stability and persistence is *normative*. Unlike causally forcing herself to pursue a goal, as Shannon would through a pre-commitment, normatively constraining her future self is less costly and allows for more flexibility. Pre-commitments and promises provide genuine guidance while leaving open the possibility of rational reconsideration, should our circumstances change.

Nonetheless, even if personal commitments are like promises in underwriting cooperation, the way they play this role cannot be explained by appealing to a raised expectations theory. Once again, on a raised expectations account, promises create new reasons for action in virtue of a duty not to disappoint those you have led to rely on you. One cannot, however, disappoint oneself in the same way that one can disappoint others. Even when one breaks a promise to oneself, one has a sort of warning or notice (namely, one's own intention) that a distinct promisor will not have, and one always gets *some* sort of benefit (if not the promised good, then whatever good induced one to break the

promise to oneself). In an interpersonal promise, of course, a broken promise often means that the promisee has missed out altogether. Furthermore, as I will claim in the next section on transactionalism, the normativity of personal commitments does not seem to trade on the committor bargaining with herself – indeed, it will turn out to seem quite odd to say that a single agent can enter into that sort of transaction with herself. Even if our personal commitments do serve a coordinative role, and it seems that they do, they do not seem able to do so in the same way as promises.

Transactional Theories of Promising

Finally, let's consider a transactional model of promising. On a transactional model, the normativity of promises is explained in terms of the transfer of rights from one agent to another: to make a promise is to give the promisee the right to the promised good. To use the Hohfeldian categories⁹³, when I make a promise to you to *phi*, I am alienating what was previously my liberty-right to *phi* or not, as I like, by handing you as a claim-right to demand that I *phi*. And if you have a claim that I *phi*, then I have corresponding a duty to you to *phi*. Broadly put, for the transactionalist making a promise is about handing over a particular bundle of claims and normative standings. So if I promise to meet you for lunch, the transactionalist will say that I have handed you a claim that I will come to our lunch date, with the attendant standing to demand that I

⁹³ Of course, on the classic Hohfeldian pictures, these rights are understood as either legal or moral. As I have suggested above, however, I think we have reason at least to leave the door open to the possibility of a distinct kind of normativity. I hope to provide some additional support for recognizing that kind in the next chapter.

show up, and the power to waive your claim if you decide you don't want to meet for lunch after all.

This model very clearly picks out the sort of normative transfer that occurs in promising. When I make a promise to you, I really am giving you a certain (limited) authority over me. That authority explains both why I have to do what I have said I would do, and why you alone have the right to release me from my promise. Furthermore I do not, as with raised expectation theories, indirectly transfer this authority by manipulating your psychological states. Rather, as an agent, I am licensed to give others certain claims over my future behavior, via employment contracts, marriages, and also promises. As Shiffrin said, these powers seem to be precisely part of what it is to be autonomous.

But how would this model apply to the case of personal commitment? On a transactional model, making a promise to oneself would mean giving a claim over oneself, to oneself. In making a personal commitment to vegetarianism, for example, Jason would be giving himself a claim (against himself) that he not eat meat, at least in the absence of more weighty considerations or him revoking his own personal commitment. In order to revoke this personal commitment, of course, Jason would simply have to hand that claim back to himself. If this language sounds tortured, it is for good reason. Even if we can imagine this sort of internal handing off of rights, it is not clear what normative upshot it would have. As we've said, I can "give" myself ten dollars, if I take the ten dollars out of my own desk and put it in my pocket – but it is not very interesting if I do so, since there is no *normative* change. The physical location of

the money may have shifted, but any normative claims (here, about who owns it) have not. For normative changes to occur, on a transactional picture, ownership of some normative elements (whether claims, or liberties, or so on) has to change hands. Similarly in the case of personal commitment, simply passing rights from my left hand to my right (metaphorically speaking) doesn't seem to make a normative difference. If we accept this picture, we seem to be left once again with the result that the committed agent could release herself at will.

Perhaps the most obvious response here would be to reassert a parallel between personal commitments and promises by dividing an agent into distinct "selves," either different temporal parts, or contemporaneous but distinct parts of the self (like Freud's division of the mind). If I am made up of multiple "selves", then perhaps a personal commitment is one part of the self handing a claim to another part of the self, like my present self making a promise to my future self. This would allow us to recapture some idea that the transaction has a normative effect, since rights would actually change hands (in some sense). On this suggestion, personal commitments would underwrite intra-agential cooperation by giving different "slices" of the self rights over each other.

Ultimately, however, this proposal seems a little metaphysically odd. Of course, there is a sense in which agents are made of distinct "parts": most people are not perfectly internally unified across time, or even at a single moment.⁹⁴ It is a significant jump, however, from saying that an agent's point of view needn't be univocal, to saying that she can engage in normatively binding transactions with herself. The fact that I am

⁹⁴ See, for example, Philip Bromberg's *Standing in the Spaces* and Jennifer Radden's *Divided Minds and Successive Selves*.

sometimes quite shy and reticent, and other times voluble and outgoing, shows that my personality reacts to different circumstances (my environment, my underlying mood, etc), in interesting ways, but that does not begin to address whether my shy and outgoing sides can enter into contracts with one another. And whether or not it turns out that these sorts of internal transactions are possible, it would be a much simpler account simply to say that agents are capable of making various kinds of future-oriented decisions, some of which bind quite gently (like intentions) and other of which (like personal commitments) bind us more tightly. As I'll argue in the next and final chapter, I think rejecting this simpler account comes with a fairly high cost for our understanding of agency.

Thus, in the end, analyzing personal commitments as promises to oneself doesn't seem to work. They are indeed similar in some ways: personal commitments seem like promises in resulting from the exercise of a normative power, and in a particular a normative power that allows an agent to bind her own self. Furthermore, reflecting on the nature of promising has led to some insights on the nature of personal commitment: our personal commitments aid us in coordinating ourselves over time, and even when outweighed, can leave behind a type of normative residue. Perhaps most suggestively for our positive defense, the normative power of personal commitment seems like it may be part of the most robust picture of agential autonomy: it may be the case that we can't fully understand the way in which agents can normatively define themselves unless we recognize the power to make private normative commitments. But despite these lessons, personal commitment seems to be distinct from the normative power of promising. It is

characterized by a distinct set of exit constraints, a distinct kind of normativity, and a distinct grounding story. Of course, no one disanalogy on its own should be enough to show that personal commitments and promises are really distinct, rather than just variations on a theme. As the disanalogies between personal commitment and each of our three contenders have added up, however, it becomes less clear what we have to gain by lumping the two normative powers into a single category. In the final chapter, I will suggest what I think we have to lose if we fail to recognize personal commitment as a *sui generis* normative power.

Towards a Theory of Personal Commitment

So far we have considered and at least provisionally rejected three different prospective accounts of personal commitment. Appealing to a psychologically strong intention, the demands of integrity, or an inwardly-directed promise each shed light on the phenomenon, but ultimately just cannot recapture the initial experience of personal commitment with which we started. Of course, as I've noted along the way, one response would be to say "so much for our intuitions", and simply revise our intuitive response to acts of personal commitment. Thus, the question now is whether we can come up with a better proposal that retains our original picture of personal commitment: what *do* personal commitments look like, and *why* do they bind us?

In this final chapter, I have two main goals. The first is to offer a fuller positive proposal as to just what a personal commitment is. For our purposes, I'd like to focus on three general features of the normative power of personal commitment: the normative consequences of making a personal commitment, the conditions that define a successful act of personal commitment, and the virtue that guides when and how one ought ideally make or revoke a personal commitment. To give just a preview, I want to claim that the normative power of personal commitment is best thought of as generating robust but extra-moral requirements constraining future deliberation, through an act of the will that builds upon and crystallizes an antecedent value, and whose best use is guided by the virtue of integrity.

After going through each of these claims in more detail, my second goal is to articulate just what the costs would be of rejecting this proposal. That is to say, I aim to give voice to the philosophical work this conception of personal commitment can do, and to the useful insights we would lose without it. Thus, rather than attempting a head-on argument for my model of personal commitments, my strategy is a bit more round-about. We have seen throughout this dissertation that personal commitments do have features in common with other normative phenomena, including intentions, integrity, and promising, and that appealing to these other models can shed some light on what it is to make a personal commitment. Nonetheless, I will argue in this chapter, analyzing personal commitment as *nothing but*, e.g., promising, would come at the cost of accepting vaguer, less conceptually powerful theories of both personal commitment and promising. The same will hold true, I claim, of any analysis that fails to recognize personal commitment as a unique kind of normative power. If, instead, we hold on to the characteristics that make personal commitment distinctive, we can make the power of personal commitment itself less mysterious while also shedding light on broader questions about the nature of agency and normativity. And in showing which conceptual tools and philosophical insights we stand to gain by accepting my positive proposal, I hope to earn some room for such a normative power.

