

Self-Creating Reasons: The Normative Implications of Identity

by

Erica Kathryn Lucast Stonestreet

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Philosophy)
in The University of Michigan
2008

Doctoral Committee:

Professor Stephen L. Darwall, Chair

Professor Elizabeth S. Anderson

Professor Scott J. Shapiro

Associate Professor Sarah Buss Paulson

© Erica Kathryn Lucast Stonestreet
2008

To Aaron,

without whose love and support I would be a very different person.

Acknowledgments

I am unendingly grateful to my committee and especially Steve Darwall for patient readings of and comments on both good and bad material during the creation of this thesis. Elizabeth Anderson suggested many of the literary examples and helped me to keep things realistically human; Sarah Buss was willing to join the committee just a few weeks after joining the Michigan faculty, reading the whole thesis in a week and spending several afternoons in patient discussion of my ideas; and Scott Shapiro agreed to be on the committee without having met me, and has also been a supportive interlocutor.

Peter Railton, David Plunkett and Steve Campbell were also patient conversants as my ideas developed; Vanessa Carbonell and Marie Jayasekera provided untiring moral support. I am also grateful to many people, too numerous to mention, whom I met during my job search and who provided many helpful challenges, suggestions and criticisms.

And finally, I'm grateful to my family: Karen, Don, Ellen and Rebecca Lucast, and Aaron Stonestreet. They have with equanimity listened to me complain, worry, moan and groan, sigh and struggle through the project from start to finish.

Contents

Dedication	ii
Acknowledgments	iii
Chapter 1: Introduction	1
Chapter 2: The Anti-DBR View and its Variations	10
Chapter 3: A Puzzle for the Anti-DBR View	45
Chapter 4: Care, Identity and Ideals: Breaking the Impasse	71
Chapter 5: Self-Creating Reasons: An Analysis of Normativity in Willing and Identity	118
Chapter 6: Conclusion	147
Appendix: Further Examples	158
References	164

Chapter 1: Introduction

What are we doing when we deliberate? To give a rather uninformative and formal answer: a practical deliberator is attempting to decide what to do. Some situation has arisen in which she has a choice to make; she could do this or that (or a third, fourth, fifth option...). Some decisions are fairly trivial, including what laundry detergent to buy or what to have for lunch. Some are more momentous or complex, such as where to buy a house or whether to send one's children to public or private school. Others are so important that they shape the course of one's life because the action one takes reflects and shapes the values which govern it; choices of careers and life partners are convenient common examples of such decisions. There is a range of seriousness involved in different instances of practical deliberation, but every deliberation arises when action has stopped because the way forward is no longer clear. The task of the deliberator is to find a way forward again.

So I ask again: What are we really doing when we deliberate? How does deliberation make the way forward clear again? Ask a deliberating agent and he will probably tell you that he is looking for reasons. We are reasoning creatures, and we like to be able to defend and justify our actions to one another and to ourselves. These are the kinds of motives that drive our agent to want to know which of his options is best supported by reasons. His deliberation, then, appears to be a project of discovering reasons, observing how much they “weigh” (or should weigh) in the decision, and understanding how they interact with one another to render a verdict concerning what he should do. He might even describe his project as “trying to figure out what (there is reason) to do”—as if he is approaching the problem with the idea that there is an answer out there and he is working to discover it. He just has to make sure he puts all the relevant reasons on the scales, with their proper weights, and then all he does is read off the balance.

Surely this is more or less what we do for many of the decisions we face on a common basis. We frequently make pro and con lists or decision analysis charts for

complex choices such as which new car to buy, which courses to take next semester, or which school to send our children to. Some decisions don't merit even this much thought, of course; in trying to decide whether to have french fries or a salad with my sandwich, I probably won't actually list pros and cons. But I will consider the facts about the taste and nutritional content of the options, my hunger, and my less immediate plans and goals (such as losing weight, satisfying a craving, or enjoying an occasional treat). I may decide that the salad is the better choice this time, but tell myself that next time fries will be allowed. Whether I actually act on this decision is another question, but I can at least say I've arrived at what I have most reason to do.

In this "common sense" picture of deliberation, reasons seem to be "out there." They are facts that count in favor of an action, to some extent fixed by the particulars of the circumstances (so the strength they have may vary depending on what other reasons are in play). It is not legitimate for the agent to put her thumb on the scales, so to speak. That is, the responsibly deliberating agent¹ is not allowed to influence the outcome of her deliberation by consciously giving more weight to some factor than the circumstances dictate. (If she does so unconsciously, then she may not be irresponsibly deliberating, but her thinking is nonetheless "incorrect.") Wanting the fries may sometimes mean I have an additional reason to eat them over the salad, because this is a fact about me in my circumstances that is relevant to the decision of what to eat for lunch; wanting them more may indicate that I would enjoy them more, and this is a reason to eat them. But wanting the fries does not mean that the fact that they taste good is more important, or should weigh more, than the fact that they have a great deal of fat and sodium in them. According to this picture of deliberation, there is no arbitrary or subjective element to the responsibly made decision, for reasons are what they are independently of what the agent may wish they would be. She does not create reasons; she seeks them out and uses them as tools to guide her action. After all, there would be no point to this weighing exercise if the agent could create her own reasons, because any decision she made would be arbitrary and indefensible from a

¹ I emphasize "responsibly deliberating" because it is clear that at least in low-stakes decisions, few of us are always fully responsible deliberators. We get impatient, we get emotional, we get lazy, and we do not do all the research and reflection we would ideally do in order to make good decisions. There need not be anything wrong with this; surely it can be more important to get on with one's life than it is to be certain to choose the best laundry detergent off the shelf, or to have a little variety rather than eat the most nutritious lunch available every day.

common or objective point of view. To tip the scale is not to use it at all, according to this view.

But note a common phenomenon of behavior with respect to our lists and charts: sometimes the deliberator finds that, after he has done his best to come up with an objective and defensible understanding of the situation, what they tell him is not to his satisfaction (and it is not always clear to him why not). Sometimes when this happens he follows the verdict anyway, but just as often he goes back and adjusts the list so that it better reflects what he “really” thinks. Is this tipping the scales, bending the reason “data” to fit an unjustified feeling? Not necessarily. We could explain such adjustments as repairing mistakes in the judgments the agent used to list his reasons or to give them weights relative to one another. The refinement of the chart or list is just a further phase of discovering the reasons the agent already has—that a shorter commute time is actually more important to him than having a big yard, for example, even though he initially thought it wasn’t. Deliberation is still decision analysis, or the discovery of reasons which purport to direct our actions.

Again, this seems right for most of our ordinary decisions. Many reasons can be understood as heavily dependent on our motivations, and we are certainly familiar with occasionally being unfamiliar with our own motivations. So it is not unreasonable for us to need to do some digging in order to understand our reasons and act accordingly.

Nevertheless, in what follows I wish first to fill out and then to challenge what I have here called the “common sense” view of deliberation. This view paints the dilemmas which prompt deliberation as something more like patches of dense jungle that obscure the path the agent should take rather than as forks in which it is genuinely up to the agent to determine, and not just to discover, which path to take. The agent is indeed in control of clearing the way, and may do so poorly and thus go wrong; but where he would be going were he deliberating well is always supposed to be defined independently of his deliberative process, according to the view I have been describing. Though adequate for the great majority of the mundane decisions we make in our lives, in this dissertation I will be arguing that this picture of practical reason, and the conception of reasons that goes with it, leaves too large a gap between the discovery of reasons and what we have (most) reason to do. The picture does not say enough about what brings us to make our decisions the way we do in certain important practical conflicts.

Briefly, I will argue that it *can* be legitimate for the agent to put a thumb on the scales.² I build on the intuition that one's decisions simultaneously reflect *and* shape who she is, stemming as they do from the experiences of her past while also building the habits of thought and action that become part of her character.³ In slogan form, my claim is that “what to do” is sometimes a question of “who one wants to be,” and when it is, the agent may have a freedom to choose that the view under discussion does not allow for. Now, it is important to note here that not all cases in which the practical question amounts to “who to be” will be cases in which who to be is up to the agent. Sometimes, for example, there will be moral reasons that require a certain character on the part of the agent. Jean Valjean's choice of whether to confess that he is prisoner 24601, thus saving an innocent man's life, or whether to remain silent and continue as mayor and benevolent factory owner, is arguably such a choice. He sees that his decision is one of character, asking himself “Who am I?” He is tempted by the chance to bury his past once and for all, but decides that he is—and must be (“How can I ever face myself again?”)—the responsible man who owns his previous actions even at great cost to the town and his workers. Thus Valjean does not have the freedom I am emphasizing here.

On the view I will be articulating in this thesis, then, there are what I will call “self-creating” reasons, which (unlike “ordinary” reasons) are not independent of the deliberative process. They are self-creating both in the sense of creating an agent's self *and* of being created out of elements that make up the self. Thus they represent a genuine freedom of *reasoned* choice. They are reasons related to the agent's ideals and are future-directed, as they justify by aiming at who the agent is to be or become—what sort of self she is creating in making her decision. Self-creating reasons are thus closely related to commitment and Frankfurtian care or love, and they allow a reasoned justification of the choice she makes when it seems as though reasons (in the sense given by the “common sense” view) have run out. Deliberation involving self-creating reasons is not simply discovery, but invention as well.

² This is an idea Nozick takes seriously, but does not provide adequate argument for, in his *Philosophical Explanations* (1981, p. 294ff).

³ This is a Deweyan point. See *Human Nature and Conduct* for his discussion of the way choices affect habit formation and thus the formation of the self.

The idea is easiest to illustrate with an example. Consider the case of John Smith in the film *The New World*.⁴ Smith is among the party who first land at Jamestown in 1607. In the film's version of the story, Smith is known for his courage and capability to lead, but has an independent streak that makes him chafe under other leadership. He is ambitious and perhaps somewhat jealous that he was not made the leader of the Jamestown party. Nevertheless, he is open to new experiences and develops friendly relations with the native people, coming to understand and learn from them better than any of the other settlers. In his extensive interactions with the tribe, Smith falls in love with the chieftain's daughter Pocahontas. But the settlement's supply ships return in 1608 with the news that the king and queen have commissioned Smith to lead a new expedition in search of the Northwest Passage farther north up the coast of North America. Smith is now torn between his love for Pocahontas and his ambition to go down in history as the discoverer of the route west to India, and the film makes clear just how agonizing this choice is for him. It is a choice he cannot ultimately share with anyone, for the one who is closest to him is the one with the deepest stake in his decision. The very fact of having to decide puts a distance between him and Pocahontas, because this decision must be wholly his even though she is entitled to be consulted. He must bear the burden of the choice entirely on his own.

Smith's dilemma is a conflict between two deep commitments. He chooses to lead the discovery expedition, but it is not obvious that another man in Smith's position would have chosen the same way—and more to the point, it is not obvious that another man in Smith's position would have had *reason* to choose the same way. The reasons on both sides are quite strong; either choice involves the sacrifice of something deeply important. Not only this, but in each case he has a chance to do some good that goes beyond his personal reasons: in staying with Pocahontas, he might work as an ambassador between colonists and natives, enhancing the lives of both through cooperation; and in leading the expedition west he might contribute a great deal to the wealth and knowledge of his country. Let us assume, then, that there is no fact of the matter as to what Smith should do as a matter of reason, at least as the view sketched here would see things. No one could really fault him for either choice.

⁴ Malick (2006). I am not in a position to judge the historical accuracy of this rendering of the tale of Pocahontas, Smith, and Rolfe, but for my purposes historical accuracy is irrelevant.

Now, one might say that of course he has reason to choose as he does. Anyone would, in his position, since each option is well-supported. This is true, but my point is that these reasons, strong as they are, are not enough for someone in Smith's position to make the choice *between* his options. A chance to fulfill his life's ambition and make an important contribution to the world is certainly a reason to accept the king's request, all else held equal. But it is not a reason to choose to go *over* remaining in Jamestown with Pocahontas in these circumstances.

Here is why. Given his opposed commitments, either choice Smith makes would leave its mark on his life; either choice would shape him deeply in a new way, at least from his own internal perspective (if not also from a third person's). Either way he decides, he will be affirming one commitment and loosening another, and (on my interpretation) there is nothing in the situation that breaks the impasse between them. Yet too much is at stake for Smith's choice to be arbitrary. So he must choose for a reason, and he does: he chooses on the basis of who he wants to be—to himself, if no one else; what matters here is that he is not making a judgment on the basis of something external to him. He is not conforming his life to a pattern given by someone or something else. His decision thus expresses a powerful sort of autonomy. In leading the expedition, Smith reinvents himself as a man who must make the effort to fulfill an ambition (and not only his ambition but that of an entire nation), even if it means leaving the woman he loves (and the good he could do as an ambassador between colonists and natives). In the very act of choosing he reaffirms in a decidedly concrete way his already-standing commitment to exploration and leadership, and in doing so necessarily (under the circumstances, which are certainly cause for deep regret) lessens his commitment to his lover. His commitments, and as a result his character and his reasons, are different now.

It is important to differentiate the kind of choice Smith makes, with its deliberate basis in shaping character, from other choices which have character-shaping consequences. A worker may make a choice to embezzle funds from her employer, and this may well have deep, long-term consequences for who she is (especially if she is caught and goes to prison for it). But most of the time a choice to embezzle will be made for the sake of things like greed, frustration or revenge, or out of boredom, just to show that she could do it. In such cases, the worker is not making her choice *on the basis of* who she will be. This is not to say that she *could* not make the choice this way; perhaps she embezzles because she wants to be

the sort of person who doesn't conform to "the system." Nevertheless, the point is that there are plenty of decisions that shape character, but the kind that interest me here are made *on the basis of* "who to be" in such a way that this gives the agent important reasons that the understanding of reasons as independent of one's will overlooks.

By making the claim that in deciding what to do Smith faces a question of "who he wants to be," I intend to argue that his decision to give one commitment a priority over another is a decision that shapes his reasons in a fundamental and *subjective* way. He is, in a sense, choosing what his reasons are.⁵ This is different from each of three other possibilities. First, there is a consequentialist interpretation: Smith could simply examine the circumstances and see that choosing to head the expedition makes him the sort of person whose priorities are on courage, glory, and duty to country, and this is the best consequence (however "best" is cashed out), so it is what he should do. Second, there is a more virtue-ethical interpretation under which Smith is asking what a good, virtuous person would do, and choosing his commitments on the basis of this. Third, he could take a more contractarian stance, asking what someone in his particular position, with his particular relationships and abilities, would have most reason to do.

All three of these possibilities are rooted in objective circumstances, however, and they do not leave Smith the room to shape his reasons that I claim he has. The problem with claiming this is that it starts to look like choices made on these bases, when the reasons *do* dictate, do not stem from the powerful autonomy that seems to be present in Smith's case. Were he choosing as any of the above three interpretations suggest, Smith would be attempting to conform his actions to an external standard whose source is beyond his control. His reasons, and their weights, would—*all*—be given to or imposed on him. Note that "objective" as I'm using it here is not restricted to agent-neutral values and reasons; it is possible to give each of the three interpretations I'm calling "objective" a reading on which it

⁵ It will not usually be the case that an agent will have the appropriate power to shape the facts that give her reasons, so most of the time, the freedom she has is freedom to weigh the reasons—that is, to give them weights. Furthermore, there may be some perspectives from which a certain fact is a reason, when it is not a reason from other perspectives. Insofar as the freedom I want to defend is a freedom to choose one's perspective on a situation, the agent may have some control over whether some particular fact is or is not a reason for her.

takes account of agent-relative reasons. But the plain existence of agent-relative reasons isn't enough to take account of the idea that Smith can shape his reasons.⁶

Another way of putting my claim is that reasons, for an agent or from an agent's point of view, are not fully given by will-independent considerations.⁷ Smith is not thinking as a consequentialist, nor as an aspiring virtuous person, nor indeed from any objective point of view from which he sees himself as aiming to match an external standard, because from any of these points of view he must view his reasons as completely determined. Yet, so viewed, he finds himself at an impasse because the reasons as he knows them up to now leave him with two permissible options that are mutually exclusive. Given that he must decide, and decide non-arbitrarily (given the importance of the choice), he must give more weight to one set of considerations rather than another, thus tipping a balance in the *prima facie* weights of his reasons.

My project, then, is to give an account of what can be going on here. I claim that an agent may choose some reasons over others, on the basis of who she will become in making the choice. I am not claiming that an agent can simply choose her reasons "from scratch," as one might say. The kinds of dilemmas which trigger questions of self-constitution arise precisely *because* the agent already has the reasons that come into conflict, and she is committed to these reasons because of her commitments to their sources. These sources *can* certainly be objective circumstances, over which the agent has little or no control.⁸ Furthermore, as I have said, I claim that the choice is not simply arbitrary, because a practical reasoner cannot fully "stand behind" her decision unless she can give a reason for making it as she does. And given that these dilemmas arise because of, and result in a new understanding of, values to which the agent is committed in a fundamental way, she must be

⁶ Certainly, in choosing what friendships to have, for example, an agent could have some control over her agent-relative reasons; in choosing a friend she would, for example, be choosing to whom to give her loyalty. But choosing friends in this way (to the extent that we ever do this) is not the same as shaping reasons in the sense I'm after. It is possible, of course, to cook up scenarios upon which the reason-shaping choice *is* a choice of a friendship, but in ordinary circumstances, these two choices will not automatically coincide.

⁷ This is not the psychological claim that the reasons the agent *thinks* he has, or the ones he *takes* to be reasons (where he can be mistaken about this), are not given by objective considerations. Instead, the claim is a normative one, namely that the (weights of) reasons themselves are not entirely given by objective considerations.