A Positive Picture of Personal Commitment

In this first half of the chapter, I'd like to sketch out my positive proposal. First, what are the normative consequences of making a personal commitment? That is to say,

once we have made a personal commitment, how does that commitment bind us *synchronically*? Second, what are the conditions that determine when we can make (and revoke) personal commitments – what counts as a successful exercise of the power of personal commitment, and under what conditions can we “take back” a commitment? And finally, even when I am licensed to make or revoke a personal commitment, are there better and worse ways of doing so? On occasions when I can successfully exercise my power of personal commitment, what determines whether and how it is really appropriate to do so?

The “Normative Footprint” of Personal Commitment

Let’s start with the claim that a personal commitment has a distinctive kind of normative output, that the nature of a personal commitment is in part defined by the kind of normative requirements it generates, the characteristic “flavor” of the normative force of these requirements, and the attitudes of holding responsible that it licenses. In particular, as we have already seen, one thing that seems to characterize personal commitment is that they give rise to *requirements*. When Jason considers whether or not to have a ham sandwich for lunch, his personal commitment to vegetarianism doesn’t just weigh strongly against the taste of the sandwich, the cost of the alternative meals, and so on. His commitment *demand*s that he choose a vegetarian option. As I’ve already suggested, I take these requirements to be structurally and functionally analogous to Raz’s protected reasons: in addition to giving Jason a first-order reason to choose a vegetarian lunch, his personal commitment directs Jason not to act on a range of other,

excluded reasons (again, with one example being a hankering for ham today).⁹⁵ To say that a personal commitment generates *requirements* is to say that it does more than just make the object of commitment valuable, or even very valuable. Making a personal commitment is to make it the case that one *must* follow through on that commitment, and hence puts some future options off the table altogether.⁹⁶

Corresponding to the norms that they generate, personal commitments seem to demand certain characteristic reactive attitudes on the part of ourselves and others. At least first-personally this seems relatively straightforward: we ought to hold ourselves to our personal commitments, and experience some sort of negative reactive attitudes if we fail. To take a philosophical example, when Descartes says of his method of doubt, “and not having restricted this Method to any particular matter, I promised myself to apply it as usefully to the difficulties of other sciences as I had done to those of Algebra”⁹⁷, I take Descartes to be saying that since he has committed himself to examine the full scope of his beliefs in the light of radical skepticism, that he will at the very least hold himself responsible for doing so so. And if Descartes allows some of his standing beliefs to slip by without serious reconsideration, he should feel disappointed in himself and perhaps embarrassed that he has failed to live up to his own standards. To take another example:

⁹⁵ Though clearly, there is a scope that defines which reasons are excluded. If Jason having a ham sandwich can somehow save his family’s life, for example, then he really ought to ignore his personal commitment.

⁹⁶ In addition, there are often likely to be correlative requirements on third parties to defer to my personal commitments in various ways. If David has made a serious personal commitment to learn German, it seems wrong of Liz to try to get him to learn Spanish instead. This seems most obvious if Liz is trying to get David to cheat on his commitment – there seems to be something underhanded and devious about leading a committed person astray – but this may also be the case if she is simply trying to convince him to drop this personal commitment in favor of another one. At the very least, Liz seems guilty of encouraging a certain capriciousness on David’s part.

⁹⁷ Descartes, 15

if I myself fail to follow through on a personal commitment to spend the weekend writing, I may reasonably feel that I have done something wrong, and that reactive attitudes like shame or even guilt⁹⁸ towards my own behavior are in order. My desire to watch television or go out with friends instead of working should appear as a *temptation*, rather than as simply one urge among others, and I should feel honor-bound to resist that temptation.

Just which reactive attitudes other agents are licensed to take towards my failures is a slightly more complicated question. I think there will be lots of cases – perhaps most – in which second parties are simply not positioned to hold me responsible for my personal commitments.⁹⁹ If I violate a personal commitment to spend the weekend writing, my gut reaction is that this is simply not the business of my students, most of my friends, or my mother. I might rightly feel that friends or family have overstepped their bounds if they criticize me without any invitation on my part to do so. Just as it can be meddling for a third party to attempt to adjudicate a contention between two friends, so it can be overstepping for someone else to try to hold me to my personal commitments.

And like the interpersonal case of a relative stranger sticking his or her nose into the dispute between a friend and me, notice that the inappropriateness of this act is more than an epistemic issue. If someone knows about my commitment, he can of course

⁹⁸ Whether or not guilt can be a fully private attitude (held toward oneself, about fully self-regarding wrongs) is a matter of some dispute: as contrasted with shame, which is commonly understood as a reaction to one's failure to live up to some sort of ego ideal, guilt is a more legalistic notion. Nothing I say in this project really hangs on which way we fall, but personally I have the intuition that we can genuinely feel guilt over fully self-regarding issues. After all, as I think Kant persuasively argues, we can violate obligations to ourselves just as easily as we can violate obligations to others.

⁹⁹ Coleen Macnamara, unpublished dissertation, *Beyond Praise and Blame: Toward a Theory of Holding Others Responsible*.

correctly judge that I have failed in some way, but more is required to go beyond mere judgment to some sort of active holding responsible.¹⁰⁰ To hold me responsible, in this sense, is to take up an active attitude of blaming, and perhaps even to express that blame outwardly.¹⁰¹ Even if casual friends from college did happen to be aware of David's personal commitment to learn German, I think it would be inappropriate of them to presume to blame David for failures to follow through. To be brief, it would be presumptuous.

In fact, I take a third party who has the standing to hold me responsible for my personal commitments to be rather the exception than the rule. Unlike a promise, to which most or perhaps all members of the moral community are licensed to hold me responsible¹⁰², fewer people seem to have the standing to hold me to my personal commitments. It's plausible to think that Jason's wife may take him to task for eating a ham sandwich, and Annie's very best friend may express her disapproval if Annie fails to raise money for the community theater. But notice that these are close and relatively rare relationships: to have standing to hold someone else to his or her personal commitments, I must be invited to take up this position in some way, either explicitly in the case of this commitment or as a more general and standing intimate relationship. It would be out of place for a mere acquaintance or co-worker to attempt to hold me to my personal

¹⁰⁰ Angela Smith, for example, makes use of this distinction in her article "On Being and Holding Responsible".

¹⁰¹ Of course, my personal commitments may overlap with distinct obligations, obligations to which other agents can hold me. My dissertation advisor may feel disrespected if I slack off in my plan to write, but arguably this is because of obligations I have to her, not because of my personal commitment. Though the content overlaps here, we can distinguish between my personal commitment, which I owe myself, and the due diligence that I owe to my advisor.

¹⁰² Of course, not everyone will be appropriately placed to *say* something, but this is different from being able to judge.

commitments. Besides being private in the sense that no one else need know about them, my personal commitments are “my business”. Annie’s co-workers may be justified in using any failure on her part as some evidence of her character and strength of will, but it is a further step to blame her for this failure. Such blame on the part of a mere co-worker seems not only uncalled for, but actually impertinent. More than simply having knowledge of my personal commitment, having the standing to hold me responsible seems to require having some sort of stake or invitation into my personal life, and indeed into this particular bit of my personal life.

Besides the structure of their normative output, and the ways in which we can be held responsible for them, personal commitments are also characterized by a distinctive *kind*, or perhaps to use a gentler term, a different *flavor* of normative force. I argued throughout Chapter 2 that instrumental reasoning alone is insufficient to recapture the sorts of requirements we have been describing; personal commitments can certainly bind us against mere desires to do otherwise. On the other hand, as I started to argue in the last chapter, personal commitments do not seem to be typically moral, either.¹⁰³ Again, if Carin makes a personal commitment to run a marathon, then it does seem to be the case that Carin really *ought* to do her daily training, that she really *ought* to register and run in a marathon once she is able to run the full distance, and so on. These “oughts”, however, just don’t seem like moral “oughts”. Though Carin would still count as failing in some

¹⁰³ Of course, some personal commitments may be morally significant. If Carin’s commitment to run a marathon is to raise money for charity, as part of fulfilling her imperfect duty of beneficence, then this particular commitment will also have moral weight. But in this case, what seems to be doing the moral work is the morally weighty content of her personal commitment, not the fact of commitment itself.

way if she doesn't train regularly for her marathon, it still feels odd to say that she is being a morally bad person, or that her moral character is implicated. It strikes me as too strong to say that someone's personal commitment to running, or to participating in a community garden, or to learning a foreign language, could implicate his or her *moral virtue*.