⁸ That is, we certainly want to be able to say that the fact that the woman is your wife is a reason to save her; similarly, the fact that she is a person—never mind her particular relation to you—is a reason to save her. These circumstances give you reasons whether you like it or not; I am not contending that you have control over whether objective circumstances give you reasons. Rather, the claim is that you may have some freedom to decide just how strong the reasons they give you are.

able to stand behind her decision. So I will give an account of the self-creating reasons an agent may cite for choosing to privilege one commitment over another in light of the fact that many values do come with (*pro tanto*) weights.

The account I want to give must, then, explain two concurrent aspects of an agent's values in such cases: that her deliberation discovers them, on the one hand, and invents them, on the other. The discovery aspect consists in several facts: she *finds* her reasons to conflict in some situations; the options open to her are then constrained by the reasons she already has, which come from internal commitments and external circumstances; the deliberation on which she proceeds will begin by examining and elucidating those reasons in order to determine whether they do, after all, provide a clear guide to action. This process may lead her to think she is discovering what she thought all along. Yet her elucidation may not leave the reasons undisturbed; elucidation may involve some interpretation, and there may be freedom to interpret. To the extent to which this is so, the agent will be shaping her reasons. Furthermore, supposing that either option is rationally permissible from an external point of view and that she must make the choice for some reason, whatever allows her to decide will shape her fundamental commitments and thus the course of her life. It is, in this sense, a self-creating reason.

I begin by spending time elaborating what I will call the “anti-desire-based-reasons” view of reasons and practical deliberation, which I take to be an extension of what I started by calling the common sense view. This discussion will be followed by an exploration of complications of the standard view, which serve as amendments that better accommodate the way practical reasoning feels from the inside. I will go on, however, to show that even these variations do not adequately handle the problems presented by conflicts such as John Smith's, as I will illustrate with a number of case studies intended to bring out just what it is for a choice to be self-creating, and how this relates to shaping one's life around an ideal. With all of this background I will finally be ready to lay out the idea of self-creating reasons, elucidation and defense of which will modify our understanding of practical reason.

Chapter 2: The Anti-DBR View and its Variations

A venerable tradition in philosophy upholds three claims about the demands of practical reason: that reason can settle all questions about what to choose; that it requires the global maximization of value; and that the grounds for rational choice must be fully and decisively articulable, leaving no room for judgment and hence none for dispute.

—Elizabeth Anderson¹

[W]e seem to think moral questions have correct answers; that the correct answers are made correct by objective moral facts; that moral facts are wholly determined by circumstances; and that, by engaging in moral conversation and argument, we can discover what these objective moral facts determined by the circumstances are.

—Michael Smith²

In this chapter I explicate more thoroughly the view of reasons described in the introduction, laying out a philosophical view of practical reason and practical reasons that I will call the “anti-desire-based-reasons” picture, and which I take to be basically correct. This picture, then, serves as the basic framework from which my view represents a departure. In taking the picture presented in this chapter as a starting point, I do not wish to suggest it is uncontroversial or even the most commonly held view in the field, though I aim to present it in such a way that it is acceptable to a wide variety of philosophers who hold views on the matter.

¹ *Value in Ethics and Economics* (1993), p. 45.

² *The Moral Problem* (1994), p.6.

2.1 *The anti-DBR view*

2.1.1 *The basic idea and the desire-based reasons thesis*

What is a practical reason?³ Practical reasons are the sorts of considerations that can guide actions (and, sometimes, attitudes⁴). Agents typically respond to reasons without deliberating or, at times, even thinking. That there is a stop sign at the upcoming intersection is a reason for me to put my foot on the brake, and I do it without really thinking about it. That the bear is threatening me is a reason to be afraid, but I don't need to do any reflection before I get scared. Responding to reasons is an everyday activity, then, because we operate with a set of goals and beliefs that often do not need to come up for review. Reasons become much more salient when we reach decision points at which we must think deliberately about what to do next or review the ends we have set for ourselves.

Standardly, reasons are stated in the form of propositions describing facts. Thus, “that it is raining” is a reason for me to take an umbrella to work; “that my son is in the burning house” is a reason for me to fear for his life, and to run into the flames to rescue him; “that she would enjoy the trip” is a reason to help my sister buy a plane ticket. These reasons are naturalistically describable, but in ordinary parlance not all reasons are uncontroversially naturalistic in this way. Ordinarily, we also take considerations like the following to be reasons for acting: “that it is the right thing to do” is a reason to admit that it was you who broke your mother's vase; “that it is kind” is a reason to give to charity; “that the painting is beautiful” is a reason to admire it; “that I promised” is a reason for me to do what I promised. It is a matter of philosophical controversy just how to understand claims like these. Some philosophers hold what has been called a “buck-passing” view, according to which terms such as “wrong,” “good,” and “beautiful” get their normative force from reasons for attitudes such as guilt, anger, approbation, and admiration—and reasons for

³ For ease of exposition, I will usually drop the “practical” label, assuming it to be implicit unless the context calls for clarification.

⁴ There are different ways of understanding the connection between reasons for attitudes and reasons for actions. (Anderson's pluralist-expressive theory of practical reason (1993) is one that focuses directly on explicating this connection, but there are many—including Gibbard (1990), Darwall (2006), and Scanlon (1998)—who take some connection between reasons for attitudes (such as guilt) and actions (such as making amends) for granted.) The connections are complex, and what action is called for when a particular attitude is warranted depends on a number of contextual factors. Nevertheless, sometimes an attitude is warranted but this says nothing about whether or what action is warranted. Though fascinating, these issues are a side track from my present purposes, and so for simplicity I will generally neglect reasons for attitudes, despite the fact that they are, to some extent, practical.

these attitudes are generally naturalistic. Others hold that there are primitive normative facts and properties involving goodness, rightness, beauty, to-be-promotedness, etc., which have their normativity “built into” them, constituting a further, non-naturalistic reason for action over and above any reasons for attitudes. While I tend to favor a view of the former sort, I cannot claim that it is standard, so I will leave open the possibility that properties such as goodness and beauty provide reasons for action in and of themselves. Meanwhile, there is a bit more to be said here about the relationship between naturalistic facts and reasons.

Because the notion of a reason is at bottom a normative one, a naturalistic fact isn’t a reason for anyone to do anything until it stands in some relation to the agent and the situation in which the agent finds himself; naturalistic facts that record or describe states of the world are not *in themselves* normative. A naturalistic fact acquires normative significance when it can play a role in an agent’s deliberations⁵ and actions, and this is true regardless of what we take “normative significance” to consist in (e.g. whether or not it is tied to motivation). That the cup is on the table has no significance until, say, someone wants to show it to a friend, is thirsty and needs a vessel in which to pour some lemonade, is washing dishes, etc. Then it is a reason for him to, say, go to the table to get the cup; and the fact *that it is a reason* is a normative fact, though the simple fact of its sitting on the table is not normative. This of course still leaves open the question of just what normativity consists in and whether it is real—that is, nothing I have said speaks directly to the “problem of the normative.” This is a difficult and gripping problem, but one I cannot tackle here. For purposes of this project, therefore, I take normativity for granted.⁶

With all of this in mind, we might give a basic definition of a reason as follows: a fact (in the form of a proposition) is a reason to perform an action if and only if it “counts in favor”—whatever this ultimately consists in—of performing it.⁷ This simplistic way of describing reasons is fairly obvious, but further understanding of reasons gets immediately complicated when we go on to ask: “Counts in favor *how?*” Some, like T.M. Scanlon or

⁵ This construes “deliberation” broadly to include reflection on the right way to feel, where this is expected to guide how one actually feels.

⁶ For a thorough discussion of the problem of the normative, see Jean Hampton’s *The Authority of Reason* (1998), especially chapters 2-3.

⁷ Scanlon uses this specific formulation (1998, p. 17), as does Raz (1990, p. 186), and Parfit (1997, p.121) recognizes it as common. Of course, there is also the negative version: something is a reason *not* to perform an action iff it “counts against” performing it. For ease of exposition, I speak only of the positive case, with the understanding that this is meant to include negative cases as well.

Derek Parfit,⁸ claim that there is no non-circular answer to this question, because the normative relation reasons represent is primitive and does not depend on facts about the agent's motivation or capability of motivation. Others, such as Michael Smith⁹ or Stephen Darwall,¹⁰ tell a (usually complex) story about the connection between the agent's possible motivations and the facts of the situation that together give the designated proposition its status as a reason. A great deal of ink has been spilled in the debate between these camps (externalists and internalists, respectively), and these few sentences do not begin to do justice to the issue. For purposes of this project, however, I wish to remain neutral in the debate and will say no more here, except that the position I will eventually be taking lends itself most readily to internalism.¹¹

It is often said that reasons have both an explanatory and a normative capacity.¹² These two roles have a complicated relationship. Here I focus primarily on reasons' normativity, because my concern is with the way deliberation looks from the deliberator's point of view, not from a third-person perspective. A third-person perspective allows the separation of two questions: "Why did (or will) the agent do *x*?" and "Were (are) there good reasons for the agent to do *x*?" The first seeks to make the action intelligible, asking about motivations and causes, whereas the second is concerned with normativity and justification. For the most part, these questions collapse within the point of view of the agent making the decision; given that she is not just asking about what will happen in the future, asking "What shall I do?" amounts to "What *should* I do—what is there good reason to do?" These questions collapse because in asking "What shall I do?," she expresses an interest in finding good reasons, so that when she finds them, she is motivated to act on them. Then at least part of her explanation for her behavior will often be that there were reasons to act in that way. Or, if it turns out that there are no good (or good enough) reasons after all, we often explain the agent's mistake by observing how the agent *could* have seen what she chose as required, justified or intelligible from the perspective she had at the time; we assume the agent was motivated by perceived justification, for even when mistaken the questions of

⁸ Parfit (1997), (2001).

⁹ Smith (1994).

¹⁰ Darwall (1983).

¹¹ Despite my occasional externalist intuitions.

¹² See, e.g., Raz (1990), p. 15ff. See also Broome (2000) and Dancy (2003, 2000b).

explanation and justification coincide from the point of view of the deliberating agent. So the two aspects of reasons are certainly related, and one might argue that the normative sense is more important because of its role in explanation.¹³ Because of this, and since I am concerned with the perspective of the deliberator, I concentrate on the normative questions.

Now, because reasons do play a role in explaining action from a third-person point of view, and because the question of what reasons there are in support of an action most often arises when an agent desires to do something, it can be easy to take the desire itself as favoring action in the normative way a reason does. That is, on the surface common sense seems to suggest that desires¹⁴ are reasons for action, as we sometimes say “he (or I) just wanted to” as a way of explaining or justifying someone’s action. Hume articulated this way of thinking when he argued that reason can never discover moral facts, because moral facts are motivating in a way natural facts cannot be and reason does not motivate. Since sentiment is what motivates (and desire is one aspect or result of sentiment), moral facts are grounded not in nature but in sentiment. And indeed, the “Humean” view that all reasons are in some way based on desires was philosophical orthodoxy for some time (notably in Davidson’s “Actions, Reasons and Causes” and the papers that developed the themes there).¹⁵ Nevertheless, some reflection can show why this is not actually the case. In fact, in recent years it has become common to think that good reasons for someone to do something typically do *not* include that person’s desires.¹⁶

Internalists and externalists alike have argued strenuously against the Humean/Davidsonian thesis, which Darwall labels the “desire-based reasons” (DBR)

¹³ See Smith (1994, ch. 5) for a more detailed discussion.

¹⁴ The word ‘desire’ can be used to describe a range of phenomena. For instance, there is a sense in which anything I do consciously is something I “desire” to do (see Darwall (1983, p. 26)); in this sense ‘desire’ describes motivation quite broadly. I use the term much more narrowly than this, following a common philosophical usage (e.g. Scanlon (1998, pp. 37ff), Parfit (1997)) according to which a desire is a mental state that directs our attention to its object in a “conative” rather than “cognitive” way. Desires in this sense can be either “motivated” or not (to borrow Thomas Nagel’s terminology), that is, some desires arise as a result of the agent’s seeing reasons to have them; others arise first and draw the agent’s attention to their object and the reasons that favor it, and she might subsequently seek reasons supporting or defeating that desire.

¹⁵ Davidson (2001). Davidson’s view has been the basic foil from which Stephen Darwall, Gary Watson, Harry Frankfurt, and David Velleman have consciously departed.

¹⁶ And, indeed, to think it difficult to understand how the fact that someone desires something can ever be a normative reason for her to do it. (See Dancy (2000a) and Darwall’s “Because I Want It” (2001).)

thesis.¹⁷ Since this view has had such a wide following, I wish to spend some time surveying convincing arguments against it. First of all, one argument begins from the observation that agents sometimes do things they have “no desire” to do because they perceive the reasons to do them and are motivated by these.¹⁸ Desires certainly include a motivational element, but it has been argued (Hume to the contrary) that they are not the only sources of motivation, and we have reason to think they are not even the primary sources, for desires are to some extent responsive to reasons.¹⁹ Darwall gives a helpful example²⁰: Roberta has led a sheltered life and until going to college, she has thought of the world as a basically congenial place. In college, however, she sees a film documenting the hardships of working in the textile mills of the South and is moved to do what she can to alleviate the plight of the workers. Darwall’s point is that we can easily conceive Roberta as having no desire, up until she sees the film, to help the workers; and not only that, but he has set up the example in such a way that we can plausibly think that Roberta also had no particular latent dispositions to do so, innocent as she was. The example then illustrates a case in which an agent was moved by something other than a desire or other motivational disposition. We can be moved by our perception of reasons independently of our desires.

Of course, this motivational point only goes as far as to show that desires are not the *only* reasons for action. More needs to be done to show that they are not reasons at all. Arguments for this point have been given in detail elsewhere,²¹ so I will give only a sketch here. The main reason for thinking that desires are not in themselves reasons is that, unlike true reasons, desires do not *justify* an agent’s actions even if they do explain them.²² A desire draws one’s attention to an object and presents certain facts as reasons to choose it. These

¹⁷ See Scanlon (1998), ch. 1, sections 8-10; Darwall (1983), ch. 2-7; Dancy (2000a), ch. 2; Smith (1994), ch 5; and especially important was Nagel’s discussion of motivated and unmotivated desires in *The Possibility of Altruism* (1970).

¹⁸ This is the observation of Kant’s known as the “Fact of Reason” in the *Critique of Practical Reason* (5:30).

¹⁹ Parfit gives a quick sketch of an argument along these lines in (1997, p. 105).

²⁰ *Impartial Reason* (1983), pp. 39-40.

²¹ See Darwall (1983); Scanlon (1998, ch. 1.9); Dancy (2000a).

²² This is true even for internalists. Internalists hold that a consideration must be capable of moving the agent in order for it to count as a reason for her to act, and this is a necessary condition on its being a reason—i.e. a necessary condition on there being a justification at all. But this is not enough for a desire to actually *justify* an action. See Dancy, *Practical Reality* (2000a), Ch. 2.1 for a discussion of this point.

facts do not (generally) consist in the desire.²³ Furthermore, desire may present its object as a reasonable choice, but agents can step back from desires and review them, asking whether they *are* reasonable, i.e. whether there really is reason to do as a given desire directs. Under this scrutiny, the desire can play no role in justifying itself without circularity. Even if an agent reliably judges that she is justified in doing as her desires direct her to, i.e. even if desires are a reliable epistemic indicator for her of the “counting-in-favor” that constitutes reasons’ normativity, the desire itself is not doing any normative work.²⁴ In short, desires themselves are subject to normativity, so it is difficult to understand how they can be normative.

A second reason for thinking that desires are not reasons is that they are “transparent” to their objects. As they deliberate, agents seldom (if ever) have in mind their desire as such; instead, they are concentrating on the object of the desire (such as that shiny new computer, the delicious piece of carrot cake, or being free of the burning building). They seek reasons—most commonly properties of the object—that can guide action and that are separate from this desire. It may be true that an agent has reason to satisfy a desire because of the pleasure or relief it will bring to do so; but then it is the pleasure or relief that serves as the reason, not the desire itself. “That my action will satisfy my desire” is rarely (if ever) the cited reason for the agent to take action, from her own point of view; as Darwall observes, this fact is apt to be “beside the point or, at best, ...a trivial bonus.”²⁵ Scanlon concurs: he discusses doing something “just because you feel like it” as “special, rather trivial cases” or cases in which you act for no reason at all.²⁶ These observations suggest that even when an agent does act on a desire, it is the non-desire-based facts—perceived as reasons—that are doing the motivational work; the desire’s role is merely to draw the agent’s attention to these facts and the object of the desire.

²³ This is the “backgrounding” point that Pettit and Smith make (1990).

²⁴ See Darwall’s “Because I Want It” (2001) for a thorough argument for this point.

²⁵ (2001, p. 129). As I said, if the agent thinks that desire satisfaction is something to aim at, the root of his motivation is not the desire itself, but the satisfaction or relief that comes from having satisfied it. The reason for his action is not “that it will satisfy my desire” but “that satisfying my desire will give me pleasure [or relief].”

²⁶ (1998), p. 48; see also pp. 69-70.

2.1.2 *Universality, requiring power, and weighing*

I take the arguments against the “Humean” desire-based reasons thesis to be convincing. Having surveyed them, then, I turn now to further details of the anti-desire-based-reason (anti-DBR) view. As I have said, simply insofar as he is deliberating, a deliberator seeks reasons to guide his choice; the point of deliberation is to arrive at a justified decision in a non-arbitrary way, for arbitrariness undermines the point of deliberating in the first place. The guidance he seeks must therefore have some normativity, some staying power, for without this it is difficult to see how the choice is *justified* rather than explainable, random, or merely rationalized (in the sense of making excuses, or providing considerations that others (and oneself) might be persuaded to regard as reasons). What can contribute to a reason’s normative “staying power”? Two things come readily to mind: one is a kind of universality, and a second is requirement, i.e., that reasons that count in favor of some action in some sense *require* that action on the part of the agent.

When I claim that reasons must be in some sense universal, I do not mean that if r is a reason for me to a , then it is a reason for anyone else to a , or even that it is a reason for anyone else to promote *my* a -ing. Although there certainly are reasons that fit these descriptions (that a landscape is beautiful is a reason for anyone to preserve it; that Julie is doing charitable work is a reason for anyone who can to give her help or money), there are too many easy counterexamples for these agent-neutral reasons to be the *only* kinds of reasons there are. For instance, suppose my child is drowning. That she is my child is a reason (not necessarily decisive) for me to jump into the swollen river to save her, but it does not have the same force for you, a concerned friend who has my welfare in mind. While the fact that she is a child in need of help provides some (agent-neutral) reason for you to join or assist me, it is not as powerful a reason for you as the fact that she is my child is for me. This becomes clear when your child is also in the swift water; in that case, you clearly have more reason to rescue your child than mine, as I have more reason to rescue mine than yours (all else held equal). In light of examples like this, we must understand the universality of reasons to be given by a formula more like “if r is a reason for me to a , then it is a reason for anyone in relevantly similar circumstances to a .”²⁷ In the example just given, then, the

²⁷ As a particularist, Jonathan Dancy (2004) denies the claim that if some consideration has weight in one case, it must have the same weight in all cases. As I understand Dancy’s particularism, however, this is consistent with my universality claim, because my version of universality hinges on “relevantly similar” circumstances; where Dancy would claim that a consideration has a different weight, I can claim that the circumstances are

fact that it is my child is a reason for me to jump into the river partly because it *would be* a reason for you to save her *if* you were in my position, despite the fact that it does not have the same force for you as things actually stand. People may have different reasons according to their positions in a situation (which are given by relationships, preferences, motivations, and so on), but the important point is that the reasons they have do not depend on their being the specific individuals they are. Anyone else similarly situated in the relevant ways (having the relevantly same relationships, preferences, motivations, etc.) would have the same reasons.²⁸

A reason must be universal in this sense because if it were not, it would be impossible for it to be action-guiding in the way a deliberator requires. For, as I have pointed out, the point of deliberation is to decide what to do in a non-arbitrary, justified way. And without the situational universality outlined above, that is, if anything hangs on the agent's being the distinct individual he is, this sort of justification is impossible.

A second quality that reasons seem to need in order to serve in their normative capacity is requiring power.²⁹ Reasons need to have some claim on the agent; they seem to be tied to claims about what an agent "ought" to do. I mean this here in a rather simple sense: an agent who ignores or in some other way fails to conform to a reason makes a mistake, and it is a mistake of a rational sort, though it may not be egregious enough to warrant the epithet of *irrational*.³⁰ Reasons are supposed to tell us what to do.

This requirement thesis can have varying strengths. For example, if I am deciding whether or not to *a*, and I determine that I have some reason to *a* and no reason not to, then I would be committing some error not to *a*; and in that case, I am rationally "required" to *a*. Obviously, it doesn't make a lot of sense not to duck when a rock is thrown at your head, if

therefore not relevantly similar. In any case, my purpose here is to explicate a very common conception of reasons, and if particularism denies this kind of universality, it represents a significant departure from this "standard" position.

²⁸ This is the position Darwall endorses in *Impartial Reason* (1983); I take it that Darwall accepts this version of universality because he ultimately claims that the judgment of what he calls an "internally self-identified subject of a rational normative system" that *p* is a reason for her *just is* a judgment that *p* is a principle on which all agents ought rationally to act. Scanlon also accepts this version of universality: "We must be...concerned [with the reasons of others], insofar as we take ourselves to have any reasons at all, since any judgment about our own reasons entails claims about the reasons that others have or would have in certain circumstances" (1998, p. 73).

²⁹ My exposition in this paragraph is drawn partly from Gert (2004, especially p. 85ff).

³⁰ Note that this is a different sort of mistake than violating, say, a moral or prudential requirement. Below I will discuss ways of understanding requirement in much more detail.

there is no reason not to. One could make a stronger claim, however, by observing that ordinarily, a deliberator's concern is not merely which of her options is good enough—explainable, excusable, or supported by *some* reasons—but which of them is “best”—*most* reasonable, most justified. A family might be evaluating vacation spots on the basis of factors like expense, weather, available activities, etc. and conclude that on the whole they would most enjoy a camping trip in the Boundary Waters. It would strike us as inconsistent if they then made arrangements to go to Disneyland. One might thus claim that the thing to do is by definition the thing most supported by reasons, and that from this it follows that the agent has gone wrong in some way if she does not then perform that action. This claim entails the former, but not vice versa; in this sense it is stronger. I take it as part of the anti-DBR view, then, that reasons are requiring in some sense, though different formulations of requirement may be stronger than others.³¹

Now, it can happen that more than one available action (and hence none at all) is “most” supported by reasons. That is, sometimes there is as much reason to choose one option as to choose the other, and no reason to choose one option *over* another. This is the realm of rational permissions. In such cases, according to the anti-DBR view, either action is rationally permitted, and the agent has no recourse but simply to pick among his options arbitrarily, lest he be rendered immobile as Buridan's ass.³² If forced to explain a choice between vanilla ice cream and strawberry, or between going to one friend's party rather than another's, the agent for whom the reasons “tied” may be unable to say more than “I had to go with *one* of them, so I just picked the strawberry. There's no rhyme or reason to it, and next time I may very well choose the other way.” Arbitrary picking may also be forced when alternatives appear “incommensurable” or “incomparable,” but because incomparability is a complicated and controversial subject it will take us too far afield to discuss it here.

One thing still needed to complete the anti-DBR picture is how to *use* reasons once the agent has gathered or uncovered them. A common metaphor for the process is that the agent weighs the reasons against one another. The idea is that reasons present themselves as

³¹ Versions of a requirement thesis can be found in Broome's “Normative Requirements” (2000; on p. 80 he writes: “If you have a reason to *q* and no reason not to *q*, then you ought to *q*.”), Korsgaard's *Sources of Normativity* (Reply to Williams) (1996, pp. 225-6), Nagel's *The Possibility of Altruism* (1970), Gibbard's *Wise Choices, Apt Feelings* (1990, p. 160), Parfit's “Reasons and Motivation” (1997, p.99, 121), Raz's *Practical Reason and Norms* (1990, pp. 28, 32), and Smith's *The Moral Problem* (1994, p. 148 and throughout ch. 5).

³² One of my eventual objections to this view will stem from its rather flimsy guidance in the realm of rational permissions.

having more or less import—more or less strength or weight—and the agent’s job is to total up the weights favoring each alternative and choose the option with the weightiest set of reasons in its favor.³³ Of course, the process is rather more complicated than this, because reasons come not with “absolute” weights, but relative ones; moreover, sometimes one reason can cancel out another altogether. If Maria is deciding how to spend an evening, the fact that the play in the nearest theater is excellent is a very good reason to go to it.³⁴ If her options are attending that play and staying in to watch reruns on television, let us suppose that the play clearly wins. If her choice is between going to the play and visiting her dear but ailing mother, however, perhaps the reasons to visit her mother win. The excellence of the play is in this sense *pro tanto*: it is a reason that can be outweighed by other reasons. *Pro tanto* reasons purport to have a certain import, but may be counted for more or less upon reflection.

Note, however, that an outweighed reason is still *some* reason for action. A *pro tanto* reason is still a reason, all else held equal. Ceasing to be a reason simply by being outweighed would undermine whatever normativity the reason had. So it is also standard to assume that reasons are *pro tanto*: that is, they retain whatever requiring power they have even when that power is not as strong as that of competing reasons.³⁵

2.1.3 Conclusion: analogy to theoretical reason

Given all I have said about practical reasons, a particularly salient feature of the “anti-DBR” story of practical reason I have been telling is that it bears some analogy to a simple model of theoretical reason. Reasons for belief take the form of propositions describing facts that count in favor of the belief (by providing evidence for it). That is, if q is evidence that p then q is a reason to believe that p . Reasons for belief can serve in both explanatory and normative roles, with the normative role more “primary” in that it can serve as part of an explanation *post hoc*. Reasons for belief are also universal in the sense I have outlined; they have the same *pro tanto* qualities; and they more or less *require* belief on the part

³³ Raz (1990, p. 36) recognizes this as an intuitive conception of practical reasoning.

³⁴ This example is a modified version of one Raz gives (1999, p.99ff). I will be mentioning his actual example later.

³⁵ The picture gets more complicated, for sometimes reasons can be cancelled or defeated entirely, and not just outweighed. I will discuss this further below.

of the agent (in the sense that there is something wrong if you do not believe some proposition to the extent warranted by the evidence).

Now, there are disanalogies between theoretical and practical reason, which can be exploited to illustrate the special problems involved in understanding normativity in practical reasoning. For instance, desire and belief are often said to have two different directions of fit: belief's aim is to fit the world, and desire's aim is that the world fit it. Thus, a disanalogy between theoretical and practical reason is rooted in practical reason's lack of a substantive aim to regulate it, where theoretical reason is regulated by the way the world is.³⁶

But there is one aspect of the analogy that the standard picture retains despite acknowledging and making use of disanalogies. Note that the main model for good theoretical reason is scientific discovery: a believer should model his beliefs on the world, and the more evidence he has, the better, so a central task he has as a believer is to gather as much evidence as possible by investigating the way the world is. Extending the analogy between theoretical and practical reason would suggest that the standard view of practical reasoning is also essentially cognitive and investigative, seeking to discover the reasons as they are, independently of the will of the agent. This is the case even though "counting in favor" on the practical side differs from "counting as evidence" on the theoretical side; a deliberating agent is seeking not only facts, but norms—nevertheless, he is seeking to *discover* those things in features of himself and the world that are independent of what he (even rationally) wants. Though his reasons may have a connection to his motivations and motivational capabilities, he is not in the business of inventing them, just as the scientist is not in the business of inventing her evidence.

All of this presents a picture of practical reason that is cognitive and systematic in a way that resembles science; we seek to discover what reasons exist independently of our deliberative process of discovery. Of course, real practical reasoners do not live up to the ideals presented in this picture of practical reason; we are lazy and fail to gather information or reflect on its import, and we let emotions color our judgments instead of assessing the warrant for those emotions. But this "anti-DBR" picture that I have drawn should sound more or less like an ideal that we at least try to hold ourselves to when we attempt to make

³⁶ Darwall discusses this disanalogy in "Because I Want It" (2001). The disanalogy is not necessarily the end of the story, however; David Velleman defends the claim that practical reason *does* have a substantive aim, which is that the agent "know what he is doing," i.e., that his actions make sense to him (see "The Possibility of Practical Reason" (2000)). But since my concerns are elsewhere, I will not pursue this point here.

decisions responsibly. Against this background, then, I will go on to present several variations on and complications of this anti-DBR picture which try to account for phenomena it misses. They, too, seem to have this discovery model of deliberation in the background. Ultimately, it is to this model of deliberation that I object. The challenge I raise in the next chapter and my response to it suggest that we do not, and should not, always step back from desires into “discovery” mode—instead, what I will suggest is that there is also a “creation” mode of deliberation.

2.2 *Exclusionary reasons*

As I have laid it out, the anti-DBR view of reasons casts all reasons on the same level, all competing (so to speak) on the same scale, with the outcome of practical reasoning depending only on the collective weights of reasons for each option available to the deliberating agent. Although it can be argued that the same reason may weigh differently in different contexts depending on what is at stake in the decision,³⁷ the basic idea so far is that decisions are made by the “pile” method: pile up all the reasons (with their particular weights in the particular context) on their respective sides of the scale, and read off the balance.

Practical reasoning often seems more complex than this, however. Reasons seem not only to outweigh one another, but also cancel³⁸ or exclude one another. Consider again Maria from the above example, trying to decide between going to an excellent play and visiting her fragile mother. If Maria learns that her mother has been rushed to the hospital because her condition has worsened drastically, it seems as though she now has reasons not to weigh the excellence of the play in her calculation at all—in fact, it would seem rather heartless of Maria to do any weighing, given the place her mother has in her life. The excellence of the play does not stop being a reason—it is not cancelled—but in this scenario we might think it is not outweighed but rather *excluded* from consideration under the circumstances.

³⁷ Scanlon argues (1998, p. 53) that “[d]epending on the place in life that an aim is to be given, different reasons will be relevant to the decision whether to adopt it.” He notes that considerations that support taking up an activity for amusement or entertainment will often seem absurdly trivial or out of place when cited as grounds for fundamental personal commitments.

³⁸ Raz (1990, p. 27) notes that canceling conditions, such as a friend releasing you from a promise, are different from outweighing or being excluded, for the conditions themselves may not be reasons at all, and in any case do not involve conflict with the reasons they affect. Thus they do not affect the way reasons interact with one another, which is my primary interest here.

With this kind of scenario in mind, Joseph Raz argues that not all reasons compete on the same level. Instead, there are both first-order and second-order reasons; second-order reasons are reasons to act for a (first-order) reason or refrain from acting for a (first-order) reason.³⁹ He calls second-order reasons to refrain from acting for some first-order reason *exclusionary reasons*. So, to borrow his examples, a woman who knows she is too tired and upset to make a sound decision may have reason not to act on an investment opportunity that she can only take advantage of on that particular day. As a result, she decides against the investment—but not because she concludes that there is not enough reason to make the investment. Her decision is not made on the “first-order” merits of the case, but rather on the basis of a reason to exclude the reasons for and against making the investment from consideration. Other exclusionary reasons include certain promises, or the orders of an authority, as when a soldier sees good reason to disobey his commanding officer but does not allow himself to base his action on that reason. The officer’s authority excludes other considerations without canceling or outweighing them. Similarly Colin, a man who promises his wife that he will choose a school for their son based solely on the son’s interests, has given himself exclusionary reason not to consider the convenience of saving money by sending the boy to a public rather than a private school. Conflicts of interest and decisions already made can also function in an exclusionary capacity.

The conflict between an exclusionary reason and the reason it excludes is somewhat indirect. Let r_1 be a first-order reason to a and r_2 be an exclusionary reason not to act on r_1 . Then r_2 is not necessarily a reason for not a -ing. It may turn out that Colin’s son’s interests coincide with his own, so that even with his own interests excluded from consideration, the school he chooses for his son is the one that costs the least. His promise is not that he will not *act* in a way that satisfies his interests, but only that he will not make his choice on that basis.

According to Raz, if exclusionary reasons exist, then one may act against the balance of first-order reasons if the reasons that tip the balance are excluded by some exclusionary reason. This leads him to conclude that one ought to act, not on the balance of reasons (as in the basic view), but “for an undefeated reason”.⁴⁰ Reasons that are either outweighed or

³⁹ Raz (1990), ch. 1.2 and Appendix. This particular definition is given on p. 39.

⁴⁰ (1990, p. 40). Some care is in order here. To some it may sound strange to say that one is to act for *an* undefeated reason, when there are several considerations contributing to making an option the most reasonable

excluded are defeated in Raz's terms. On this variant of the anti-DBR view, then, reasons are still requiring, but they have a more intricate structure than the "pile" model of the standard weighing metaphor.

However, exclusionary reasons can function on both a first- and second-order level and, furthermore, they may conflict with one another. Because the resolution of such conflicts is typically done by comparing weights, one may wonder whether purported exclusionary reasons are reducible to first-order reasons that are much stronger than other reasons, so that the structure introduced into practical reasoning by the concept of exclusionary reasons is illusory. For example, a manager in the process of hiring excludes the fact that an applicant is a relative, even though such a consideration is normally a reason to help the person out. But, the objection goes, perhaps what is going on here is not exclusion; rather, the reason to help out a relative is simply outweighed by, e.g., considerations of fairness. Against this objection, Raz argues (I think rightly) that the exclusionary structure better captures the way conflict resolution looks from the deliberator's perspective; in the case of the soldier who obeys a command despite seeing good reason not to, he does not think that the commander's order is a stronger reason than those that support disobedience; on a first-order level he may even think it weaker. He simply considers himself bound not to consider anything other than the command.⁴¹

Nevertheless, we might think that purported exclusionary reasons are not reasons so much as constraints on the situation that make certain (first-order) reasons irrelevant: their presence is not some kind of higher-order reason, but a parameter that makes the situation altogether different. Thus we might think that Colin's promise is not a reason for him to ignore his interests, but rather creates a new situation in which his interests simply don't count as reasons. Or the woman's tiredness creates a situation in which complicated decisions simply do not come up for consideration.

Again, this does not seem to capture the way the agents think about the situation. The woman may feel the pull of the reasons to consider the proposed investment; she does

but each of which, on their own, would not sufficiently support the option. Raz slides between taking reasons individually and taking them collectively: he will speak of what we might call "contributing" reasons, but he will also speak of *the* "complete" reason for an action, which is something of a composite of all of the contributing reasons. When he says one ought to act "for an undefeated reason," then, I take it he means a *complete* undefeated reason, i.e., a set of what I call "contributing" considerations.

⁴¹ In the appendix to his (1990), Raz has a good deal more to say about why such things as authority work as exclusionary reasons. For my purposes here, however, the basic intuitive examples are sufficient.

not *feel* as though she is in a situation in which these reasons do not even come up for consideration. She feels as though they have weight, but her mental fatigue simply excludes them from consideration. Colin, too, may feel the tug of his own interest (in saving money), but know that his promise excludes his interest from consideration in choosing his son's school. In both cases, the reasons seem to be present, with all of their import, but simply set aside, as if behind a glass wall.