Indeed, lumping things like personal commitment in with our moral duties seems to water down our intuitive sense of what counts as moral. If everything aside from instrumental reasoning and perhaps also prudence falls into the realm of the moral, moral norms will end up sharing fewer substantive characteristics in common. To call a norm "moral" would be not only less evocative, but less informative.

Of course, whether or not personal commitment is a moral phenomenon will depend on one's theory of the moral. Most obviously, personal commitments (at least as a class) fail to count as moral on any theory that defines the moral as essentially other-regarding or interpersonal, such as what we have seen of Scanlon and Darwall. If morality is centrally about our obligations to one another, then personal commitments are something else. Other moral theories, of course, are more inclusive. Kant, familiarly, was more than happy to recognize moral obligations to the self, but then these are universalizable moral duties, not particularistically imposed personal commitments. On other models, however – such as an Aristotelian sort of approach – personal commitments may very well fall within the purview of the moral. The sorts of commitments one makes, and how one does or does not follow through on them, seem to

be very much a part of whether we can call one's life flourishing. And, though I have argued that personal commitments issue deontic rather than aretaic norms, we'll see later that there is an intimate relationship between the virtue of integrity and the power of personal commitment.

Now, at the end of the day, the real issue is obviously not whether or not we call personal commitments "moral". My point, rather, is that personal commitments don't fall into a certain core understanding of morality, one which is defined by the issuing of claims, a fairly widespread or even universal liability to be held responsible, and a very hefty sort of significance. This core understanding includes things like promissory obligations, which are almost always trumping considerations, and to which more or less the entire moral community can hold us. To contrast, personal commitments also give rise to requirements, but these requirements are generally defeated by moral and often even legal considerations. Compare a moral obligation to defy an immoral law, like racial segregation, to a personal commitment that contravenes a morally neutral but socially useful law, like a commitment to enjoying the feeling of the wind in one's hair rather than wear a motorcycle helmet. Even though there may be circumstances in which it may be morally permissible (if somewhat reckless and self-indulgent) not to wear a helmet – one is tooling around a lightly-traveled country road – for most, a legal requirement to wear a helmet can at least sometimes trump a commitment in a way that an immoral law to obey even some relatively trivial segregation clearly cannot. Furthermore, it feels counter-intuitive to claim that just anyone can hold us responsible for our personal commitments, as opposed to the quite broad license to hold us

responsible for our moral commitments: a sort of intimacy seems required that is not necessary to hold one another to our moral obligations. Recognizing this sort of distinctively commitment-based flavor would complicate our picture of normativity, of course, but it would also allow us to hold on to a richer, more detailed picture of what makes a *moral* norm distinctive, as well as what characterizes personal commitment as a kind.

Personal Commitment as Building on Value

Second, we can ask under what conditions I am able to effect this sort of normative change. That is to say, what does it take to have a personal commitment “come off”, and how can I then get out of commitments that I’ve made? In this section, I want to motivate the claim that there are two intersecting components needed to commit oneself successfully: first, that one wills oneself to be committed (the volitional act), and second, that what one commits oneself to is an antecedently valuable end (what we might call the “value constraint”).

To see how these two conditions intersect in a personal commitment, let’s return to the example of Annie making a commitment to raise the money needed by the local community theater. Now, the first component – the volitional act – is pretty straightforward. In order to make her commitment, Annie clearly needs to *will* that she be under such an obligation, as opposed to commitments with which we simply find ourselves, (as when Frankfurt describes “love” and “caring”), or the various ways in which we can passively end up committed, (as when we come to define ourselves

through built-up patterns of behavior). Annie needs to commit herself actively. And unlike social commitments, including promises and legal contracts, Annie seems able to exercise this normative power on her own.

Of course, unlike an interpersonal promise, there isn't just one clear formula by which we do this. Besides being overly legalistic, (even in the case of promising, I can sometimes implicitly bind myself without actually using the words "I promise), insisting on a single magic phrase by which we commit ourselves would be false to the phenomenology. In making a personal commitment I may internally say, "I am committing myself to *phi*", or "I promise myself that I will *phi*", or "This time, I am really going to *phi*." Indeed, on occasion I may make personal commitments without really verbalizing the act at all; on these occasions, I may simply recognize myself as having settled on an end in a committed manner.¹⁰⁴ What all of these various acts have in common is that, in each case, I settle upon a commitment and decide to bind myself through it. Like any other normative power, I make a personal commitment through an

¹⁰⁴ This may sound odd: given that there is no one locution defining the act of personal commitment, and that the various turns of phrase that I've offered can be employed in other ways than making a personal commitment, one might wonder how an agent can tell whether she has really bound herself. If the commitment is to myself, and constituted by no more than an act of my will, how can I be mistaken or confused about whether I've made one? In fact, however, cases in which we fail to be transparent to ourselves are not unfamiliar. A father may not realize that he believes his son is physically tougher than his daughter until a third party points out the different ways in which they treats each of them; Freud's "Rat Man" fails to recognize his latent hostility towards his own father because he unconsciously buries that feeling under what could be loosely called neurotic behavior (Velleman, 2002, 101). Sometimes it can take a while, not to mention a helpful conversation with a friend, to come to see that one has been mistaken about one's commitments. Ultimately, I'm not certain there is any one simple and foolproof way to ascertain that one has made a personal commitment: at the very least, it will require a realistic view of the antecedent importance of the object of commitment and the authenticity of one's act of committing. But the fact that we can sometimes be mistaken in that belief needn't impugn that there is a fact of the matter as to whether we are committed.

(And, indeed, the fact that there can be different particular internal acts by which we commit ourselves will turn out to be important, as we will shortly see.)

exercise of my will. Or to borrow the motto of the London Stock Exchange, “my word is my bond”, though the “word” here may be internal and silent.

The second component of a successful personal commitment is that it builds upon an end with some antecedent value.¹⁰⁵ That is to say, the volitional act of committing oneself comes off only if one is trying to commit oneself to an objectively worthwhile end.¹⁰⁶ To begin to make this claim plausible, think of the sort of language we have used in describing personal commitments: to make a personal commitment is to “take a stand”, or to “stake oneself” on something. Reflecting on these locutions seems to suggest that there must be something on which one has taken a stand, or on which one has staked oneself. To continue the metaphor, one might ask whether I can really stake myself on an illusory value, or if the commitment is to bind me genuinely must there be something to which I am binding myself?

¹⁰⁵ Here I mean something like “objective value”; a personal commitment only comes off if it is to an end that really is valuable. (We will address the way in which the agent ought *subjectively* to value the end to which she commits in the next section.) See Ruth Chang and Susan Wolf for interesting recent work on the way in which agents can play with and inflect these sorts of objective value.

Of course, just what counts as the minimum threshold having sufficient value will depend on the particular theory of value one brings to the table, something about which I am agnostic here. As long as one has some version of a realist theory of value – that there are at least some values that don’t need to be grounded in what a particular agent contingently happens to desire – the idea is that this theory will give us some way of distinguishing ends that are valuable from those that are not, or are only negligibly so. I am claiming that we can make personal commitments to the former and not to the latter.

¹⁰⁶ This invites the question as to whether we can successfully make personal commitments to immoral ends. In the vast majority of cases, I take it to be obvious that an agent will in no way be required to act on an immoral personal commitment, but this leaves open the question as to whether an immoral personal commitment is binding by being outweighed, or whether it fails to bind in the first place. As with promises, there will be disagreements as to whether an immoral commitment comes off and is defeated, or doesn’t come off at all.

Though there will generally be relatively little practical difference between the two claims, I for one am inclined to say that an immoral commitment fails to come off – that we are not even *prima facie* bound by such personal commitments, though as always, we can still be criticized for attempting to make these commitments, or for failing to follow through on them if we mistakenly take ourselves to be bound.

It is certainly the case that the most plausible examples of personal commitment are to ends with their own antecedent value. While we may readily take ourselves to be bound by commitments to improve oneself (in mind, body, or character), to play a more active role in one's community, and so on, we are likely to have a different reaction when someone claims to have made a personal commitment to a valueless end. It sounds at least odd to claim that agents can make binding personal commitments to things like counting the pieces of fringe on a rug, or avoiding use of the preposition 'at', (at least without a significant backstory that reintroduces some value). Our more likely reaction in those cases is to take the agent to be mistaken or confused about what she only thinks is a binding commitment, or perhaps even to take her confusion as evidence that she may be mentally ill.

On the other hand, one might worry that limiting personal commitment to ends with antecedent value empties the act of commitment of any significance, or else fails to reflect the significance of commitments that are made to trivial or even problematic ends. If we can only make personal commitments to valuable ends, are we just pinning all of the importance on the antecedent end?

Now, I don't think this is true. At the very least, as we have already seen, the act of commitment isn't redundant because it can shift a valuable end (one worth pursuing) to a required end (one that one must pursue). To go back to a familiar example, while supporting the community theater is already a worthwhile thing to do, Annie's personal commitment makes it something that she is now *required* to do, and *blameworthy* for not doing.