In addition, as Raz argues, exclusionary reasons better account for our conflicting reactions to cases in which the agent acts on the balance of first-order reasons, in defiance of an exclusionary reason, or when she acts against the balance of first-order reasons in compliance with an exclusionary reason. If the soldier does disobey his commander, we feel inclined to praise the soldier for acting “on the merits” of the case (remember, by hypothesis there are good reasons—say, moral ones—to disobey), but may feel obliged to register our dissatisfaction with the fact of his disobedience; or, if he obeys, we may feel disapprobation toward his action “on the merits,” but excuse the soldier because of the nature of the command hierarchy.⁴² If it were the case that exclusionary reasons changed the situation entirely by rendering some reasons irrelevant, or if exclusionary reasons were really exceptionally strong reasons, we would not have these conflicting reactions.

I do not consider my discussion here to be a conclusive defense of exclusionary reasons, though I think the intuitive case is reasonably convincing. If they are ultimately defensible, exclusionary reasons help capture an aspect of first-person deliberation that the basic anti-DBR view, with its “pile” method of deliberating, seems unable to capture: a hierarchal structure of reasons. Nevertheless, I take exclusionary reasons to be a complication in the spirit of the anti-DBR view because they retain the crucial universality and requiring aspects of reasons that are most characteristic of it, as well as (perhaps more importantly) the discovery model of deliberation.⁴³

⁴² Of course, we need not feel torn in either of these ways; certainly there are some situations in which soldiers are blameworthy for obedience to an order, and sometimes disobedience is worse than doing the supposedly “right” thing. The point here is that there are realistic situations in which our feelings run both ways, and exclusionary reasons provide an explanation as to why.

⁴³ It is important to note that Raz makes significant departures from the standard anti-DBR view in other important threads of his thought. The concept of exclusionary reasons taken by itself, however, remains compatible with this view.

2.3 *Satisficing*

Another aspect of practical reasoning the anti-DBR picture of practical reason misses is the fact that deliberators commonly do not search exhaustively for the best reasons for their choices, and this does not always seem to be the result of laziness, fatigue, or ignorance. Instead, agents sometimes settle for something that is “good enough,” something that will satisfy the present need, goal, desire, etc. without necessarily being the optimal choice (where this means the choice best supported by justifying reasons)—and this seems to us to be rationally acceptable. Perhaps Gordon is searching for a bottle of wine for a dinner party and buys the first one he finds that will go with the chicken and pasta dish he is serving, even though he knows that if he looked longer, he would probably find something better. Or maybe Rose declines an offer of chocolates from a coworker, even though she knows they would taste delicious, saying she’s fine as she is. Both of these people turn down something better than what they have, even though it is something that is readily available to them. The anti-DBR picture suggests that rational action is that which is done for the best or most reasons, so commonplace occurrences such as these would be irrational (or mistaken in some weaker way). But we tend to think they are not.

We might speculate that these cases strike us as rationally acceptable because they are cases where the agents only *appear* to be violating the anti-DBR view of reasons. The examples could be underdescribed. After all, it could be that Gordon is in a hurry to get to a meeting he promised he would not be late to, and is therefore rational to sacrifice finding a better wine for getting to the meeting on time, and Rose is watching her weight and prefers to forgo chocolates in order to stick to her diet. This may well be true, and in that case, there is something that they are not optimizing (the gustatory pleasure of the moment for Rose, the best impression of the guests for Gordon), but although their choices fall short of the best in *some* sense, they are still acting in accordance with their best reasons—that is, we might say, they are optimizing their rationality. So we must be careful to keep satisficing with respect to goods of various kinds separate from satisficing with respect to the reasons one has for action. If we do so, it might turn out that the former kind of satisficing is rationally acceptable (if it turns out to be satisficing at all),⁴⁴ but the latter is not. That is,

⁴⁴ Some theorists argue that satisficing with respect to goods is not really satisficing in a robust sense. The debate may turn out to be terminological, but see papers by Schmidtz, Narveson, and Richardson in the Byron volume *Satisficing and Maximizing: Moral Theorists on Practical Reason* (2004).

satisficing with respect to certain goods (or at certain times) can be instrumental to maximizing overall good, and as such it is just the natural predicament of agents with conflicting ends. I wish to leave this discussion aside; here my interest is in whether “non-instrumental rational satisficing” is possible—whether it can ever be rational to satisfice with respect to one’s *reasons*, and not just one’s ends.

It has been argued that cases like Rose’s and Gordon’s need not have such a rationalizing explanation. Proponents of non-instrumental rational satisficing⁴⁵ contend that it can be rational for agents to make choices that are not in accordance with the best or most reasons they have, i.e. that it is rational to settle for being satisfied with something for which one has sufficient reason, even when there is something for which one has better reason. There are a number of kinds of cases here. For instance, let us suppose that Rose is not, after all, on a diet; that she likes chocolate; that she has just had lunch, but is not full; and so on. In the absence of positive reasons not to eat the chocolate, it may seem as though she is a little irrational to pass it up—after all, she has no reason not to take some. Yet her excuse, “I’m fine as I am, thank you, though,” seems perfectly reasonable and understandable. Rose is acting moderately, and we can understand and endorse the motive from which she acts.

This example may still be unconvincing, for we could easily interpret Rose’s moderation as a factor that makes it most rational for her to turn down the chocolate. Still, we might follow Michael Slote and argue that there is a tight analogy between morality and reason, and given that we tend to think it permissible to perform an action that is morally satisfactory but not morally best (as we do when buying small luxuries for ourselves instead of giving the money to someone in need, or failing to help an old man carrying a heavy bag of groceries into his house), we should expect to find cases in which it is permissible to do what we have sufficient, but not most, reason to do. What we need in order to establish that non-instrumental satisficing with respect to reasons can be rational are cases in which an agent does something that we can count as rational, but there is something she could have done that is *more* rational than what she actually chooses.

⁴⁵ E.g., Michael Slote (1989) and Michael Weber (2004).

If the analogy between morality and rationality holds, Slote argues, the existence of moral supererogation should lead us to expect phenomena that are recognizable as “rational supererogation.” Here I borrow a case from Slote:⁴⁶

A man is angry at his boss and believes (for good reasons we needn't enter into) that it would be in his own interest, and on the whole a good thing, to tell the boss off when they next meet. He believes that it would be best to do so in a loud enough voice so that everyone in the office will know what is happening. But both he and his fellow employees have long been intimidated by the boss, and the employee knows it will be difficult to stand up to the boss and tell him what he thinks, and even more difficult (though not impossible) to do so in the loud and angry tones he thinks most appropriate. When the time comes to confront the boss, he manages with considerable effort to speak his mind, but allows himself to be intimidated to the extent of not daring to do so loudly and angrily; and indeed from a rational standpoint and relative to the man's own values, it would have been better for him to speak angrily to the boss than to express his opinion in conversational tones. But the latter may have been difficult enough so that on the whole the man is more proud than ashamed at his performance.

The idea is that what the man actually does is not, in the end, what he had most reason to do, but it is still rationally acceptable—he did what he had plenty of reason to do, and that was good enough. To do more would have been more admirable from a rational point of view, but he still did something that counts as rational.

One might be inclined, at first, to accept this case as an illustration of non-instrumental rational satisficing. We can imagine that the man knows that what he did is good enough, and will make the needed difference, but he nevertheless can't help berating himself for not going all the way and speaking loudly to the boss. He feels the pull of the intuition that one should act as one has most reason to act. And we can imagine him discussing the confrontation with a coworker or a friend, expressing his exasperation at himself for not doing better, and being told by the friend that what he did was plenty good enough. The friend need not be “just saying” this for the sake of consoling the man; she could be fully sincere.

Yet I think that accepting this case as an illustration of what rational supererogation could be trades on a confusion between rational action and rationally excusable action. The friend consoles the man because what he did was very good, even if not the best—but this is because, under the circumstances, there were substantial motivational barriers to doing what

⁴⁶ *Beyond Optimizing* (1989, p. 119).

was most rational. The man was *aiming* to do what he had most reason to do, and can be excused from failing to achieve it because of the boss's intimidating demeanor and because what he did manage to accomplish was a reasonable improvement over the status quo. But if the man had aimed at less—if he had *aimed* only to talk to the boss in conversational rather than angry tones, knowing that he had more reason to speak angrily, no doubt the friend would be criticizing (or at least questioning) him for not doing what he had most reason to do. It would be like accepting the lower of two offers on a house just because the lower offer was good enough.

Nevertheless, I recognize that intuitions about when it is rational to satisfice and when it is not are messy and can run in different directions when the details of situations vary. This is a fact that makes Michael Weber uneasy as well, and in “A New Defense of Satisficing” (2004) he suggests a more systematic way to defend satisficing, attempting to go beyond using cases to pump intuitions. Weber's proposal is essentially that because agents are capable of occupying a number of competing perspectives, each of which may make a different action rational, it is rationally permissible to satisfice in one perspective because from another the choice simply doesn't matter. Thus Ann may choose to be a fine lawyer rather than the best lawyer she can be in order to spend more time in leisure activities or with her family—here we might think Ann satisfices with respect to the “global” perspective of a whole life in order to make each moment more pleasing. Or, to take a different case, Dave and Tim may satisfice with respect to a restaurant choice for today's lunch because, from the “global” perspective of their entire lives, the restaurant they eat at on any given day is simply insignificant.

Weber argues that he is not giving in to the maximizer, because his claim is that Ann's choice is rational *even if* the benefits of leisure and family do not, overall, outweigh the costs to her career; and similarly, Dave and Tim's choice is rational not because the costs of searching further outweigh the benefits of a better meal, but because in the long run a better meal just doesn't matter.

Even this proposal seems caught in the maximizing mentality, however; it strikes me that what Weber has done is less a defense of satisficing than an elimination of it—an explaining away.⁴⁷ For there is more to say about why it is rational for Ann and Dave and

⁴⁷ Richardson (2004) makes a related point.

Tim to make the choices they do, even if they are motivated in ways that make Weber's claims true. That "more" has to do with values and the normative role they play as regulative ideals in our practical thinking. Ann may place a higher value on leisure and family than she does on career success, and even if this predictably does not make for a better life on some objective scale, she can be seen as rational in that she holds herself to an ideal that represents something deep about who she is, or takes herself to be. Dave and Tim, on the other hand, may value keeping things in perspective, and this value may be what guides them not to search too hard to find a better restaurant for their casual lunch.⁴⁸

Weber discusses valuing moderation or "taking into account the variety of (temporal) perspectives we can occupy" as regulative ideals.⁴⁹ These values are not things to be maximized the way (for instance) pleasure or professional excellence can be. Instead, they serve as constraints on the way we maximize other goods, introducing structure in a way perhaps analogous to exclusionary reasons. So Ann's ideal of being a well-rounded person constrains her daily choices, excluding certain career-related considerations from guiding her day-to-day choices. Thus, Weber claims, "the good life...cannot be construed as maximizing anything." This is true if Weber means that the good life cannot be construed as maximizing any particular good. But in acting in accordance with regulative ideals, we optimize with respect to the reasons we have for acting. Since some reasons do not fall under the description of "promoting some good," not all of our actions can be construed as maximizing goodness. My point is that satisficing, if it is truly satisficing, must be construed as satisficing with respect to some good; it is still not clear that it can be rational to satisfice with respect to reasons if reasons require action on the part of the agent whose reasons they are.

⁴⁸ If this is what guides them, of course, it seems as though they are not really satisficing, since they are maximizing on another value (and this was my point). Still, it could be argued that the restaurant example is not supposed to be of that sort because Dave and Tim already know that restaurant A is better than restaurant B, and they go to B anyway. But if they know this, then the fact that where they eat for lunch just won't matter in the long run simply is not a good reason to choose B over A. If the insignificance of the decision plays any role, it must be in combination with a further reason for choosing B—e.g. that it is closer, or has faster service, or is cheaper. In that case they might be satisficing with respect to the best meal, but not with respect to some other value. Satisficing seems to me to be rational in a case where a search is curtailed on the basis of its being unimportant in the long run to maximize each and every choice, but not when the available alternatives are known and the choice between them is supposed to be clear.

⁴⁹ Weber (2004, p. 97).

So while it is true that the need to balance our competing perspectives is what drives behavior that appears to be satisficing from the perspective of goodness, it is the values underlying the perspectives that guide the way we balance them, and if agents are keeping true to their values, then they are not necessarily satisficing with respect to their reasons—though we might say they are satisficing with respect to pleasure, preference satisfaction, etc. I will say more later on about how reasons come from an agent’s values,⁵⁰ but for now the point is that although satisficing purports to take into account phenomena that the anti-DBR view of reasons does not (namely, the fact that we sometimes deem it acceptable to act on sufficient reasons that are not the best), it is not clear that satisficing *with respect to the reasons we have* is fully defensible unless there is more to say about the nature of reasons.⁵¹ Thus, I turn to another approach in the next section.

One further comment is in order, however. As we will see in Chapters 4 and 5, my rejection of the discovery model of deliberation in favor of a self-creating view will have many ties to a choice of perspective and ideal. In this sense I have much in common with the spirit of Weber’s project, even though I disagree with respect to his claims concerning satisficing.

2.4 *Optional reasons*

The way to take account of cases like Gordon’s buying a suboptimal bottle of wine or Rose’s refusing the chocolates *and* get around the conceptual difficulties involved in defending rational satisficing is to argue that reasons do not have the structure that is presupposed in the anti-DBR view. As I have laid it out, that view supposes that all reasons are “first-order,” and every reason has the same status as every other reason—they are all universal, they all require, etc. In exclusionary reasons we have seen one way of suggesting a different structure for reasons, but it retains the discovery model of deliberation as well as the requiring aspect of reasons that seemed problematic in explaining cases of supposed satisficing. Now I will discuss schemes according to which reasons do not all have the same status, and they do not all require action on the part of the agent.

⁵⁰ See Chapter 5.

⁵¹ See Dreier (2004) for an argument that in fact there can be no satisficing with respect to reasons.

2.4.1 *Relaxing the “requirement” requirement*

In *Engaging Reason* (1999), Joseph Raz defends the basic thesis that the primary function of reasons is to make actions eligible by making them intelligible from the point of view of the agent.⁵² “What I call *the basic belief*,” he writes, “is that this structure of reasons is most common: that most of the time people have a variety of options such that it would accord with reason for them to choose any one of them and it would not be against reason to avoid any of them.”⁵³ Reasons, then, “are optional to the extent that the fact that there are reasons for a certain response make it an eligible, attractive response, but not one which it is wrong not to adopt.”⁵⁴ This basic belief is easy enough to understand, because, as Raz rightly insists, it is entirely common. His example is Mary, who learns that a local theater is running a powerful adaptation of a good play. It is easy to agree that she has good reason to go to the theater this evening; yet we would also not think her unreasonable if she chooses not to see it.⁵⁵ If she would rather read the newspaper, take a walk, relax in a bubble bath, call her sister, or even do nothing, no one would fault her. She has an adequate reason to go to the play, but that reason is not so strong as to require her to go. She may recognize the reasons to see the play as good ones, but choose not to act on them. On this particular night, those particular reasons may not move her all the way to action, and that’s all right.

This sort of situation is readily recognizable, because we face it almost daily: there is no uniquely rational⁵⁶ action among the available alternatives. The situation is almost

⁵² Raz (1999), p. 26-7.

⁵³ Raz (1999), p. 100. How does the phrase “against reason,” which seems to be the operative phrase here, compare with such alternatives as “irrational”? I take it that he means something weaker than “irrational,” because in the next paragraph Raz indicates that even when there seem to be more reasons favoring some option, as long as there are *some* reasons favoring an alternative, an agent is not irrational to take the less favored alternative. This seems to say that she is not required to take the favored alternative, which means that Raz must take the view that in certain situations, reasons do not require action of us even when there is more to be said for some option. (At this point it is not clear what sort of “requirement” Raz means. Certainly he means “rational” requirement, but his phrasing (in the claim that Mary is not “doing anything wrong”) is such that it might include moral or prudential requirement as well.) Note, however, that Raz still allows that there are times when reasons *do* require that one perform some action. His point here is only that this is not the most common situation.

⁵⁴ Raz (1999), p. 95.

⁵⁵ Note that it is a somewhat different matter if she fully judges she has most reason to see the play. If, given this judgment, she does not go, then we have cause to call her irrational, unreasonable, or wrong in some way. She is, in some sense, subjectively irrational—something has gone wrong with her reasoning—because she is not acting in accordance with her judgment, or because she is not acting on what she has most reason to do.

⁵⁶ The phrase “uniquely rational” can be taken in a number of different ways. One may think that some action is “uniquely rational” in the sense that it is best supported by reasons, i.e. that the only thing that counts as rational is the single “most” rational thing to do. As we shall see, this is an interpretation one could hold

unremarkable—so the project of giving an account of how reasons, while guiding practical judgment, can be optional in this way is easy to overlook. It is easy to concentrate on determining what an agent has most reason to do, and to assume as a result that if he does not then do this, he is irrational (broadly speaking).⁵⁷ An optional-reasons theorist such as Raz need not deny that there are plenty of cases in which reason does require or direct us to a certain choice; after all, part of the function of reasons is to guide deliberation. If the reasons clearly indicate what the agent is to do, then it would be irrational (again, broadly speaking) for him not to do it. But optional-reasons theorists contend that we must be careful how we understand rational requirement; and they warn us that it is less common than philosophers sometimes think.

Raz diagnoses the basic belief introduced above as resulting from a certain kind of incommensurability: an incommensurability among reasons. He states that two reasons are incommensurate if neither defeats the other and they are also not of equal strength.⁵⁸ An agent with incommensurate reasons may then choose either option, not because the options are equally good or reasonable, but “because it is reasonable to choose either option (for both are supported by an undefeated reason) and it is not unreasonable or wrong to refrain from pursuing either option (for both are opposed by an undefeated reason).” Raz thinks reasons are *not* optional if they defeat or are defeated by other reasons in the situation, for defeat would involve interaction and thus comparability; so this version of optionality is relative to the other options on the table and does not stem from the nature of reasons, their function, etc.

The kind of incommensurability Raz has in mind is among reasons, rather than options. He claims that having incommensurable reasons is compatible with having

compatibly with holding that it is not “against reason” to do something else. In this case, however, I mean that there is no “uniquely rational” action in the sense that more than one action is supported by good reasons, and all it takes to count as rational is being backed by good reasons. Hence there is no action that is unique in the sense of being the only rational one. This leaves open the question of whether one is *best* supported.

⁵⁷ We will have cause to distinguish between two senses of “irrationality”: one is a very weak sense, in which an agent does not act for his best reason; the other is a stronger, much narrower sense, in which he is doing something outright crazy (e.g. putting himself in great danger with little or no chance of gaining a compensating benefit), or has contradictory beliefs, etc. (I will leave the substance of this strong sense vague for now.)

⁵⁸ Raz (1999), p. 103. Raz’s way of casting incommensurability in terms of defeat can be contrasted with the logically distinct notion of incommensurate weight. According to this notion of incommensurability, two reasons are incommensurate if neither is weightier than the other, but neither are they equally weighty. Raz’s notion is logically consistent with the possibility that one reason is weightier than another even though neither defeats the other.

commensurable options. For example, it may be that films are fully rankable according to a variety of factors, and that these multifarious factors all together can be bundled into a ranking of overall film quality.⁵⁹ Let us suppose that there are two films you are considering seeing tonight. One is overall mediocre, but amusing. The other is overall quite excellent, but treats a heavy subject matter. On the scale of film goodness, the choice is clear: you have more reason to see the excellent film. But suppose it has been a difficult week at work and tonight, amusement is just what you need. So all-things-considered, it seems it would be reasonable for you to choose either way. If this is so, then it seems that reasons of film excellence are optional. Why is this? It's not because the films are incommensurable on the scales of overall film quality or amusement, for we've stipulated that the films are commensurable on these scales. Rather, it is because your reasons for seeing each film tonight are incommensurable: amusement on the one hand, and overall film excellence on the other. These two reasons don't seem to compare fully; neither seems to outweigh the other, but neither do they seem to weigh the same. And so your choice would be reasonable either way you decide.

Of course, this incommensurability of reasons only explains the open option between alternatives on the assumption that both are supported by reasons. If Mary chooses to spend her evening "in some worthless pursuit," we will rescind our judgment that it was reasonable for her not to go to the theater. If the mediocre film did not excel in amusement (exactly the quality you were in the mood for), it would be (at the very least) very surprising for you to choose it over the excellent film.

Even though philosophical discussion of incommensurability and incomparability is ongoing, the picture of deliberation which emphasizes reasons' role in making options eligible, rather than requiring them, strikes me as eminently reasonable, and throughout what follows I will take it that something along the lines of the basic belief is true. Nevertheless, I do not think that even optional reasons theories are adequate to explain a case like the one I presented in my introduction, for they do not help us to understand the way a choice such as Smith's is made even when reasons are optional. Before I go into detail on this kind of case, however, I wish to explore two other ways of explaining the plausibility of the basic belief. These are given by Jonathan Dancy and Joshua Gert, respectively.

⁵⁹ Both Raz and I are ultimately skeptical of this. But for the sake of explaining the distinction at hand, let us assume that such a ranking is possible. See Raz's (1999) discussion, pp. 103-4.

2.4.2 Dancy's enticers and preemptors

Jonathan Dancy (2004) explains optional reasons by designating two categories: “preemptory” reasons, which have the power to require actions of us, and “enticing” reasons, which do not. These categories are to be understood as exhaustive; anything that is not a preemptory reason must be an enticing reason. Preemptory reasons seem to carry a certain kind of censure when we fail to be guided by them (though it need not be as strong as *irrationality* in a narrow sense⁶⁰). Dancy is not explicit about this; since his purpose is to introduce and explicate enticing reasons, I surmise, he assumes we understand preemptory reasons relatively well (probably not a bad assumption, given the intuitive sense of the anti-DBR view). Despite not being explicit, Dancy implies that someone who violates a preemptory reason is wrong in some way. For instance, a promise gives the promiser a preemptory reason to do what she promised; it may not be *irrational* to act against the promise, but the promiser can certainly be blamed for breaking the promise (certainly, at least, if she had no good reason for breaking it). There is something wrong with breaking a promise, because in doing so you are violating something that is required of you.⁶¹ Or, to take a non-moral example, an automobile owner has a preemptory reason to have the car’s brakes checked regularly. It may not be irrational in a strict sense not to do so, but it is certainly imprudent. Which is just to say that there is something wrong with neglecting one’s brakes.

Dancy does not make explicit his notion of requirement, leaving it as an intuitively understood heterogeneous category; but from his examples it is clear that he includes rational, moral and prudential reasons as examples of potentially requiring considerations. You may be rationally required not to have contradictory beliefs; you may be morally required to keep a promise; you may be prudentially required to have your brakes checked regularly. In each of these cases, there is something wrong if you do not take the action required of you. These are the sorts of things we often point out to one another.

Enticing reasons, however, are not like this; they are considerations that certainly count in favor of an action, but which do not require it of the agent in any of the various possible ways of requiring. Consider a decision concerning how to spend a free afternoon:

⁶⁰ See footnote 57.

⁶¹ Note that this notion of requirement is broader than even the broadly rational sort of requirement I take to be part of the standard anti-DBR view.

shall we go to the park or the art museum? Suppose it is the first sunny day in two weeks, which is a strong reason favoring a long walk in the park. On the other hand, the current exhibit focuses on one of our favorite artists. According to Dancy, our reasons in this case are enticing, and we are in no way wrong (irrational, immoral, etc.) to decide to go to the museum or to the park, *even if* one of them is supported more strongly by these enticing reasons.⁶² He allows that there is rational criticism of a mild sort available, however: if there is clearly more enticing reason to go to the park, and we go instead to the museum, we may be “plain silly”—which, I take it, is just to say that we have not acted as our best reasons dictate; but though this is criticism of a rational sort, our behavior is not strictly irrational.

Dancy argues that the categories have the different requiring powers they do because they stem from different sources and focus on different normative notions. Peremptory reasons are deontic, taking us to “oughts” which (purport to) compel us to perform actions, whereas enticing reasons are evaluative, taking us to “bests” that (purport to) beckon us toward the actions they recommend.⁶³ Both oughts and bests are normative notions, but they do not operate in the same way in our deliberations. Their styles of favoring, as Dancy puts it, are different; and we do not feel the pull of “bests” the way we do “oughts.” Thus, for Dancy, reasons may be “optional” because they entice rather than require.

⁶² I think the best way to make sense of this claim is to construe it along the lines of the case above in which two films are clearly comparable on their merits (and one is clearly better than the other), but your circumstances make it more rational for you to see the less good film. So it may be that being in the mood for a day in the fresh air, together with the merits of exercise and so on, make the park the better option for us today. If instead we decide to go to the museum, it will seem odd, because overall there seemed to be more reason to go to the park. What could explain our choice of the museum, then? All I can think of is that it is some lesser consideration that was outweighed by the considerations that made the overall verdict what it was, such as its being the last day of the exhibit. Imagine the following conversation:

“What did you do today?”

“We went to the museum.”

“Really? I thought you would go to the park—it seemed like the better option.”

“Well, we thought so too, and that’s what we were in the mood for, but since it was the last day of that exhibit, we changed our minds.”

“Huh. That seems like a strange choice, since I thought the park was better even taking into consideration the fact that the exhibit was closing, but ok.”

The interlocutor here remains unsatisfied as to why we went to the museum, even though we’ve given an explanation. And rightly so, it seems to me; even though such a choice does not seem to be out-and-out irrational, so that the interlocutor is left with nothing to object to, it remains rather mystifying.

⁶³ Dancy (2004), p. 116.

2.4.3 *Gert's justification and requirement*

Rather than differentiating them by source, as Dancy does, Joshua Gert (2004) elaborates a theory of reasons which distinguishes two *roles* for reasons to play in deliberation: justifying and requiring.^{64,65} Gert then argues that reasons have two corresponding strengths, filling in the intuition behind this distinction with a paradigm case: The reasons in favor of saving forty children from starvation are strong enough to *justify* risking death, for example by smuggling them food and medicine when they are trapped in a war zone. But these reasons are not strong enough to *require* you to do so, because they are not even strong enough to require you to donate \$100 to Oxfam rather than spend it on a bottle of wine—or nothing at all. (Note that, in contrast to Dancy, we are talking here strictly about *rational* requirement, not moral or prudential or anything else.) Assuming the donation will produce the same effect as the smuggling would in the other scenario, the reasons to donate the money are the same as the reasons to smuggle the food. Yet they can justify risking death, which would make them seem quite strong, when they cannot even require spending \$100, which would make them seem quite weak. This leads Gert to conclude, plausibly, that reasons must have two roles, and their strengths within those roles can be remarkably different. Thus, saving forty starving children has immense justifying power, but little requiring power.

How are we to understand these two different powers of reasons? According to Gert, reasons justify by making it rational to do something that would otherwise be irrational, and they require by making it irrational to do something that is otherwise rational.⁶⁶ Note that for Gert, an option is “rational” in the sense of its being not-irrational; and to this extent he will agree with Raz that reasons justify by making options eligible when they wouldn’t otherwise have been, and they require by eliminating otherwise eligible options. In order to understand Gert’s position fully, then, we must understand what it is to be irrational. To this end, Gert provides a substantive theory of what is rational—or, rather,

⁶⁴ Gert believes, in fact, that there are some reasons that are *purely* justificatory—i.e. that they do not have any requiring strength. He allows, however, that many people will disagree with him on this point, but that such disagreement does not undermine his basic point.

⁶⁵ Terminological note: for ease of exposition, I will often speak of the justifying strength of a reason, or a reason *in its justifying role*, simply by referring to a “justifying reason.” This should not be understood to mean that the reason cannot play the requiring role.

⁶⁶ Gert (2004), p. 65-9.

of what is irrational. Gert claims that it is irrational for an agent to put herself at risk of harm without a proportionate chance of thereby procuring a benefit,⁶⁷ though because this begins to look like Gert reduces rationality to prudence, perhaps it is better to retreat to a prior principle (also endorsed by Gert): that an objectively irrational action is one which cannot be sincerely recommended by anyone to anyone else.

Reasons in their justifying role are optional, on Gert's scheme. This is intuitive, and can be made especially clear by analogy with the moral case. You don't need to act on a reason that would morally justify you in acting—e.g. you don't need to take the justification of self defense and kill (or react in any violent way to) an attacker. Gert claims a parallel with practical rationality: you don't need to act on some reasons that would justify your acting—e.g. you don't need to save the orphans at great risk to yourself, even though you would be rationally justified in doing so.

⁶⁷ Gert (2004), p. 140-1. I am not sure that much of the plausibility of Gert's "two roles" idea rests on his particular normative principle—though for Gert, it is important that there be *some* substantive definition of irrationality, or else his argument that "wholesale rational status" is prior to reasons will be suspect. Certainly much more rests on his contention that an objectively irrational action must be one that cannot be sincerely recommended to someone else (138, 140); but this is also a less substantive and hence less controversial claim. Nevertheless, even this weaker claim is not entirely satisfactory. Elizabeth Anderson has helped me to articulate what seems strange about this way of characterizing the optionality of optional reasons: it seems that Gert must classify activity done for no reason as irrational. And it is much too strong to claim that doodling, whistling, fidgeting, or dancing for no reason are irrational activities. A child's play, too, seems idle without being irrational. We often improve activities (especially those of children) by giving them reason-based direction, but this does not mean that pure directionless activity is irrational. This Dewey-inspired position strikes me as correct.

Is there a way to save Gert's view from seeming like a non-starter? One thing to say might be that activities are never done for absolutely no reason, because even activities done for no apparent reason (even from the point of view of the agent performing them) are always done for the reason that activity is a default state. This strikes me as rather ad hoc.

A more plausible defense would classify all actions as rational by default unless they were clearly irrational (where "irrational" is independently defined, e.g. by the substantive principle Gert outlines). This would classify idle actions as rational. It would also cover actions such as giving money to Oxfam or public radio, where there are perfectly good reasons in favor and none against. It would also classify actions we could call "permissible but not recommended", such as giving money away for no particular reason (when there is no reason not to), as rational. This seems to me as it should be.

This defense, however, raises the question of what role the reasons favoring activities such as giving the money to charity are playing. If we assume with Gert that we are not (rationally, at least) required to donate to charity, and the only two available roles for reasons to play are justifying and requiring, these reasons must be justifying. But it is not plausible to say that they make something rational that would otherwise have been *irrational*. They do, however, give direction to something which otherwise would not have had direction, perhaps turning into something deliberate and justifiable something which was simply impulse and which therefore escaped the need for justification. So perhaps the justifying role of reasons should be suitably amended to include this kind of role. Doing so has little cost in the theory, I think, and it preserves the optional character of justifying reasons.

Ultimately, nothing important to my overall project hangs on whether Gert's particular scheme is plausible; I rehearse it here as an illustration of a different approach to optional reasons. What is important to me, overall, is that some optional reasons thesis be plausible.

2.4.4 *Comparisons*

Even from this brief exposition, it is already apparent that the two schemes are, for the most part, orthogonal to one another. Each has identified important aspects of the way reasons work in deliberation, but the accounts are not directly competing. At first it appears that Dancy's and Gert's categories are different because they come from different angles. Dancy rests his distinction on the differences between value and duty, whereas Gert is concerned with justification and requirement. But upon deeper analysis, the differences between them do not stem from these surface differences. Rather, Dancy sorts reasons using a notion of requirement broader than Gert's, so each is sorting according to a different criterion (in contrast to Raz, who does not sort reasons in this way at all because he thinks they are uniform along the dimensions Dancy and Gert identify⁶⁸).

It is worth noting that Dancy's and Gert's categories partially cut across one another. For example, given that Dancy includes both moral and prudential considerations as potentially requiring, it is possible to imagine a situation in which one is morally required to do something, but not rationally required to do it in Gert's sense. One might think of an amoralist as an example here. The amoralist might say, "Sure, stealing the jewels is morally prohibited, and I am morally required not to do it, but that doesn't say anything about whether or not I *really* have reason not to. You can't criticize me for violating any requirements of rationality. Moral reasons aren't really reasons, after all." Nevertheless, the amoralist who thinks this way might accept moral reasons as justifying in Gert's sense. And of course, it is possible for morality or prudence to coincide with rationality in Gert's sense. So what Dancy would call peremptory reasons may have either of the two strengths Gert argues for.⁶⁹ This is in part a result of the fact that the theorists, as I have said, are working with different notions of requirement on actions: Dancy's broader sense includes many different varieties of deontic consideration, while Gert's narrower sense has strictly to do with rationality.

⁶⁸ Raz (1999), p. 101-2.

⁶⁹ Enticers cannot play each of Gert's two roles—in particular, they cannot require in Gert's sense, because if they could, Dancy would classify them as peremptors, since they would be rationally requiring. This does not detract from the point that the two classifications cut across one another, however.

Despite these differences in their notions of requirement, both Dancy and Gert will accept some version of the paradoxical-sounding thesis that it is not always wrong (in the sense of being somehow mistaken) not to do as you have most reason to do.⁷⁰ But we must be careful to understand exactly what each understands by this thesis, because they are not the same. Both Dancy and Gert will agree that it makes sense to rank options as better and worse,⁷¹ even when they are supported only by enticing or justifying reasons. They differ, however, in what should be said about someone who takes a “lesser” option when a better is available.

Dancy comes close to stating the thesis directly.⁷² On Dancy’s picture, we can make sense of the claim if we distinguish between two types of rational criticism. First, there is the (stronger, narrower) sense in which an agent is criticized for violating some requirement on her (e.g. moral, rational, prudential), where Dancy leaves “requirement” to our deontic intuition, as noted above.

Second, to understand Dancy’s version of the thesis, we must distinguish a (weak, broad) sense in which it *is* unreasonable not to conform with the best reasons available. If, for example, Mary has, and recognizes that she has, most reason to go to the play tonight but decides nonetheless not to do it, we will be puzzled. We might ask her, “But why aren’t

⁷⁰ It is not clear what Raz thinks of this thesis. He comes quite close to an explicit endorsement when he says that Mary “does not have to have a child just because there are reasons for having it and little against” (1999, p. 100). But given that optionality is relative to the situation for Raz, as I argued, it is difficult not to think he is committed to the claim that if you have most reason to do something, then you should do it. In other words, it is difficult to see how he could endorse the thesis under current discussion.

Still, Raz’s exclusionary reasons seem to make it possible to understand how we need not act on our strongest reason. Because they are second-order reasons, they need not *outweigh* first-order reasons; their power is different. They override by excluding from consideration certain first-order reasons—even, perhaps, quite strong ones. Thus, we may end up acting on the basis of an exclusionary reason that defeats an otherwise very strong reason, but not on the basis of strength. In the case of the starving children in the war zone, for example, Raz might say that the danger to yourself functions both as a first-order and as an exclusionary reason. The exclusionary aspect allows you to discount or ignore the very strong reasons to enter the war zone and deliver the food. Because you may take or leave the exclusionary aspect of the reason, even if we suppose that the reasons to save the children are stronger than the reasons for you to stay where you are, you are not required to act as you have most reason to act. Raz would have to explain why the exclusion here is optional, but presumably he could understand the situation in terms of the incommensurability he accepts in first-order cases.

⁷¹ Gert thinks there is no incommensurability (102ff), and reasons have fixed normative strengths within their two roles (76), so for him it will in principle *always* make sense to talk about what you have “most” reason to do. Dancy is less clear about this; in “Enticing Reasons” he says nothing either way about incommensurability. I assume, therefore, that the possibility that options are incommensurate is open on Dancy’s view. Nevertheless, he discusses the possibility of differently weighted but still enticing options (92-3), so it is clear that he does think enticing alternatives can be compared at least some of the time.