But even if the value requirement can co-exist with the need for a volitional act, one might ask, why think it's necessary? Why insist that we are constrained in the personal commitments we can make? To compare, promises can be made to quite trivial ends. If Shannon, for example, promises her husband to speak only in rhymed couplets next Tuesday, or to hum "I'm a Little Teapot" each morning at ten, then I think that she is really bound to do so.¹⁰⁷ She is bound simply because she made the promise to her husband, and he accepted it. (Similarly, if I offer you a gift which it turns out you don't really care for, I don't just get to take the gift back. Once you accept the gift, it is now your property, whatever your reason for accepting it in the first place.)

But though I do take this view of promising, I think it's a mistake to adopt an analogous model of personal commitment. Rather, I think that it overstates our agential power to claim that we can commit ourselves in this way to just anything. Personal commitments seem to be a way of *inflecting* value, of allowing agents to be active with respect to normativity. Shannon's commitment changes what it is that she has to do. Nonetheless, that activity is not unbridled. In the case of personal commitment, we build upon a value by staking ourselves on it, and to earn its normative significance, that staking has to be meaningful. Shannon does shift the object of her commitments from something worth pursuing to something required. And while part of what makes an act

¹⁰⁷ For those who might be suspicious about the success of a "Little Teapot" promise, let me point out that Shannon's promise won't be binding unless her husband takes her up on it, and that if her husband is a reasonable agent (and hence the kind of thing to whom one can make promises in the first place), then there has to be some story as to why he would accept this promise. The story that makes his acceptance intelligible, however, needn't be directly in terms of the value of the promised good itself. He may whimsically decide to hold Shannon to her offer, or he may think that accepting Shannon's promise will help her to develop some more responsibility and learn to follow through on her word, etc. However trivial or ridiculous their content, our promises bind us just in case the other party takes us up on that promise.

of personal commitment meaningful is the agent's willingness to commit herself, her intention to bind and constrain herself, unless the end to which she binds herself has some value of its own – unless the end is *worthy* of commitment – an act of commitment simply seems simply misguided.¹⁰⁸ Though there may be some consequences to attempting to make such a commitment, that doesn't mean the commitment itself has come off.

Furthermore, I think that a recognition of the connection between personal commitment and the values on which they build is necessary to fit the phenomenon of personal commitment into a realist view of reasons. Now, of course, there are different camps as to the fundamental nature of reasons, and I recognize that this will be a moot point for some. Many of us, however, share the intuition that what it is to be a reason is, in some sense, independent of what actual human beings contingently care about or find motivating. Though I may very much feel like drag racing down a busy road, and not care much about the negative consequences, my own physical safety and the safety of others who might be in my path are reasons not to do so. On a broadly realist view of reasons, however, the idea of something like a personal commitment can sound

¹⁰⁸ Of course, to the extent that an agent thinks that he has made this personal commitment, he ought to follow through on it; though the point is both trivial and somewhat awkwardly put, we ought to do whatever it is that we think we ought to do. To take a parallel case, I ought to follow whatever I believe to be the rules of English grammar in writing. If, however, I am objectively mistaken about English grammar – for example, I think that English grammar always requires a conjugated verb as the second element in a sentence – then there is also a sense in which I ought *not* follow what I take to be its rules. (Since I actually ought to use proper grammar, and this is not a rule of English.) Similarly, we can account for an impression that an agent should act on a “personal commitment” to counting fringe by appealing to the importance of doing what one thinks one ought to do, without claiming that this person has actually changed what he ought to do. He ought to act consistently with what he takes to be a binding personal commitment, but if he were to wake up one day and realize that there is actually no independent value to fringe counting, there would be no reason for him to continue to act as if he were so bound. (Which, importantly, is different from saying that he'd be justified in revoking his commitment.)

mysterious. If reasons are something that are, (again in some sense), just *out there*, how can it be the case that an act of my will can bring new reasons into being?¹⁰⁹

Of course, even on this realist picture of reasons, the idea that agents can change what it is they have to do really needn't be odd. The claim that we can change our reasons for action through social transactions like promising, for example, isn't so mysterious, since promises hook into standing reasons – I have a reason not to harm others, and causing them to rely on a promised good only to disappoint them would be a sort of harm. In the light of the privacy of personal commitments, however, it can sound mysterious to claim that personal commitments can by themselves add importance to just any end, from developing the character trait of patience to protecting the dust bunny under my bed. And indeed, focusing on examples like a commitment to caring for a particular dust bunny is precisely part of what can make personal commitments sound too odd to accept. We will continue to address this question, but at least in part, I take insisting on a value constraint to be a compromise, one that allows us to see a normative power of personal commitment while simultaneously recognizing the limits to what agents can do.

As I have suggested above, there is naturally some variation across particular commitments in both the volitional act (the specific way the agent internally phrases her act of commitment, or the way in which she conceives of that act), and in the value to which she commits (in terms of the nature and strength of the value in question). Each of

¹⁰⁹ I take this to be a more concrete version of the general bootstrapping worry we saw in the first chapter.

these components can come in varying degrees, though both must be present in some way. Some personal commitments will involve the agent very carefully, deliberately telling herself that she is now committed to that end. Other personal commitments, while still a product of the agent's will, may be slightly less explicit, coming across as more a moment of conviction or resolution, in which the agent simply realizes that this is what she has to do. Similarly, as we have seen throughout this dissertation from our examples, personal commitments can be to very weighty ends (like developing patience that will mend one's relationships), to much less weighty ends (like learning German), and even to ends that are valuable only within a given context (like participating in certain religious or cultural practices that make sense within a very particular social context).

Given this variation, I'd like to say just a little bit more about the relationship between the two components that I've suggested are necessary for a personal commitment. Specifically, I'd like to suggest that a personal commitment can have a lesser degree of one component (value or volitional act) as long as it's compensated by a larger degree of the other. So, the idea goes, a personal commitment to an end with lesser or somewhat questionable value requires a more thoughtful and deliberate act of commitment, and a less careful act of commitment will only succeed in binding you if it builds on a more valuable end. I think we account for this effect by looking (again) to paradigmatic examples. Compare a personal commitment to learning German with a commitment to developing much-needed patience. Given the same circumstances – say, a relatively casual but nonetheless deliberate moment of commitment, we are more likely to hold Shannon responsible for a personal commitment to develop patience than David

to his commitment to German. For this, relatively speaking, more trivial commitment to learning German to be binding, it seems that David must have more seriously and clearly undertaken the act of commitment. Though the end to which David commits may, itself, be less significant, we can see the way in which a more thoughtful and deliberate act of commitment on his part could compensate. The end to which he commits still needs to be valuable, but if his commitment is to come off, David seems to need to commit himself more seriously. A relatively flighty “commitment” to a potentially interesting hobby feels more like an intention than a commitment. And, conversely, a personal commitment to something very serious seems an easier normative act to pull off. We will have an easier time holding Shannon responsible to becoming patient, in part because this may be an end she already has serious reason to pursue, but in part because she ought to take the act of commitment more seriously when more is at stake. Put more generally, given the two factors at work we can recognize how an increase in one factor (the seriousness of the volitional act, or the significance of the value) can compensate for less of the other component in committing oneself.

And turning now from how we enter into a personal commitment, we can ask what it looks like to revoke one’s commitment. Whereas making a personal commitment seems to require the right *combination* of independent value and an act of the will, I want to propose that revoking a personal commitment involves the same two components, but in a *disjunctive* way.

To start with a straightforward case, we generally do seem able to revoke a personal commitment through a fairly straightforward act of the will. “Volitional act in, volitional act out,” one might say. If David, for example, reflects on his personal commitment to learn German, and decides that the commitment requires more of him than it is worth, then he can free himself just by willing it to be so. Particularly if David *genuinely* reevaluates his willingness to stake himself on this particular end, he seems to have the power to give that commitment completely up.¹¹⁰ Indeed, I think it would be somewhat odd to claim that David could enter into the personal commitment through an act of the will, but that afterward he could only be freed from its force through the intervention of someone else, or due to changing physical or psychological circumstances. Intuitively, we can imagine David stepping back and reconsidering the merits of his original commitment – if he finds that his commitment is no longer serving the reasons for which he adopted it, or perhaps decides that he was mistaken about its independent value – and then having the power just to “take it back”.

One’s initial reaction to this claim might be that it either seems to make personal commitments only trivially binding. If we are free to revoke our personal commitments, how binding are they? And if some sort of substantive mistake is necessary for exit, are we really in control over those commitments? I’ll have more to say about this worry in the next section, where I discuss the guiding virtue of integrity, but let me just say a couple of things here. In the first place, saying that I can exit a commitment with relative

¹¹⁰ Of course, even when David genuinely revokes his personal commitment, other sorts of criticism may be available – he may have acted too hastily, or capriciously, or with a lack of integrity. (This last is a point to which we’ll return in the next section.) Raising these sorts of criticism, however, is fully compatible with recognizing that David is no longer bound by the commitment in question.

ease is not necessarily a comment on how demanding it is before I do so: as Allen Habib points out, “at will” employment is defined by the fact that either party can exit the relationship simply by giving appropriate notice, yet until such notice is given, the employee really has to perform her job duties, and the employer really has to pay her. Similarly, we can accept the claim that personal commitments are revoked with relative ease without thereby accepting that personal commitments are not robustly binding. Secondly, even a genuinely revoked personal commitment may leave behind normative residue of a kind we don’t see in the case of a mere intention. If Jason has merely entertained a plan to be a vegetarian, he can revoke this plan for any reason, even if he reflectively doesn’t think his reason is a particularly good or noble one. By contrast, (and again, this is a claim to which we’ll return shortly), if Jason’s reasons for revocation don’t really address or take into account his original reasons for commitment, this act of revocation may leave behind some normative residue, or reflect on his character in some way.

But if one way of dissolving a commitment is active, I now want to claim that there is also a second, more “passive” way. Once the personal commitment no longer leads to or shares in the value to which one originally committed, the commitment may “fall apart”, so to speak. To see more clearly what I mean, let’s add a little more detail to David’s case.

Imagine David originally made his personal commitment to learn German within the context of his broader goals to write a dissertation on Husserl, and to pursue a career

researching his philosophy. As it sometimes happens, however, David might gradually lose interest in an academic career, letting his studies fall by the wayside as he begins to pursue an alternative vocation. In such a situation, we can easily imagine that this person might not think to step back and explicitly reconsider his commitment to learning German. Particularly in the midst of undertaking a serious change in his life's course, one might simply put to the side, overlook, or even end up completely forgetting one's personal commitment. If, decades later, our onetime graduate student is now a 65 year old social worker in the American Southwest, it seems silly to say that he really ought to be studying German, and is failing when he doesn't review German grammar and vocabulary. Though that might still be a valuable use of his time for various reasons, it sounds odd to say that he *must* devote himself to studying German.¹¹¹

Rather, I think it makes most sense to say that the personal commitment has, as it were, unraveled over time. No longer having the right sort of connection to an independently valuable end, the commitment is no longer binding. This, of course, is distinct from claiming that David is still bound by his commitment but really ought to revoke it, as perhaps might be the case when someone's personal commitment is turning out to be overly burdensome or not a good investment of his time and energy. Decades later, when it no longer makes sense for David to devote himself to learning German, it's

¹¹¹ Alternatively, one might claim that our agent is still bound by his personal commitment, but that he now really ought to revoke the commitment in question. I think that there are clearly cases in which an agent is bound by a personal commitment and yet, all things considered, revoke it, as when a personal commitment imposes overly burdensome prudential costs, or when one's time and energy could be used to much better effect. Nonetheless, I want to draw a distinction between those sorts of cases, in which an agent is bound but would be better off reconsidering her commitment, and cases in which a personal commitment no longer makes sense. In these latter cases, it is mistaken and indeed stupidly stubborn to insist that the commitment continues to be binding.

really not the case that David could make the situation right by turning back to his studies. This onetime personal commitment lacks the needed context to make sense. At this point, David's personal commitment no longer binds him. (Of course, all of this is compatible with recognizing that simply letting a personal commitment's normative force peter out is a subpar way of getting out of that commitment, a point that we'll consider in more detail in the next section.)

Personal Commitments and the Virtue of Integrity

Thus far, personal commitments may sound terribly easy to make and revoke, given the relatively cumbersome normative burden they impose. After all, I have claimed that so long as there is a genuine value at play, I can commit (and free myself) through a simple act of the will. If, after some thoughtful deliberation, Jason decides that he no longer wants to be committed to vegetarianism, he can simply take back his original personal commitment. And even if he doesn't do so, it is at least possible that Jason's personal commitment may be passively dissolved over time, if his circumstances change sufficiently.

Nonetheless, I do think the ease with which we can form and revoke personal commitments is compatible with the relatively strong normative force I have attributed to them. As we have already seen, cases like "at will" employment demonstrate that fairly thin entrance and exit conditions can be compatible with fairly burdensome normative demands. That I can opt out of my obligations whenever I feel like it doesn't impugn the fact that they bind me until I do so. Furthermore, simply addressing whether one has

successfully made or revoked a personal commitment is not the end of the story. Rather, the significance of our personal commitments is also reflected in questions of whether a personal commitment really *ought* to have been made or revoked. Even if Jason can free himself from his commitment under a given set of circumstances, is he fully right to do so? To put the question another way, what does it look like to enter and exit a personal commitment in a *fully virtuous way*?

We can gloss this question of appropriate use in terms of the “spirit versus letter of the law” distinction. So, for example, after Barack Obama resigned his seat in the United States Senate, Illinois Governor Rod Blagojevich had the legal power to appoint his replacement. But though Blagojevich in fact had (and exercised) that authority, one might criticize this as an inappropriate use of his authority – given the particular accusations of corruption against him, Blagojevich ought to have had the respect for his office, the position of Senator, and also the Democratic process to let his eventual replacement fill the seat.¹¹² Though he genuinely had the power to name a replacement (as validated by the Illinois Supreme Court and the ultimate acceptance of the U.S. Senate¹¹³), it would have reflected better on Blagojevich had he not exercised that power. In short, whether one has and can genuinely exercise a normative power can, at least under some circumstances, come apart from the question of how well one has exercised it, or whether one really should have employed that power at all.

¹¹² Obviously, this is a normative claim, and thus completely compatible with a lack of surprise that someone as corrupt and arrogant as Blagojevich would go ahead and appoint Roland Burris anyhow.

¹¹³ At least, of course, until Burris had to face his own allegations of corruption – I don’t take this, though, to reflect on the legal powers Blagojevich had at the time.

Like the governor's political power of appointment, I take the normative power of personal commitment to be importantly tied to a guiding virtue of appropriate use – indeed, I take this guiding virtue to be part of what constitutes personal commitment as the particular power it is. To compare, our understanding of what friendship is goes beyond simply recognizing the minimal standards necessary to count as being a friend; part of how we understand the very concept of friendship is in relation to what an ideal friendship would look like, (even though actual friendships never attain that ideal). So, by a guiding virtue, I mean a set of norms which define the appropriate use of a particular normative power, set forth a picture of its ideal use, and in doing so, help to constitute that particular normative power as what it is.¹¹⁴ To be clear, this goes beyond the claim that the normative power in question just *happens* to be frequently associated with or relevant to some virtue or other. While many virtues can be relevant to exercising a normative power – presumably Blagojevich's choice could be assessed by the lights of any number of virtues, including justice, honesty, and so on – I am claiming that a “guiding virtue” has a deeper connection with a normative power. It helps to define a particular normative power as such, constituting the ideal contours of that power, rather than just pointing to the intersection of a normative power with freestanding virtues.

The guiding virtue for personal commitments, I claim, is best described as integrity – integrity defines the ideal contours of personal commitment, above and beyond mere success conditions. As we have seen, the literature on integrity is fairly

¹¹⁴ I suspect that not all normative powers are coupled with a guiding virtue – the power to alienate my property by selling it or giving it away, for example, may not have the same sort of relationship with any given virtue as the power to make promises or personal commitments may. I leave teasing out that relationship, however, for a future project.

diverse, but again here I'm focusing on something generally like integrity as standing by one's idiosyncratic belief system, such that maintaining integrity is a matter of acting in ways that honor and express one's deep personal values. Integrity is about doing one's best by one's own priorities, just as the best use of one's personal commitments goes above and beyond merely refusing to overstep some legalistic boundaries. So, on this proposal, determining whether one has used one's power of personal commitment in better or worse ways requires evaluating one's acts of commitment with reference to one's personal beliefs and values, and looking at how well one lives out those concrete values in making and revoking personal commitments.

Now, to reiterate my claim from chapter 3, I don't think personal commitments can be understood *solely* in terms of any virtue; rather, in the first instance, personal commitments (like promises) place demands on action rather than character traits, and these demands distribute to each relevant situation, rather than operating as holistic constraints.¹¹⁵ A personal commitment is, as I've been using the term, an essentially *deontic* phenomenon. The case I want to make here is for integrity as a virtue that guides the ideal use of personal commitment, without being identical to it.