⁷² Dancy (2004), p. 92.

you going?” If she can satisfy us, then we will think we were mistaken to believe that she has most reason to go to the play, because (as we now see) she actually has most reason to stay home. If she cannot satisfy us, we will remain puzzled and perhaps criticize her for being silly, stubborn, lazy, or something else.

I contend that Dancy recognizes this weak sense of unreasonability because the distinction between the strong and weak senses is important to his claim that it is not always wrong not to do what one has most reason to do. It is not always wrong because Dancy reserves “wrong” for the stronger sense of requirement. Thus, the seeming paradox is resolved by Dancy’s understanding of “wrong.” It *is* wrong in the weak sense not to do as you have most reason to do; for it is a mistake of a rational sort not to act on the balance of reasons, including enticers. But it is not always wrong in the strong sense, because in neglecting to follow enticing reasons you violate no (deontic) requirement.

This is a departure from Raz’s view because Raz does not divide reasons by style of favoring, nor criticism by strength. For Raz, favoring is favoring, and one reason is as requiring (or not) as another, because Raz retains the idea that reasons are unified in what they are and what they do. Instead, Raz emphasizes the principle that one ought to act, all things considered, for an undefeated reason.⁷³ (Remember that an “undefeated” reason is one that is neither outweighed nor excluded from consideration.) This creates a picture on which the primary role of deliberation is not specifically to discover what we have *most* reason to do, but merely to seek out relevant reasons for actions, which make certain options eligible for consideration.

We saw that Raz emphasizes deliberation’s function of proposing options, and from there it is up to the agent’s will to determine what he will do. I suggested above that this basic picture also stands behind Gert’s view. Recall that for Gert, reasons are optional in their justifying roles, and a reason justifies by making it rational to do something that would otherwise be irrational;⁷⁴ i.e. justifying reasons make eligible actions which would otherwise have been ineligible (or perhaps simply not on the agent’s radar). Then, with all of the eligible options before him, it is up to the agent how to proceed. In the war zone, you may take or leave the option to save the forty children: the great risk of harm normally makes it

⁷³ Raz (1990), p. 40.

⁷⁴ Though see note 67.

irrational to enter the war zone, so under normal circumstances, this is not an eligible option. But you are not under normal circumstances; you are in a position to deliver food that may save forty starving children. The well-being of the children is a large enough benefit to make the risk worth taking—but because of the role these reasons are playing, you are not required to make the delivery. Thus, you have before you two eligible options supported by very good reasons, and it is up to you which you choose. Neither decision would warrant criticism, even of the weak sort discussed above, so it is up to you which action to take.⁷⁵

Nevertheless, as I explained above, Gert thinks it always makes sense to compare options, and from there it is a short step to ranking them. Yet Gert explicitly endorses the thesis that we are not always required to act on a stronger reason.⁷⁶ If it is always possible to come up with a ranking of eligible options, how can Gert make this claim?

It is not because, as in Dancy's account, the weak kind of criticism has nothing to do with requiring. Rather, it is because of Gert's understanding of a reason's strength combined with the thesis that justifying reasons are optional. Roughly, a reason's strength is understood as its power to do what it does—a justifying reason is stronger than another if it can justify more, and a requiring reason is stronger if it can require more.⁷⁷ So it can be the case that there is a very strong justifying reason for an action, and only a weak one for a different action, but it is not irrational to act on the weaker reason, because both are equally optional in virtue of their playing a justifying role in the deliberation. To take Gert's example: suppose you are choosing between donating \$200 to Oxfam and donating the same amount to public radio. The reasons (such as starving children) to donate to Oxfam are no doubt stronger than those in favor of donating to public radio, because they can justify more (remember that saving the children could have justified entering a war zone; would anyone be justified in dying for public radio?). Yet you are free to donate to public radio because

⁷⁵ This sort of scheme appears to allow for a much more robust notion of permissions than does the standard view, for the agent in Gert's scheme may legitimately cite (optional) reasons for her choice of one over the other, whereas on the standard view the choice of one *over the other* must be arbitrary. Although of course the agent will *cite* the reasons favoring the option she chooses, according to the standard view, she must be incorrect to think that they are winning considerations, if the choice is truly up to her; if it is not up to her after all (and she merely thinks it is), then her action is not *permissible* but rather either required or prohibited. In the coming chapters, however, it will become evident that even this notion of permission is not enough to account for the resolution of the conflicts I am interested in.

⁷⁶ Gert (2004), p. 87.

⁷⁷ Gert (2004), p. 65, 68.

both of these donations are justified,⁷⁸ but neither is required. Thus, you are not required to act as you have most (strongest) reason to act.

2.4.5 *Lessons*

It speaks to the complexity of an agent's relationship to her reasons that these three accounts are defending similar theses, yet doing so by turning to different phenomena. The two views I have just discussed will allow that there *can* be (rational) requirement, but that this requirement is not generated by the justifying power of any reasons. Gert in fact specifically denies the thesis that sufficient justifying reason will always yield requirements; that is, he denies that "any reason that can *justify* can also *require*, if it is instantiated strongly enough or in sufficient numbers, or if countervailing reasons are weakened or removed."⁷⁹ This is something Dancy would deny as well, because enticing reasons are meant to be a kind of reason that normatively justifies some course of action, but yields no rational requirement (in the strong sense). While they may differ on some overlapping details, the two accounts are arguably consistent, or can be made so with minor modifications.

Raz, of course, will agree with some of the claims made by Dancy and Gert, though I believe he would stop short of endorsing the thesis that no reason that plays the "justifying" kind of role in Dancy's or Gert's schemes could never play a "requiring role," since Raz does not sort reasons along these lines, but instead analyzes the optionality of reasons as resulting from incommensurability.

2.5 *Conclusion*

What can we learn from optional reasons theories? First, it is important to note that these optional reasons theories are not a return to the "Humean" picture of practical reason according to which desires are, or are the only source of, reasons. Rather, they operate against the backdrop of the anti-DBR view, which adopts the Pettit-Smith stance that desires are "backgrounded" when it comes to our practical reasoning. Even when the reasons are optional, a desire is not a reason of any sort—even tiebreaking.

⁷⁸ Again, see note 67.

⁷⁹ Gert (2004), p. 20.

Second, optional reasons theories capture an important aspect of our practical reasoning that the anti-DBR view, as I have laid it out here, does not. That is, they make sense of the fact that there does not always seem to be a single correct choice as to what an agent should do in a way that the plain anti-DBR view does not. As we shall see in the next chapter, my focus is on how we go about deciding among the alternatives that are open to us, once we know that they *are* open to us. The options before us may involve many combinations of enticing and peremptory reasons in justifying and requiring roles, or they may simply be incommensurate (I am content to leave the details open). For the moment, the lesson of this section is that there are ways of relaxing the “requirement” part of the anti-DBR view of reasons to accommodate situations in which reasons seem optional to us and we have some freedom to choose. Nevertheless, as I will show in the next chapter when I develop the puzzle that motivates my view, even optional reasons theories do not go far enough in their flexibility to accommodate certain wrenching cases of practical dilemmas. I diagnose this problem as stemming from a failure to drop the “discovery” aspect of the anti-DBR view of reasons: the view that reasons are wholly objective—i.e. desire-independent—and can only be discovered by an agent, rather than created by him. Optional reasons views suggest that the most common picture of a deliberative situation is one in which there are multiple eligible options. In such cases, what is to pull us toward one option or another, normatively speaking? In the next chapter I argue that unless this is addressed, an important gap remains in our understanding of practical reasons. Eventually, of course, my purpose is to fill in that gap.⁸⁰

⁸⁰ In a forthcoming paper, Ruth Chang defends the existence of what she calls “voluntarist” reasons, which also fill in the gap in our understanding of practical reasons. Her view has much in common with mine, though we come at the problem from different angles and differ on some of the details.

Chapter 3: A Puzzle for the Anti-DBR View

In the last chapter I discussed an “anti-DBR” view of reasons, along with some variations on it that take account of our intuitions about common cases. I ended that discussion with the claim that none of the variations described there provides an adequate explanation of how agents reason in certain practical conflicts. In this chapter (and in the appendix) I describe the conflicts I have in mind in detail, giving examples intended to illustrate just where the story of reasons I told in the last chapter falls short. The conclusion will be that there is a problem with the anti-DBR view’s discovery model of deliberation.¹

3.1 *A Case*

I take as my paradigm case the one I introduced in Chapter 1: the case of John Smith in the film *The New World*.² To summarize: Smith, a born leader, falls in love with Pocahontas so that when he is offered a royal commission to head an expedition in search of the Northwest Passage, he must make a terrible choice between the woman he loves and the life of adventure to which he is called. I have stipulated that, as a matter of reasons independent of what he ultimately cares most about, there is no *best* choice for Smith; but the decision before him is of too much import for him simply to pick an option arbitrarily.

In telling this story, I have emphasized that Smith faces a choice in which it is not clear what he has most reason to do, at least if we hold (as the anti-DBR view does) that reasons must be independent of what an agent wants or wills. This is not just a matter of epistemology; we can suppose that agents like Smith have thought of everything and still cannot decide, for the matter is unclear on the level of reasons. That is, the problem is not

¹ For ease of exposition, throughout the chapter I speak specifically of the anti-DBR view. Since, however, most of what I say hinges on the claim that all reasons are “desire-independent,” an aspect retained by optional-reasons variations on the anti-DBR view, this locution is meant to encompass the optional-reasons variants as well.

² See Appendix for alternative cases.

that there is some truth about what the agent should do and he cannot see it; rather, we may suppose that he can see all of the relevant considerations and there is no truth of the matter as to what the reasons indicate.³ In this sense, perhaps, we are in the territory of optional reasons: Smith may rationally choose either way.

Such “evenly poised” or optional reasons are thus one distinctive feature of the cases I have in mind. By “evenly poised” here I do not mean to imply commensurability or comparability; the reasons may be comparable and perfectly balanced, or they may be incomparable in the sense that neither side outweighs the other, but neither are they equal.⁴ As we saw in the last chapter, there may be different diagnoses as to why both options seem perfectly acceptable, from the point of view of practical reason. Thus, the difficulty of these cases need not rest on any claim about whether the reasons on each side balance one another out perfectly.

Second, the example is also different from cases in which the agent’s desires and preferences yield indirect reasons. For example, if I am deciding between an apple and an orange, the fact that I prefer apples, or that today I want the apple, gives me a reason to choose the apple. Putting it this way is misleading, however. The preference itself does not provide the reason, for the preference seems to result from the fact that the agent would enjoy the apple more. If this is the case, the enjoyment is the reason to eat it, not the preference.⁵ But there is more than this going on in the examples I outline. In our case, the agent’s preferences may indeed matter in this indirect way, but (we can stipulate) these reasons are taken into account in the weighing, and the agent is *nevertheless* wondering how he is to proceed.⁶

³ This may be because it is unclear from any point of view how the reasons weigh against one another. Alternatively, it may be that under particular ways of framing the question, it *is* clear—but there seems to be no one point of view best supported by reason. See Chapters 4 and 5 for discussion.

⁴ Joseph Raz (1999) and Thomas Nagel (1979/1991) blame the problem with this sort of case on incommensurability or incomparability. Ultimately, I do not think that incomparability as such is at the heart of the problem, though it may often be present. As I argue below, the heart of the problem lies in what is at stake in the choice.

⁵ This is the Pettit-Smith point in “Backgrounding Desire.” See previous chapter.

⁶ In putting the point this way, I overstate it somewhat in order to make the intuitions clear. This overstatement invites a question as to whether the self-creating reasons I defend in the next chapter function only as tiebreakers. I do not ultimately intend them to be so restricted, and in chapter 6 I will address this question. For the moment, because I think the clearest way to motivate the problem is in terms of evenly poised reasons, I will continue to frame the discussion as it is here.

(Note that it is acceptable, even on the anti-DBR view, to follow a preference in a situation like this. If reasons are evenly poised, then the anti-DBR view hold that one may go either way, rationally speaking. And in that case it is perfectly acceptable to follow an inclination even though it does not strictly speaking provide a normative reason. Now, if one's inclinations are also evenly poised, then one has a further problem—and this will be a problem even for the DBR view. This is why Smith has a problem on either the DBR or anti-DBR view.)

The other major feature of the example is that the situation is momentous for the agent who must decide. The decision at hand is not just a matter of preferences in the sense of decision-theoretic preferences; the motivations in these cases run much deeper than mere preferences. Not just any sort of moment is enough to make these cases distinctive, however. Consider a case suggested by Thomas Nagel⁷: you are a cancer patient faced with a choice between having surgery and undergoing chemotherapy, and your doctor cannot predict with confidence which will be more effective. Now, a great deal rides on your decision; it could be a matter of whether you live or die. But this does not seem like the same kind of moment as the decision facing John Smith. While the patient's choice may be hard because of the pain of carrying out whichever alternative he chooses (and because of the fear and sadness of the situation), he may be willing (if reluctant!) to settle his decision arbitrarily because it is not a matter of deciding what is most important to him. *That* question is already settled. For Smith, however, it is not settled.

As a result, what is at stake for someone in a situation like Smith's is not only what life he will lead, but that *he* lead it himself. That is, part of what's at stake is the agent's authorship of his life. This is perhaps the more important characteristic of Smith's problem, as it helps to distinguish Smith's case from the cancer patient's, even though a great deal is at stake for each of them.⁸ Flipping a coin, for Smith, would express indifference to his own authority, and allow him to drift through his life in a way that, as I will eventually argue, is out of the question for him. He must settle his question in a conscious and deliberate way.

Consider the nature of the situation. Smith's motivations stem from different sources: on the one hand, his love for Pocahontas and the good he could do as an

⁷ "The Fragmentation of Value" (1979/1991).

⁸ I will say more about this in Chapter 5.

ambassador between peoples; on the other, the ambition to do service for his country and become a famous discoverer. The reasons on both sides seem quite good, neither side seems clearly weightier than the other, and yet they do not seem to weigh the same—Smith is neither decided *nor indifferent* (if we take indifference to indicate that it doesn't matter to the agent how the choice comes out, or that the choice itself doesn't matter). This is a characteristic mark of incomparability, but I do not think that it is the essential feature of the situation; the lack of indifference is the result not of a kind of inequivalence between the options that seems to put them on different scales, but rather because of what is at stake. What matters here is that either path Smith takes will amount to a reorientation of his life, because his action expresses his dedication to one set of reasons (and hence a certain ideal or set of values) over another. I will return to this point in the next chapter to discuss what it is to “reorient a life” in more detail.

So the difficulty of true practical conflicts—as opposed to mere difficult decisions—stems from the fact that the decision the agent makes will express a priority of value that reflects not just a change in the way her life will go, but in the path she herself *takes* in her life. They are cases in which, as I will phrase it in later chapters, the boundaries of the will have become unfixed. In short, her choice is a choice of who she is to be. Such changes run deep, and they lend a particular kind of moment to some decisions. That moment can be distinguished from the moment of making a choice that will affect the rest of one's life but not necessarily change it internally, such as choosing a cancer treatment.

It is important to note that the gravity of a life change that I have just described is not enough by itself to distinguish our examples from other kinds of case. For some decisions are momentous in a self-defining way without being true practical conflicts of the kind in our examples. An agent who decides out of greed to embezzle funds from his employer, for instance, makes a deliberate decision that shapes who he is. And he may well know this ahead of time—but if he does it out of greed, he does not take his becoming a criminal as a reason to embezzle.⁹ I have underdescribed this case, but I think we may fairly assume that it is not a practical conflict in the same way Smith's is, because in general there

⁹ Of course, some people do commit crimes precisely because it will make them criminals; consider someone who is required by a gang or the mob to kill someone as a rite of initiation. He would perform the act precisely in order to belong. But this need not be the case, and many of the people who commit crimes are doing it for reasons they take to be strong enough to do something criminal even though they desire strongly that they not be caught and known as a criminal. (Thanks to Elizabeth Anderson for this point.)

are good reasons not to act out of greed, and (unlike Smith) the embezzler goes wrong, either by not fully considering the reasons not to steal, or by failing to be moved by those considerations. Similarly, a friend who decides to intervene in his friend's destructive behavior may have a momentous and difficult decision on his hands, but there may be no doubt as to whether taking action is the thing for him to do. Thus, the cases which are true practical conflicts of the sort that interest me require *both* self-defining moment *and* indeterminacy as to what all of the relevant, will-independent considerations indicate is to be done.¹⁰

I will claim that an agent in such a position can gain some traction on her problem by reframing it as a question of “who she wants to be” rather than “what to do.” These two questions are slightly different ways of framing the same issue, distinguished more by how the agent thinks about the problem than by the considerations that are relevant. In asking “what to do” an agent seems to be wondering more directly about what state of affairs to bring about. But as Elizabeth Anderson contends,¹¹ states of affairs are only extrinsically valuable; they are valuable as expressions of what an agent values intrinsically, such as a person, animal, landscape, or thing—or the direction of her own life, as in cases like Smith's. States of affairs are simply means to expression of valuation, and expression of valuation is expression of one's self.

In thinking about who to be, an agent puts her values and ideals up for scrutiny quite centrally (more, perhaps, than she might when framing the question as “what to do”). And in our everyday lives, I believe we have a tendency to think that our very personal attraction to or repulsion from an ideal of how to live provides us with reasons for making that choice. “I couldn't face myself if I...” or “I just can't make myself do that” or “I don't want to be the person who...” are familiar ways of explaining choices. It is important to note that although we can understand these expressions as expressing moral compunction, they may also express non-moral judgments.¹² A gay person, for example, may decide to come “out” because he couldn't face himself if he stayed in the closet. Such an act is identity shaping, as well as an affirmation of a pre-existing identity. This seems like a perfectly good reason for his decision. So if, as in Smith's case, the will-independent reasons do not yield clear

¹⁰ Though see later discussion of when the boundaries of the will may become unfixed (section 4.2).

¹¹ Anderson (1993, p. 24).

¹² Again, thanks to Elizabeth Anderson for this point and example.

guidance as to what to do, why not endorse the intuition that these sorts of “excuses” are legitimately action-guiding after all? I will elaborate these last points later. Presently I wish to turn to related but distinguishable cases in order to bring out more clearly what I’m getting at.

One note before I turn to these contrasting cases, however. Recall that “optional” reasons are meant to be the sorts of considerations that one can take or leave and still incur no criticism, and yet they do justify a choice when taken. Perhaps we are to understand “I just couldn’t” or “I don’t want to be the sort of person who” as optional in this way: that it makes me such-and-such a kind of person is an optional reason for going with one alternative rather than another, from the point of view of rationality. But as I intend to show, this doesn’t seem to capture the first-person feel of the choice; when what is at stake is the choice of a whole life, the reason isn’t just optional. It is something compelling about who the person is. Few people are concerned with “being rational” per se;¹³ they are, however, concerned with who they are. So in the sense of justifying by making an option eligible or intelligible, these reasons—what I will be calling “self-creating” reasons—are optional, since they make each alternative rational. Thus, an agent could rationally decide arbitrarily which alternative to take. But an arbitrary decision, even when rational, is inadequate for someone in Smith’s position because too much is at stake. Smith will need a justification not just for choosing one of his alternatives, which the optional reasons provide, but for choosing one *over* the other. A self-creating reason isn’t “that it will make me such-and-such a person” so much as “that *I want* to be such-and-such a person (rather than another).” The next two chapters elaborate and defend this claim; before I begin that task, however, I consider contrasting cases in order to get clear on just what defines Smith’s predicament.

¹³ (Except decision theorists and philosophers, perhaps.)

3.2 Moral dilemmas

The case of John Smith bears some important similarities to other cases that are somewhat problematic, or at least puzzling, for the anti-DBR view of reasons. In particular, they have features in common both with cases of “moral dilemmas,” and with cases in which integrity is an important issue. In this section I compare my cases with moral dilemmas, and in the next I compare them with integrity cases.

In Baroness Orczy’s *The Scarlet Pimpernel*, Marguerite Blakeney is the clever, beautiful French wife of Percy Blakeney, a splendid British aristocrat reputed to be fashionable but dull. She and Percy have, however, become somewhat estranged from one another. Marguerite’s beloved brother Armand St. Just, who raised her and with whom she is very close, works for a gang of young British aristocrats who plot and execute daring rescues of the endangered French aristocracy from the Reign of Terror. The gang’s leader, known only by his signature flower, a scarlet pimpernel, is greatly admired among the rich on both sides of the English Channel—Marguerite included. Through no fault of her own, however, a fanatical acquaintance of Marguerite’s puts her in a terrible position: she must help him discover the identity of the Scarlet Pimpernel, or he will denounce her brother Armand and leave him to the guillotine. Knowing that the gang can function without Armand but not without the Scarlet Pimpernel, Marguerite must now choose between the brother who (she believes) is the only person in the world who loves her, and the mysterious, heroic leader of a gang that takes immense risks to save innocent lives.

Marguerite is caught between two alternatives that are each backed by powerful reasons. On the one hand, her love for Armand has been the one sustaining presence in her life for a very long time, and their bond is a weighty reason for her to protect him. Furthermore, as a member of the Scarlet Pimpernel’s band, Armand is doing good work. On the other hand, the Scarlet Pimpernel is a crusader in a cause in which she, too, deeply believes, namely saving innocent lives from the guillotine;¹⁴ and not only that, but in helping discover his identity she would be betraying a hero.¹⁵

¹⁴ There may be some ambiguity in the source of Marguerite’s first reason. It can be seen as stemming from the *value* of human lives in general and heroes in particular, which is a consequentialist way to view the matter, or it can be seen as stemming from some *obligation* everyone has to aid fellow persons. (I suppose something similar can be said about her reasons for saving her brother, though I suspect that the pure consequentialist reading has less plausibility here—it is far more likely that she is acting *for Armand’s sake* rather than because of the value he has to her.) In any case, either way we understand Marguerite’s reason for protecting the Scarlet

Here it is unclear what Marguerite should do as a matter of reasons. Her reasons on each side are quite strong. One might object that it *does* seem clear what she should do: she ought not betray the Scarlet Pimpernel because he is doing so much good, and sometimes we must sacrifice for the greater good. But despite its initial plausibility, this claim does not do justice to the situation. For one thing, it relies on a controversial consequentialist view that impartial reasons should take precedence over personal ones. Although more lives will be saved if she protects the Scarlet Pimpernel, it is not *obvious* in this situation that this is a reason that outweighs the fact that her brother's life will be saved if she betrays the hero. We generally recognize special obligations to loved ones, and although those obligations can be overridden in certain circumstances, it is difficult to tell whether this is one of those circumstances. (How many strangers are we allowed to let die in order to save a brother, child, spouse, etc.?)¹⁶

There is a great deal of controversy over the existence of true moral dilemmas, and I wish to remain officially neutral on the matter, though as I have already said I do think that it can be unclear *as a matter of reasons* (and not just epistemically) what an agent should do. Whether this translates into endorsing moral dilemmas would take some further argument, however, and I will not pursue the matter here. Assuming for the moment that there is such a thing, a moral dilemma is canonically characterized as a clash between moral obligations where it is impossible for the agent to fulfill both, and neither is defeated by the other.¹⁷ If it is wrong not to fulfill one's obligations, the agent in a dilemma faces a situation in which any choice she makes will be morally wrong. So Marguerite is in a classic moral dilemma, for it seems on the face of it that either way she chooses, she will do something wrong. She does not seem to face optional reasons.

Pimpernel, the reason is impersonal in the sense that it does not stem from Marguerite's special relationship to him (at least, for those who have read the book, not at the time she must make this choice).

¹⁵ No doubt her personal admiration for the Scarlet Pimpernel's cunning and daring play a part here also, but although Marguerite is aware of this, her main motivation is directed at the man himself and the importance of his cause.

¹⁶ If this case does not satisfy as an example of a moral dilemma, consider the case of Abraham and Isaac. Abraham must choose between sacrificing his son and disobeying his God; if we assume that there are good reasons to obey God (other than just the threat of punishment), either way he chooses Abraham will be betraying someone else as well as himself.

¹⁷ This is often alternatively expressed by the claim that a moral dilemma is a conflict between two all-things-considered obligations.

A dilemma like Marguerite's, if it is genuine, could be problematic for the anti-DBR view. If there is more reason to do one thing rather than the other, there is a puzzle concerning whether it would be wrong to take that option, for it may initially seem as though it cannot be wrong to do what one has most reason to do. So either it *can* be wrong to do what one has most reason to do,¹⁸ or there is no real dilemma in the first place. On the other hand, if there is no more reason to choose one option over the other, the official position of the anti-DBR view is that the agent has no choice but to pick arbitrarily, to “plump” for one or the other of the options; and if that is the only thing you can do you cannot be blamed for it, and therefore you do not do anything wrong. But if we suppose that dilemmas exist, then by hypothesis the agent *has* done something wrong, so the anti-DBR view must be incorrect—or at the very least, there is a puzzle here.

Now, rather than accept this as a problem for the anti-DBR view, one could argue that there is no such thing as a true dilemma in the sense given above.¹⁹ As Peter Railton and others²⁰ have suggested, what we call and experience as moral dilemmas may not be restricted to the results of strict obligations, but may arise from many different clashes among personal values, ideals, and responsibilities that in some circumstances (but not all, as Railton illustrates in a series of examples) allow for making reparations as a way of recognizing the importance of an obligation that must be violated because of the circumstances which give rise to the apparent dilemma. Thus so-called dilemmas are not true dilemmas and therefore need not be a count against any particular moral theory (and hence create no problem for the anti-DBR view of reasons, either). Furthermore, for those who take “moral residue” such as the appropriateness of guilt or remorse as evidence that true dilemmas exist, the idea of recognizing lost value or the importance of violated obligations accounts for this appropriateness without undermining moral theory.

¹⁸ This is a theme of Bernard Williams'; a clear but brief statement of it is found in his “Practical Necessity,” in *Moral Luck* (1982), pp. 124-5, and it is prominent in both the title essay of *Moral Luck* and in Chapter 10 of *Ethics and the Limits of Philosophy* (1985). Gilbert Harman (1996, p. 61) has also advanced a similar thesis. Susan Wolf (1982) may be thought to express a similar idea in her contention that moral saints are missing out on some important non-moral virtues.

¹⁹ One could also argue, as Ruth Barcan Marcus (1980, 1996) does, that there are dilemmas *and* that this is not a problem for a theory. I will not discuss this option here because the other is more closely related to my present interests.

²⁰ Railton's 1996 paper is entitled “The Diversity of Moral Dilemma;” the “others” referred to include Christopher Gowans, H.E. Mason and perhaps Mary Mothersill in their 1996 papers.

Despite the merits and interest of the debate over dilemmas, I will not dwell on whether or not there are true dilemmas or whether or not they create problems for theories of moral and practical reasoning. What I want to get at here is how the dilemmatic case of Marguerite Blakeney differs from the cases I have in mind. An obvious difference is that John Smith is not *necessarily* wrong, does not violate any obligations, etc. if he chooses any of his available paths. He may disappoint either Pocahontas or the king, but it is not morally wrong to do so. He probably has responsibilities to both Pocahontas and to the king to (at the very least) explain his choice, but again this says nothing about which choice he should make or what his reasons are. As I said above, Smith's choice looks as though it comes down to a question of who to be or what he cares most about after careful consideration. It is genuinely up to him.

Because Marguerite does violate an obligation no matter what she does, however, unlike Smith her choice cannot come down to a question of who she is in the sense that it is up to her to choose an ideal by which to live.²¹ Her authority to live her life is not a part of what is at stake in her situation. Marguerite's case seems to be moral in a way that Smith's is not, for the most salient features of the case are the duties Marguerite owes to Armand and to the Scarlet Pimpernel. The boundaries of her will have not become unfixed in the way Smith's have. Because of this it is suspect to claim that Marguerite's decision is ultimately about *her*—about what she ultimately cares about—in a way it is not suspect to claim that Smith's is about him. For Marguerite to make this choice on the basis of who she will become seems self-indulgent and based on the wrong sort of reasons, given that the problem is about others and thus at heart a moral one. So the kind of case that interests me—a case like Smith's—is not a moral dilemma in the classic sense, where the agent is apparently wrong whatever he does.²²

²¹ I mean this as a normative claim. Descriptively, Marguerite could well see her plight as a question of who she is to be; no doubt people frame important choices in this way quite often. Although a case can be made that Robert E. Lee more or less backed into his choice to lead the Confederate troops rather than the Union's, prior to choosing he framed his problem as a question of whether he was more a citizen of the United States or of Virginia. Now, since we can see the moral choice in Lee's case as being fairly clearly in favor of joining the Union, it may be that we should not think of Lee's identity-based decision as being a normative one either. Yet it is possible to imagine a case of civil war where the identity question really is legitimate because the moral question is not so clear. See appendix. (Thanks to Elizabeth Anderson for this point.)

²² On broader construals of dilemma advanced by people like Railton and Mary Mothersill, Smith's case may still count as a dilemma, because it involves a clash of values and ideals that are not necessarily moral. This is all right because my interest is only in differentiating Smith's choice from cases of moral dilemma in the narrower, classic sense that is concerned only with strictly moral obligations.

3.3 Integrity

I mentioned above that cases like Smith's are different both from moral dilemmas and from cases in which the agent's integrity is a factor. I turn now to integrity cases.

In Henry James' *The Portrait of a Lady*, Isabel Archer is a bright and independent young woman, an American whose cosmopolitan aunt takes a fancy to her and brings her to Europe to see the world. Isabel wins the hearts of both her cousin Ralph and his friend, an English lord, but, treasuring her freedom, remains single and independent. Not long after, Isabel's uncle dies, leaving a substantial part of his immense fortune to her. Isabel's aunt brings her to Italy where she keeps a villa, and there Isabel meets a man named Osmond. Eventually, she consents to marry him, over the objections of nearly everyone in her acquaintance, and particularly her cousin Ralph and friend Henrietta.

In marrying Osmond, Isabel chooses to give up her freedom for a man who appears to her to be "independent" and "individual," "good enough," "interesting enough," and "clever enough"; he "knows everything, he understands everything, he has the kindest, gentlest, highest spirit." She insists more than once that her reasons for choosing Osmond are that she likes him, and that her ambition now is "to be free to follow out a good feeling." Yet despite the good opinion she comes to have of Osmond, it is Ralph and the others who are right: Osmond is "small," "narrow," and "selfish," a "sterile dilettante."²³ Because of her stubbornness, naïve self-assurance, and this difference of opinion, the match causes a cooling between Isabel and her friends. But after only a few years, Isabel is quite unhappy.

Eventually, Ralph is on his deathbed in England and summons Isabel to come to him. Osmond forbids it, but Isabel disobeys him and makes the journey. This is a turning point for Isabel; it is an act of defiance unusual for a woman of her time, and furthermore she knows it is her chance to separate from Osmond. Once away, she might break with him and never return. The thought is powerfully tempting, and the decision to go involves a gathering of will and the self-defining choice to remain independent from her husband in spirit if nothing else. Yet Isabel has always been a creature of principle, and she feels as though she ought to remain with Osmond and accept the fate she chose of her own free will:

²³ The descriptions in this paragraph are found in Chapter 34.

It was impossible to pretend that she had not acted with her eyes open; if ever a girl was a free agent she had been. A girl in love was doubtless not a free agent; but the sole source of her mistake had been within herself. There had been no plot, no snare; she had looked and considered and chosen. When a woman had made such a mistake, there was only one way to repair it [as she thought]—just immensely (oh, with the highest grandeur!) to accept it.²⁴

Although she does break away from Osmond to visit Ralph, Isabel is concerned with her propriety, with living up to the “theories” she formed.²⁵ Others such as Ralph and Henrietta are concerned with her character and high spirits—the things that lead her to go to Ralph—and how the marriage to Osmond will ruin them.

After Ralph has died, a former American suitor tries to persuade Isabel to run away with him. Perhaps he tempts her when he says that it would be an insult to her “to assume that you care for the look of the thing.”²⁶ He may be right about that, but she does care for the *thing* of the thing, and she sticks to her principles rather than her feelings. She chooses to return to her husband, however cheerless that life is for her.

In choosing to marry Osmond, Isabel makes the “tragic” choice which proves to be her undoing; and in defying him and fleeing to England, she performs the self-defining act of asserting her independence. Ultimately, however, she does not capitalize on this decision and complete the self-constitution it has the potential to be, because she does return to Osmond. In choosing to return when she could escape, then, she chooses to be a woman who remains true to the principles she has, and to face the mistakes she made freely, rather than allow herself the relief of following her true feelings. From now on her life is defined by her choice of herself as Osmond’s wife, true to her principles, rather than as an unhappy woman who fled from her mistake in order to have a chance at a happy life.

Unlike Marguerite Blakeney, Isabel faces a situation in which her authority to lead her life *is* part of what’s at stake here. The contrast between Isabel and John Smith is instead that the boundaries of Isabel’s will never become unfixed.

That is, both in leaving and in returning, Isabel acts so as to preserve her integrity. Her choice is thus different from two other kinds of choices in the same neighborhood. On the one hand, Isabel chooses on the basis of principles—principles which hold for anyone in

²⁴ Chapter 40.

²⁵ (See ch. 51 for a glimpse.)

²⁶ Chapter 55.

her position. That she has chosen with her eyes open and made a marriage vow are very good reasons for her to return to Osmond, and they are the sort of reasons that apply to anyone in circumstances like hers. But this is not, precisely, why she makes the choices she does. Isabel chooses first to assert her independence in order to be true to herself, and later to obey the principles that tell her she must return because they are *her* principles. She would be betraying something important in herself if she did not see Ralph for the last time, but she also finds she would be betraying herself if she gave in to the temptation to leave Osmond permanently and return to America with her old suitor. It is one thing to assert a measure of independence in a contained incident, but it is quite another to escape altogether. She could not live with herself if she did this, and so she resigns herself to the misery she has chosen. Thus, because Isabel's choice is based on the personal character of her principles, it is different from a choice made on principles which hold for anyone because they hold for anyone.

On the other hand, however, Isabel's final choice is also different from one like John Smith's. In both cases, the issue before them is who to be. Nevertheless, they are different in that Isabel acts so as to preserve (or perhaps regain) her integrity; she chooses to preserve a self which has already been constituted by past decisions. Had she not returned to her husband, then her act of defiance in leaving him would have been a self-constituting choice that changed what constituted integrity for her. But since she resigns herself to her marriage, she is not reconstituting herself, but rather acting out of the constitution already established. Smith, however, faces a choice which will be self-constituting. Integrity is not at stake for him because unlike Isabel he has not established just which of his alternatives *is* most important to him. He faces a decision as to what his principles are to be, so his choice is about who he is to be in an inventive sense; Isabel's choice of who to be is made in a preserving sense.

Now, identity-conferring projects and principles can change; why is Isabel's case one of preserving integrity rather than changing her identity? Why claim that she acts so as to preserve her self, whereas Smith is in a position to *form* his? I have claimed that it is clear to Isabel what is more important to her. The misery of her life with Osmond calls this importance into question, as misery is usually a good reason to get out of a situation; but even after the daring act of defiance, upon reflection her misery cannot overcome her attachment to principle. Her journey to Ralph's deathbed gives her the space she needs to

reconsider her identity and the opportunity to change who she is. But she finds that if she gives in to temptation she will betray herself, and so she returns to her husband. For Smith, however, even upon reflection it is not clear what is more important to him. He does not examine himself and *discover* that the life of an explorer is the life he must lead; instead, he discovers that neither the commission nor Pocahontas is antecedently more important to him. Unlike Isabel, he must *make* one more important by making up his mind. He must invent himself.