Integrity seems to be a natural choice for this role given the significant overlap between the demands of integrity and those of personal commitment. Recall that both personal commitments and integrity (particularly on our understanding) give normative weight to what agents, as concrete individuals, idiosyncratically take to be important. As

¹¹⁵ As a reminder, Annie's personal commitment to helping fund her community theater primarily shows up in a requirement that she raise the needed money, rather than implications for the state of her character. And Jason's commitment to vegetarianism operates each time he deliberates over what to eat, not as a holistic demand that he avoid meat a sufficient number of times.

we saw throughout the second chapter, whether or not Jason, Annie, and Shannon genuinely try to live up to their respective commitments can reflect on their integrity. And again, integrity and personal commitments also invoke similar sorts of metaphors: if I make a personal commitment to environmentalism, then I am “staking myself” on that end. Borrowing Cheshire Calhoun’s language, when I make a personal commitment, I am “standing for something.” Taking integrity to be the guiding virtue of personal commitment allows us to recognize this overlap without trying to analyze the latter as merely a matter of the former.

And as it turns out, appealing to this sense of integrity does nicely recapture our intuitive sense of the conditions for ideally creating and revoking personal commitments. Whether Shannon, for example, should make a personal commitment to developing patience depends largely on whether this end is one on which she can authentically stake herself. As integrity, on this understanding, is a matter of staying true to one’s idiosyncratic beliefs, so ideally comporting oneself with respect to personal commitments requires proceeding in accord with one’s other evaluative beliefs.

Imagine that Shannon makes her personal commitment for reasons which she doesn’t reflectively endorse at the end of the day, perhaps because of gender norms valuing submissive women which she would consciously reject, or out of social anxiety that makes her hyper-aware of even slight displays of anger, but which she rationally knows to be over the top. In these cases, I think we should say that Shannon has genuinely bound herself, but that in terms of the norms of personal commitment itself, it would have been better had she not. Particularly if Shannon was aware as she made the

commitment that she was acting on less than ideal reasons, I don't think we want to let Shannon completely off of the hook, just declaring her commitment null and void. This would be overly restricting our power of personal commitment, meaning that we could never successfully commit ourselves when we were feeling at all ambivalent or uncertain. Indeed, given the pervasive nature of confusion and conflict in human lives, this might mean that actual agents could never make personal commitments at all. Nonetheless, it is totally fair to say that it would have been better for Shannon not to make a personal commitment in these circumstances, that it would have been more ideal for Shannon to take up the end of patience in some other, less weighty way, or perhaps for her to have focused on analyzing her motivation to pursue patience first.

As distinct from other sorts of criticism about making a personal commitment – for example, that this commitment will impose too high of a prudential burden, or that a different commitment would be more rewarding – appealing to the guiding virtue of integrity is to raise a sort of *internal* criticism about her act of commitment. That is to say, while there are many considerations (including other virtues) that can intersect with questions about personal commitment, integrity seems to have a constitutive role to play in personal commitment. One can be more or less *honest* about whether one has made a personal commitment; one can also be more or less *generous* in discharging those commitments. These virtues, however, simply happen to overlap with the demands of personal commitment, just as one can be more or less honest and generous in discharging one's contractual obligations. Again, my promises can be generously given or begrudgingly fulfilled, but these sorts of assessments are about the intersection of two

independent sorts of demands (the normative power to make promises and the demands of the virtue of generosity). A guiding virtue, by contrast, fleshes out the contours of the normative power itself; it helps to define the *telos* of that power.

Let's return to a concrete example. Even if we agree that Jason has made a genuine personal commitment to vegetarianism, and has thereby bound himself to follow his commitment, appealing to integrity allows us to point to better and worse ways in which Jason might have done so. We can assess whether he really ought to have made that commitment by asking whether it stems out of his own sincere values, or by contrast, a mere desire to fit into a certain social group. The demands of integrity also mean that it is better, not just prudentially better but more laudable for Jason to make a personal commitment that he authentically feels should be a priority, and also for which he has the appropriate affect. If Jason goes ahead and (in a clear-headed moment) commits to being a vegetarian, he can bind himself even if he fails to feel a particular sort of pull or resonance with this commitment. I think, however, we can recognize it as a better use of Jason's power if his personal commitment is rooted in Jason's own value system – while there are many worthwhile ends to which Jason *could* make a personal commitment, it seems more appropriate that he choose those ends which are connected to his other commitments, values and priorities. There seems to be something more genuine and thoughtful about this sort of act. Furthermore, an ideal personal commitment is connected to the right affect; we should hope that Jason has the right emotional response to this end. Of course, a committed environmentalist may be reasonably pleased if her boyfriend commits to commuting by bicycle rather than car largely out of his love for

her, and take him to be bound by that personal commitment. And as I argued in chapter 3, I think it's a mistake to try to analyze personal commitment as any complex of emotional attitudes. Nonetheless, our environmentalist (and we) may reasonably take her boyfriend's personal commitment to be lacking in a significant way if it's not accompanied by a deep sense of caring for this end.

Similar points apply to exiting a personal commitment: as long as Shannon and Jason decide to revoke their respective commitments in a clear-headed moment (e.g., not drunk, or clinically depressed), and have decent reasons to do so, then they are no longer committed. They each have the ability to free themselves. (And, as I have claimed, it would be fetishistic to insist that they were irrevocably bound.) But unless this revocation is grounded in considerations which each personally and genuinely finds important, and which reflects a sort of respect for the original act of commitment, we can say that Shannon and Jason really ought not have opted out. Merely intellectual recognition that Shannon's personal commitment is taking up a disproportionate amount of time, or that Jason cares more about the convenience of an omnivorous diet seems to leave something lacking. Rather, ideal revocation of a personal commitment involves a sense of regret that one can no longer be committed, or perhaps regret that one committed to an untenable end in the first place. It should involve a sense of sacrifice or otherwise significant renunciation. These attitudes can clearly exist alongside a sense of disinterest in one's commitment, relief to be rid of it, and excitement to try new things. One may regret the need to give up a time-consuming and unrewarding commitment to a local charity, all the while feeling profound relief to be free from a draining obligation to a

poorly organized group. Nonetheless, the ideal revocation of a personal commitment should be accompanied by some affective recognition of the significance of this move. Such recognition should also (ideally) be grounded in a certain sort of genuine change of heart, a change in how one perceives the commitment's significance, and in one's feelings of attachment for it. These are signs of respect for the particular personal commitment, as well as for the agential power to make it in the first place. Even when we have the power to bind ourselves through personal commitment, we can appeal to integrity to make judgments of better or worse uses of that normative power.

Personal Commitments and Our Picture of Agency

So far, I've put forth a picture of personal commitments as a distinctive sort of normative binding, defined by a characteristic normative output, success conditions, and guiding virtue. Personal commitments issue requirements on action, and are characterized by a distinctive flavor of normativity. Commitment-based normativity is weightier than the demands of means-ends reasoning, but still falls short of the sort of significance and categoricity we generally see in moral demands (including the demands of promising). This demandingness builds upon something of antecedent value, rather than creating norms "out of thin air". And finally, the best use of these demands, both in their creation and revocation, is guided by the virtue of integrity – we can recognize occasions on which agents use that power when it would have been better that they not.

Now, as I already argued in the very first chapter, I don't think there are good in-principle reasons to be wary of this notion of personal commitment: specifically, I think

we can respond to worries that they necessarily involve problematic bootstrapping or private languages. Nonetheless, one might agree that both our pre-theoretical experience and the model I have been starting to develop fail to conform with our standard notions of intention, integrity, or promising, and yet conclude “so much the worse for personal commitment”. At the end of the day, one might just decide to sacrifice this conception of personal commitment. Why, instead, should we complicate our ontology by taking on this particular normative power?

In response, I think we can point to the value of this model by pointing to the ways in which it can enrich our understanding of reasons, values, and agency. Thus, I take the dissertation so far already to be a partial defense of the model I’ve been building. To the extent that analyzing personal commitment as an ersatz intention, issue of integrity, or promise has been insufficient, this is already reason to look elsewhere. But furthermore, I want to argue that rejecting a power of personal commitments forces us into a relatively impoverished picture of agency. In particular, it is an oddly asymmetrical picture on which we have the normative power to change our reasons for action simply by saying the right words to another agent, and yet on which we are normatively powerless on our own. It also, consequently, really limits the ways in which agents can be active creators of their own selves, and in doing so, flattens our understanding of normativity. Rejecting personal commitment leads us towards an understanding of normativity as exhausted by a few very broad types, such as the

instrumental, the moral, and perhaps the prudential. In the remaining pages, I'd like to sketch out each of these concerns in more detail.

Building a Self-Conception

For a wide variety of philosophers there seems to be an intuition that agents can be at least partly self-constituting. What it is to be an agent, for this broad camp, is not just to be a particular sort of mammal with advanced cognitive skills. To be an agent is, at least in part, to be able to decide who one wants to be in a very concrete sense. Am I an intellectual or a person of action? Shall I be the best I can be at one chosen pursuit, or aim for being “well-rounded”? Though these schematic sorts of questions obviously put the choices in starker terms than we often encounter in real life, they are the sorts of self-defining choices with which many of us are familiar.

Oftentimes, however, attempts to theorize this sort of self-constitution end up going to one of two extremes, each of which has its own implausibilities and costs. On the one hand, (to schematize somewhat roughly), we have someone like Sartre, who conceives of the individual as radically free, unconstrained by the past and always able to redefine himself as fully as he sees fit. This view, however, comes at a cost. It sacrifices the feeling that many of us have, (and which to some extent I think has been best explored by some of the feminist work on narrative integrity¹¹⁶), that we cannot conceive of agency as fully independent from one's own past and social context, and that without these frameworks, it is easy to lose grip of what counts as the “self” at all.

¹¹⁶ Including work by Cheshire Calhoun, Hilde Lindemann, and Margaret Urban Walker.

On the other hand, we have someone like Harry Frankfurt, who also conceives of the agent in a particularistic and concrete way, but who cashes that out in terms of the contingent psychology and will with which a given agent finds herself. Frankfurt's picture of the will pulls out the way in which agents can be defined by their actual objects of care and love, and not just by some bare choice-making faculty. At its limits, my will is defined by the things I cannot help caring about, the people and projects I simply find myself loving. But in tying who one is so closely to one's contingent psychology, I think a view like Frankfurt's ends up having to sacrifice important elements of activity.¹¹⁷ One's selfhood becomes importantly a matter of discovery, more than something to be created. Who I am is tied to what I fundamentally care about, but "care" in the sense of being unable to move myself otherwise. While it is clearly true that some parts of our lives are like this – in particular, I find Frankfurt's analysis of parental love to be spot on – this is not the only way in which agents can be a part of determining their own identity. At least sometimes, we seem able to *decide* who we ought to be. And personal commitments, I want to suggest, may provide one route to a middle way, helping us to work towards a sensible picture of agential self-constitution.

In our power of personal commitment, we can recognize a way in which understanding who an agent *is* requires looking to the ends on which she contingently, idiosyncratically decides to confer importance. In making a personal commitment, an

¹¹⁷ "...[I]t is a necessary feature of love that it is not under our direct and immediate voluntary control. What a person cares about, and how much he cares about it, may under certain conditions be up to him. (...) With regard to certain things, however, a person may discover that he cannot affect whether or how much he cares about them merely by his own decision" (*The Reasons of Love*, 44). As he adds, "The origins of normativity do not lie, then, either in the transient incitements of personal feeling and desire, or in the severely anonymous requirements of eternal reason. They lie in the contingent necessities of love" (48).

agent has staked herself on the value of this end; in a matter of speaking, we can say that she has incorporated that end and its normative demands into who she is. Again, I take this to point to be in a similar vein to those made by people like Sartre and Frankfurt. Who *I* am extends beyond the powers of thought and volition that I share with other agents to include characteristics that are particular to me, and beyond descriptive facts about me to include normative claims – what I am *bound* to consider or do.

But as I've said, I think our picture of personal commitment enables us to go beyond the two sorts of camps I've sketched out. On the one hand, our personal commitments bind us as a result of our *activity* – unlike the objects of Frankfurt's primary interest, personal commitments are something we freely adopt through our own deliberative choice. We have a creative role to play in our self-understanding. Whereas some could simply end up at vegetarianism as a result of their other values and decisions, Jason decides to confer this special value on that end. Yet, on the other hand, his choice is not radically free. We cannot simply make any old personal commitment we'd like, since a commitment must build upon at least a somewhat antecedently valuable end. But nor can we simply abandon or revise our personal commitments on a whim. At the least, I've claimed, revoking a personal commitment requires thoughtful reconsideration and some justification for doing so. Both getting in and out of a personal commitment require working with the values out there in the world, not just creating norms *ex nihilo*. Furthermore, even when I successfully commit myself or revoke a personal commitment, there can be normative residue if I do so in a less than ideal way, residue which doesn't show up on a picture like Sartre's. Annie may genuinely revoke her personal

commitment if she gets into a feud with the community theater director, but this will certainly reflect on her character, and likely require some sort of amends (though not necessarily to the director herself!).

My argument in favor of this picture has been relatively simple: skepticism about this picture of personal commitment comes at the cost of viewing agency as fundamentally asymmetrical and normativity as relatively limited. Whereas I can freely bind myself to others, often merely on the strength of my word, the claim would be that I cannot do the same on my own. Imagine what this would look like: if Annie wanted to commit herself to raising money for her local community theater, two courses of action would be open to her. One option would be to make this commitment to someone else: for example, Annie could turn to her friend Saul and try to pull off a promise to raise the money. Saul, however, might refuse to accept the promise, perhaps because he really doesn't care, or perhaps because he doesn't want the responsibility of having this normative hold over Annie. If Annie can't make a commitment to herself, then her only option will be to find a different friend to whom she can make a promise. If this second friend also refuses to hold her to her commitment, we can imagine Annie going door-to-door in her philosophy department, vainly searching for a promisee in order to commit herself to raising this money. And if we reject the possibility of personal commitment, Annie's only alternative to this somewhat absurd behavior is to put in place some causal constraints to force her to raise the needed funds, such as giving the community theater the legal right to garnish her wages should she fail to follow through. In rejecting a

power of personal commitment, we consign Annie to either the mercy of those around her, or else to the physical costs she can manage to impose upon herself.

The agency we are imagining in this scenario is somewhat silly, certainly impoverished and not, I think, the sort of agency that most of us take ourselves to have. Now, oftentimes I change my normative position through my relationships with others – I make a promise to my mother, accept a favor from a friend, sign a contract with the electric company. But if part of being an agent, (and indeed, I think it is a central part), is the ability to do more than just respond to the demands of norms – also affecting and interacting with those norms – then why should I only be able to exercise this ability by transacting with others? Again, to be clear, there are many ways in which we can define ourselves: indirectly, through built-up patterns of action, passively, through the feelings of care and importance we happen to have, and socially, through our interactions with and commitments to others.

The claim I have been pushing here is that only personal commitment can account for the occasions when we seem able to define ourselves directly, actively, and solely through our own power. If an essential part of agency is some sort of normative activity, it seems to be a radical revision to insist that this aspect of agency – normative self-definition – can only be exercised indirectly, through the help of other agents or jerrymandering of physical and psychological facts. As I have argued, mere psychological facts are insufficient to recapture the normative changes we seem able to effect, and it feels oddly asymmetrical to consign purely normative changes to the social

realm.¹¹⁸ In responding to the value in the world, we seem able to inflect and mutate the values we encounter, going from impersonally good to personally valuable, and valuable to required. Accepting that we have a power of personal commitment allows us to recapture a richer picture of agency and normativity, helping to illuminate the way in which agents are responsible to, and yet also build upon, the value they find in the world.

¹¹⁸ In addition to the other differences we have seen, I suspect that some of what makes personal commitment a distinctive means for self-constitution comes from what we might call their “dual nature” – coming from an act of my will, yet constraining my choice; particularly valuable because of my activity, but the possibility of my activity depending on their value – that these aspects place personal commitments at the intersection of an agent’s activity and her responsibility.

Works Cited

- Aquinas, Thomas. *Summa Theologica*. New York: Benziger Brothers, 1948, IIa 35.1.
- Audi, Robert and Patrick E. Murphy. The Many Faces of Integrity. *Business Ethics Quarterly*, January 2006, 16(1): 3-21.
- Austin, J.L. *How To Do Things With Words*, 2nd Edition. Ed. J.O. Urmson & M. Sbisá. Cambridge, Mass.: Harvard University Press, 1962.
- Bratman, Michael. Anchors for Deliberation. In: *Intentionality, Deliberation and Autonomy*. Ed. Christopher Lumer and Sandro Nannini. Burlington, Verm.: Ashgate Publishing Company, 2007.
- Bratman, Michael. *Structures of Agency: Essays*. Cambridge: Cambridge University Press, 2007.
- Bratman, Michael. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press, 1999.
- Bratman, Michael. *Intention, Plans, and Practical Reason*. Cambridge, Mass.: Harvard University Press, 1987.
- Bromberg, Philip. *Standing in the Spaces: Essays on Clinical Process, Trauma, and Dissociation*. Hillsdale, N.J.: Analytic Press, 1998.
- Broome, John. Normative Requirements. *Ratio*, 1999, 12: 398–419.
- Broome, John. Should We Cope with Incommensurable Values? In: *Practical Rationality and Preference*. Eds. Christopher W. Morris & Arthur Ripstein. Cambridge: Cambridge University Press, 2001.
- Calhoun, Cheshire. Standing for Something. *Journal of Philosophy*, May 1995, 105(5): 235-60.
- Chang, Ruth. Voluntarist Reasons and the Sources of Normativity. *Practical Reasons and Action*. Eds. David Sobel & Steven Wall. Cambridge: Cambridge University Press, 2009.
- Cicero. *On Duties*. Ed. E.M. Atkins. Trans. M.T. Griffin. Cambridge: Cambridge University Press, 1991.

- Cox, Damian, Marguerite La Caze, & Michael Levine. Integrity. *The Stanford Encyclopedia of Philosophy (Fall 2005 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2005/entries/integrity/>>.
- Cox, Damian, Marguerite La Caze, & Michael Levine. Should We Strive for Integrity? *Journal of Value Inquiry*, 1999, 33(4): 519-530.
- Davidson, Donald. Actions, Reasons, and Causes. *The Journal of Philosophy*, 7 Nov. 1963, 60(23): 685-700.
- Davis, Wayne. A Causal Theory of Intending. *American Philosophical Quarterly*, 1984, 21: 43-54.
- Den Hartogh, Govert. The Authority of Intention. *Ethics*, 2004, 115(1): 6-34.
- Descartes, René. *Discourse on the Method and Meditations on First Philosophy*. Trans. Elizabeth S. Haldane and G.R.T. Ross. New Haven, Conn.: Yale University Press, 1996
- Downie, R.S. Three Accounts of Promising. *The Philosophical Quarterly*, 1985, 35(140): 259-271.
- Elster, John. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press, 1979.
- Ferrero, Luca. *Making Up One's Self: Commitments, Agency, and Identity*. Unpublished dissertation.
- Frankfurt, Harry. *Taking Ourselves Seriously & Getting It Right*. Ed. Debra Satz. Stanford, Calif.: Stanford University Press, 2006.
- Frankfurt, Harry. *The Reasons of Love*. Princeton: Princeton University Press, 2004.
- Frankfurt, Harry. *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999.
- Frankfurt, Harry. The Faintest Passion. *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999.
- Frankfurt, Harry. *The Importance of What We Care About*. New York: Cambridge University Press, 1988.

- Gauthier, David. Commitment and Choice: An Essay on the Rationality of Plans. In: *Ethics, Rationality, and Economic Behaviour*. Eds. Francesco Farina, Frank Hahn, and Stefano Vannucci. Oxford: Clarendon Press, 1996.
- Gert, Joshua. Two Concepts of Rationality. *Southern Journal of Philosophy*, 41(3): 367-398.
- Habib, Allen. Promises to the Self. *The Canadian Journal of Philosophy*, forthcoming.
- Hart, H.L.A. Are There Any Natural Rights? *Philosophical Review*, 1955, 64(2): 175-191.
- Hill, Thomas E. *Autonomy and Self-Respect*. Cambridge: Cambridge University Press, 1991.
- Hobbes, Thomas. *Leviathan*. Ed. Richard Tuck. Cambridge: Cambridge University Press, 1996.
- Hume, David. *A Treatise of Human Nature*. Ed. L.A. Selby-Bigge. Oxford: Clarendon Press, 1978, 3.2.5.9-12.
- Jaworska, Agnieszka. Caring and Full Moral Standing. *Ethics*, 2007, 117(3): 460-497.
- Jaworska, Agnieszka. Caring and Internality. *Philosophy and Phenomenological Research*, 2007, 74(3): 529-568.
- Kant, Immanuel. *The Metaphysics of Morals*. Ed. & Trans. Mary Gregor. Cambridge: Cambridge University Press, 1996.
- Korsgaard, Christine. The Normativity of Instrumental Reason. *Ethics and Practical Reason*. Eds. Garrett Cullity & Berys Gaut. Oxford: Clarendon Press, 1997.
- Lindemann, Hilde. (as Hilde Lindemann Nelson). Context: Backward, Sideways, and Forward. *HEC Forum*, Special issue on narrative, 1999, 11(1): 16-26.
- McClennen, Edward. *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: University of Cambridge Press, 1990.
- McFall, Lynne. Integrity. *Ethics*, 1987, 98: 5-20.
- Mian, Emran. The Curious Case of Exclusionary Reasons. *Canadian Journal of Law and Jurisprudence*, January 2002, 15(1): 99-124.

- Mill, John Stuart. *A System of Logic*. New York: Harper & Brothers, 1871. Bk.5, Ch.7, §2.
- Owens, David. A Simple Theory of Promising. *The Philosophical Review*, 2006, 115(1): 51-77.
- Radden, Jennifer. *Divided Minds and Successive Selves: Ethical Issues in Disorders and Personality*. Cambridge, Mass: MIT Press, 1996.
- Rawls, John. *A Theory of Justice*. (Rev. ed.) Cambridge, Mass.: Harvard University Press, 1999.
- Raz, Joseph. *Value, Respect, and Attachment*. Cambridge: Cambridge University Press, 2001.
- Raz, Joseph. *Practical Reason and Norms*. Princeton, N.J : Princeton University Press, 1990.
- Raz, Joseph. *The Authority of Law: Essays on law and morality*. Oxford: Clarendon Press, 1979.
- Raz, Joseph. Voluntary Obligations and Normative Powers II. *Proceedings of the Aristotelian Society*, 1972, 46 (suppl.): 79-101.
- Rescorla, Michael. Convention. *The Stanford Encyclopedia of Philosophy*. (Fall 2007 Edition), Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/fall2007/entries/convention/>.
- Sartre, Jean-Paul. *Existentialism is a Humanism*. Trans. Carol Macomber. New Haven, Conn.: Yale University Press, 2007.
- Scanlon, T.M. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass.: Harvard University Press, 2008.
- Scanlon, T.M. Reasons: A Puzzling Duality? *Reasons and Value: Themes from the Moral Philosophy of Joseph Raz*. Eds. R. Jay Wallace, et al. New York: Oxford University Press, 2006.
- Scanlon, T.M. Promises and Practices. *Philosophy and Public Affairs*, Summer 1990, 19(3): 199-226.
- Shiffrin, Seana. Promising, Intimate Relationships, and Conventionalism. *Philosophical Review*, 2008, 117(4): 481-524.

- Singer, M.S. Duties to Oneself. *Ethics*, 1959, 69: 202-5.
- Smith, Adam. *The Theory of Moral Sentiments*. Eds. D.D. Raphael & A.L. Macfie. Oxford: Clarendon Press, 1976.
- Smith, Angela. On Being and Holding Responsible. *The Journal of Ethics*, 2007, 11: 465–484.
- Solomon, Robert C. *Ethics and Excellence*. New York: Oxford University Press, 1992.
- Stocker, Michael. Desiring the Bad – An Essay in Moral Psychology. *The Journal of Philosophy*, 1979, 76(12): 738-753.
- Strawson, P.F. Freedom and Resentment. *Proceedings of the British Academy*, 1962, 48:1-25.
- Taylor, Gabriele. *Pride, Shame, and Guilt: Emotions of Self Assessment*. New York: Oxford University Press, 1987.
- Velleman, David. Identification and Identity. In *Contours of Agency: Essays on Themes from Harry Frankfurt*. Eds. Sarah Buss & Lee Overton. Cambridge, Mass.: MIT Press, 2002.
- Velleman, J. David. *The Possibility of Practical Reason*. New York: Oxford University Press, 2000.
- Walker, Margaret Urban. Picking up Pieces: Lives, Stories, and Integrity. In: *Feminists Rethink the Self*. Ed. Diana Meyers. Boulder, Colo.: Westview, 1997.
- Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press, 1994.
- Watson, Gary. Free Agency. *Journal of Philosophy*, 1975, 72(8): 205-20.
- Wellman, Carl. *A Theory of Rights: Persons under laws, institutions and morals*. Totowa, N.J.: Rowman & Allanheld, 1985.
- Wenar, Leif, "Rights", *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), forthcoming URL = <http://plato.stanford.edu/archives/fall2008/entries/rights/>.
- Williams, Bernard. *Moral Luck: Philosophical Papers 1973–1980*. Cambridge: Cambridge University Press, 1981.

Williams, Bernard. Integrity. In: *Utilitarianism: For and Against*. Eds. J.J.C. Smart & Bernard Williams. New York: Cambridge University Press, 1973.

Williams, Bernard. Ethical Consistency. *Proceedings of the Aristotelian Society*, 1965, 39(supp): 103-124.

Wolf, Susan. The Meanings of Lives. In: *Introduction to Philosophy: Classical and Contemporary Readings*. Eds. John Perry, Michael Bratman, & John Martin Fischer. New York: Oxford University Press, 2007.