3.4 *Tension with the anti-DBR view*

Now that I have defined cases like Smith's by contrast to related kinds of cases, I would like to concentrate more carefully on the problem that cases like Smith's, and those that appear in the appendix, pose for the anti-DBR view and the discovery model of deliberation.

As we saw in the last chapter, on the anti-DBR view the kind of "reason" expressed by "I just couldn't" and the like is no reason at all. Part of the motivation for thinking this is that such wants and willings are responsive to reasons, which might lead us to believe that wants do not *justify* actions, though they might explain them. Desires, for example, can be misplaced—one may want revenge even while knowing that it is better (that there is most reason) to forgive and forget. We often do our best to adjust our wills to fit our reasons (though we may not always succeed), precisely because we think that only a will that is backed by reason is truly justified. Without rational backing, we can want all sorts of crazy things.

A second, and related, motivation for thinking that really wanting something is no reason at all to pursue it is that such wants and willings are normally transparent to their objects. Agents do not think to themselves about how good it will be to have their desire satisfied or their will fulfilled; they think about how good it will be to see their brothers or captain their own ships.

Thus, according to the anti-DBR thesis, deciding on the basis of an inclination is arbitrary, given that desires, inclinations, and so on are viewed as transparent to the reasons that support them. That is, since an inclination contributes nothing above and beyond the reasons there may be to have that inclination, on the anti-DBR view, there may be an *explanation* as to why an agent decides as she does, but it is not a *justification* of that decision. So deciding to act on an inclination in a case like Smith's amounts to an arbitrary picking on

this view. Of course, this doesn't seem to be the full story, and in the next two chapters I will argue that it is not. Although I think that the anti-DBR view and the model of deliberation it embraces are correct for the majority of practical questions that face agents in their everyday lives, in this chapter I have presented the example of John Smith to test the adequacy of this view. Smith gives us reason to wonder whether perhaps what an agent wants, all-things-considered, sometimes provides a reason after all. If all of the available desire-independent considerations leave the agent wondering what to do, couldn't we admit a desire-*dependent* consideration as a way to break the impasse?

Here is a reason to think we could, and should. If all of the available will-independent considerations leave a responsibly deliberating agent²⁷ without a clear course to follow, according to the anti-DBR view it seems as though she must simply pick one of her options arbitrarily; after all, if will-independent considerations are the only available reasons, then in cases like our examples, reasons have run out. Still, the agent must do *something*, so why not just pick arbitrarily? Yet something about picking when the stakes are so high seems repugnant. Why is this? Part of the problem lies, I believe, in the fact that the discovery model of deliberation, with its air of detachment and objectivity, represents the wrong sort of view of oneself. The anti-DBR view and the discovery model of deliberation that stems from it suggest that agents view themselves only as practical reasoners, rather than liver of lives. But we are unavoidably tangled up in some of our deliberations.

John Smith's problem as I have sketched it here is a conflict of cares (as I will elaborate in Chapter 4), in that it arises in part because of commitments the agent has to values on each side of the conflict that arise from care. But the fact that the dilemma arises from previous cares or commitments is not an essential feature. What matters is that the choice the agent *now* faces is a choice that will shape his life, so that different things may count, or count more, as reasons than they would have had a different life been chosen. For actions tend to express an agent's valuing and ideals,²⁸ so the choices made by people in situations like Smith's express a commitment to one set of reasons over another, that is, to

²⁷ This loaded term deserves a clarificatory footnote. Just what falls under the heading of "responsible deliberation" is a matter of much discussion, of course. I think something along the lines of Rosati's "normal ideal conditions"—all available or reasonably attainable information, clear head, lack of coercion, etc.—are a good basic description (Rosati 1995, 1996). See also Scanlon (1998, chapter 1.12) for discussion of the review process on purported reasons.

²⁸ See Anderson (1993).

one ideal over another. Thus, these choices will change who they are in ways they can foresee and for which they can take full responsibility. As a result, such agents are unable *merely* to pick among their options in an arbitrary way, because a great deal—a whole life—is at stake.

To make this claim clearer, I would like to make a few remarks on an important distinction. I have pointed out that in the anti-DBR story, an agent who has reached an impasse of (desire-independent) reasons has no recourse other than simply to pick among the open alternatives. For my purposes, I mean “picking” to be opposed to “choosing”;²⁹ to pick an alternative is to decide arbitrarily, without any particular guidance from reason, whereas to choose is to decide on the basis of reasons.³⁰ In this sense, choice has an expressive element: a *choice* of an alternative expresses an agent’s endorsement at some level. Picking expresses indifference, certainly between the two alternatives, and sometimes toward the decision itself.

It is important to differentiate the use of a picking procedure from what we might call “pure” picking. “Pure” picking is something one does when one gives up on making a reasoned decision and throws oneself at one alternative or the other simply for the sake of making a decision. Pure picking can express indifference toward the decision as well as between the alternatives. It is what Buridan’s ass should have done, rather than starving to death between its two piles of hay—such a decision is simply not worth starving over. Pure picking is a way of making a decision when the relevant reasons do not yield a determinate verdict as to what to do. It is a decision to go with a particular alternative, but it is not a decision to go with one alternative *over* another.

As in “pure” picking, a picking procedure is a way of making a decision without really making up one’s mind, without actually choosing some option over another. Buridan’s ass could have flipped a coin to decide between its haystacks;³¹ we also draw straws, roll dice, throw darts, and so on, both literally and figuratively. But note that a picking procedure *need* not be an instance of what I have called “pure” picking, as it can specifically express a *lack* of indifference toward the decision itself. For example, Adam might be torn between two

²⁹ This is a distinction originally made by Ullmann-Margalit and Morgenbesser in “Picking and Choosing” (1977).

³⁰ At least as seen by the agent. Agents may not always take all relevant considerations into account, but I wish to count them as “choosing” in the sense described here when they are making some effort to deliberate.

³¹ Well, it could if it had opposable thumbs.

friends' requests for help. Both have equal claim to his help, and he wants to treat them fairly and show no favoritism, but he cannot do both tasks. So after unsuccessfully looking for ways to change the situation so that he can help both friends, he may flip a coin and explain to his friends that he had a conflict, and no good reason to choose one over the other, so in order not to insult anyone, he simply flipped a coin so that neither of them would have any basis for hurt feelings. Next time around, he might say, he'll choose the loser of this decision's toss. The picking procedure Adam uses here cancels any implication that his decision favors one friend over the other, so he does justice to the reasons he has to be fair to both of his friends and shows that he takes the decision seriously. But because of this, Adam does not pick in the "pure" sense, though he uses a randomizing device.

The point I'm driving at with this aside is that on the anti-DBR view, agents in situations ranging from the trivial to the immense are forced to pick arbitrarily when the relevant standard reasons—reasons *to* choose a given option, independently of what the agent wants—do not direct them to one alternative as the most appropriate one. This is not a problem when the decision to be made is which laundry detergent to pull off the shelf, or whether to have chocolate or strawberry ice cream. These are decisions that hardly matter, so being forced to pick seems unimportant; the agents can remain indifferent to the decision itself. Picking is less and less satisfactory, however, the more meaningful the decision is.³² As agents we may feel comfortable throwing darts about laundry detergent, but we are less comfortable with it when it comes to choosing careers, schools for our children, personal relationships, and the like. In such decisions we may sometimes reach a point at which we feel we have examined all of the relevant reasons and have no basis for choosing one thing over another, and then perhaps we settle for picking as our only way of making the decision. In the cases I have outlined in this chapter, however, where an agent's very identity seems to be at stake, even the use of a picking procedure may feel quite inappropriate. John Smith cannot say to himself, "Oh, well, I don't know what to do, so I'll just flip a coin." Or rather, he *can* tell himself this; but he *will* not. Something certainly seems wrong if he does.

One may think that although Smith will not make his decision flippantly (so to speak), he may nevertheless view the situation in a way analogous to Adam's above: he may flip a coin because it is the only responsible thing to do. I cannot deny that this is possible,

³² A clearer account of why will be given in Chapters 4 and 5.

and perhaps every tough decision *could* be made by a picking procedure even if it does not involve pure picking. Nevertheless, I do not think this is the case. Someone in circumstances like Adam's flips his coin because he wants to express his commitment to fairness, and to cancel any implication that his choice is a choice of one of his friends *over* the other. In Smith's case, however, fairness to others is not at issue, so a coin flip cannot be a deliberate expression of that value. And given the nature of his options, he cannot cancel any implication that his choice is a choice of one *over* the other. To attempt to do so seems to me to (attempt to) express an indifference he patently does not have.³³

Furthermore, if an agent makes all of her difficult choices by arbitrary procedure, then she is not really owning them; the point of Adam's coin flipping is precisely to *disown* the decision. Adam has a good reason to do this, and his decision takes a stand on fairness even though it purposely takes no stand on the decision between the two alternatives. But if coin-flipping is an agent's habit, then she is not really ever defining herself, and she is not doing justice to her agency. There seems to be something wrong with this, in a non-moral sense; this agent isn't doing justice to the considerations, and doesn't seem to *care* that she is not doing them justice. Something may be wrong with this agent, but I see no reason to deny that there could be such a person. Most agents are not like this, and all I need is that there be some agents who will not resort to flipping coins in cases like the one I have outlined here.

Nevertheless, more can be said about what is wrong with ducking a decision in certain kinds of situation. One possibility is that perpetual coin-flipping is a threat to the agent's unity as an agent. After all, empirically, there is a need to avoid a fracturing of the self that can result from arbitrary decision making.

It is useful here to draw on a three-way distinction among different kinds of arbitrariness made by Connie Rosati in her dissertation.³⁴ One kind of arbitrariness arises when there is no uniquely correct outcome of rational deliberation and either option is acceptable, from a rational point of view. This kind is not particularly disturbing because alternatives are backed by *some* good reasons. This kind of arbitrariness does face Smith because either of his paths seems to be correct in this sense; still, for him, each outcome also

³³ If this example does not satisfy, the reader might look to his or her own experience to find a case in which coin flipping, even when not a *pure* picking, seemed inappropriate.

³⁴ Rosati (1989, p. 117ff, fn. 1).

seems to be *incorrect* in light of his commitments to another alternative. (Note that this “incorrectness” is not just the result of the impossibility of choosing both alternatives. What is wrong in Smith’s case is that there are very good reasons *not* to choose each side, over and above the fact that he cannot choose both. This does not seem to be the case when there is less at stake.)

Second, there is an arbitrariness born from acting on desires, cares, commitments, etc. that stem from psychological factors that arise outside of our control and are “arbitrary” as a result. These motivators are not rational in a strict sense, but this is not a problem unless we have reason to call the motivating factors into question; for they are answerable to reasons. This arbitrariness is also present in our example, but it is not a problem because, as we have stipulated, neither decision is rationally criticizable.

It is the third kind of arbitrariness that may present the problem. As Rosati puts it, the third kind “leaves us without a rational standard that determines the correct option between conflicting alternatives, thereby threatening us with fractured lives.” I take it that by this she means that the agent who decides arbitrarily in this sense does not actually *resolve* the conflict between his options; the option not taken will still have its rational (and emotional) pull, and when the decision is important the agent will be perpetually unsatisfied, unable to quiet a part of himself because he never quite feels justified in having taken the path he did.³⁵ The phrase “making up one’s mind” seems illustratively apt here. An agent who decides has made up his mind in one direction, into one unified action. But an agent who cannot make up his mind is “torn” between alternatives and hesitates because he is not “of one mind” about what to do. He feels as though there are two (or more) parts of him facing in different directions. And without a rational standard by which to make up his mind the two parts never fully reconcile. A part of him will always be looking elsewhere and not be fully present in the life he has chosen. Thus, an agent cannot truly *decide* in the face of such arbitrariness but can only go ahead.

As a matter of fact, this kind of fracturing is something agents strive to avoid, and I think this is reason enough to seek a way to resolve a conflict rather than simply move ahead. Let me stress that a “resolution” here is not necessarily a happy ending where the agent gets everything she wants in the end. Rather, by “resolution” I mean something much

³⁵ This fracturing phenomenon may well play a part in Eric O’Neill’s decision not to remain an FBI agent after bringing down Robert Hanssen. See appendix.

weaker: the agent feels justified in choosing as she does, but this does not exclude regret over the option she does not choose.³⁶ There is a sense of regret that is compatible with the belief that one has chosen well in the end. The wholeness that results from a resolution in this sense is something normal agents strive for, out of an implicit need to make sense of ourselves and our actions.³⁷ We need things to orient ourselves around, or else we cannot act as agents at all.³⁸

Nevertheless, to explain what is objectionable about the anti-DBR view's position on conflicts like Smith's we need a stronger argument than can be provided by a contingent, empirical psychological need for self-unification. The idea is that someone who doesn't make the decision *for himself* isn't fully participating in his life, and there's something wrong with this (in a non-moral sense).

That is, the problem for Smith probably isn't precisely a unification or integrity question; rather it has more to do with ducking a decision.³⁹ Ducking a decision is not a matter of actively disowning it in the name of another value; it is simply shirking it. There may be two ways of doing this. One is passive: simply ignore the situation and the choice it presents, take no action, and let the world take its course. For instance, Smith's making no active decision would amount to a passive "decision" to remain in Virginia: he could shirk the responsibility of making a choice and let the ships depart from Jamestown without him. Even if he ducked the decision in this way Smith would have to reorient his life, since he would in fact give up something he cares about and would feel the loss. He would no longer be in a position to reorient his life by *deciding* what is to be most important to him, however. The necessary reorientation would be *reactive* rather than *proactive*. It would be more like falling than jumping.⁴⁰

³⁶ See Chapter 5 for a short discussion of regret.

³⁷ This is most explicitly David Velleman's view, but it also underlies the work of many others, including Bratman, Korsgaard, Anderson, and Jaworska. See related discussion of making sense of our lives through narratives constructed around the objects of care in the next chapter.

³⁸ See Jaworska's (2007) discussion of Bratman and of Damasio. The basic idea is that the only way to be a coherent agent and be said to have any identity is to be integrated over time. This requires projects that engage the mind and will of an agent.

³⁹ Part of what seems undesirable about ducking a decision may indeed be the fracturing it causes, but as I will discuss, I think the heart of the problem lies in a related but somewhat more focused issue.

⁴⁰ Incidentally, I think grief is like this. It is a reaction to the loss of something important, and you find yourself having to reorient your life and make sense of it without this important person or project. We feel such a loss much more acutely than we do a loss that happens with regret but by choice, I think, because of the difference

The other way of ducking a decision is more active and has the external appearance of a true decision. It is the “self-launching” kind of decision that Sartre seems to advocate in his story of the young man who must choose between caring for his mother and joining the French resistance during World War II.⁴¹ But according to Sartre, the young man cannot decide, in the sense of resolving his conflict, but can only go ahead. Sartre writes that the young man can only make his action the right one after the fact, for “existence precedes essence.” I find it difficult to understand such self-launching fully, though I am willing to grant that it is possible. Still, the reorientation of the young man’s life in self-launching is reactive rather than proactive. If indeed he simply acts his way out of his decision, he will have done so by turning off deliberation and ducking the tough choice, rather than engaging with the situation. It is likely that this rather careless way of proceeding will come back to haunt him later on. He may well find himself someday wondering regretfully how he ended up where he did.

Both of these modes of shirking choice seem less than desirable, and this is not only because of the dangers of *post hoc* regret. While the nature of the conflicts that these characters face does not admit of a truly satisfactory, “happy-ending” kind of resolution, it seems that both passively letting the world go by and actively launching without deciding are inferior ways of living one’s life, precisely because they do not seem to be ways of *living* one’s life. Such lives are not guided by the agent. To be guided by the agent in the most robust sense, a life must be actively and continually chosen by an individual. Such an active choice is much less likely to fracture the self because it will involve a kind of (regretful) resolution.⁴²

The idea that there is something wrong (in a non-moral sense) with less than full participation in one’s life has a history. One of the more famous formulations comes from John Stuart Mill in the third chapter of *On Liberty*, “On Individuality, as One of the Elements of Wellbeing.” In this chapter Mill argues for individual decision and originality, against conformity for conformity’s sake, as the thing that makes the liberty that is his main subject worth having. The argument is that individuality is a good in itself, and not merely instrumentally, because there is intrinsic worth in the development of humans’ special

of control. It is easier to come to terms with a loss you are prepared for, just as it is easier to land on your feet when you jump rather than fall.

⁴¹ In “Existentialism as a Humanism” (2000). I discuss this example further in the appendix.

⁴² This is not to say that such an active choice could never be the wrong one. Even the most careful deliberators can sometimes misunderstand the reasons they consider. See Chapter 5 for a further discussion.

faculties. Liberty is to serve individuality, not the other way around. The problem he abjures is not the fact of conformity, but conformity for no other reason than that one wishes to blend in. It is one thing to adopt customs because one finds them suitable after exercising one's discrimination, but it is another to adopt them out of "ape-like imitation." The danger Mill sees is that conformity for its own sake leads to an atrophy of the "highest and best" in individuals; those who do not exercise their discrimination lose it and become no more than animals.⁴³

It is relatively uncontroversial to claim that it is intrinsically good to develop one's faculties, at least with the qualification that they be exercised in the service of some worthy purpose. Mill's particular point lies in where he locates the source of this good. As his chapter title suggests, the point is that the self-development that constitutes individuality is an essential element of an individual's well-being. This is not the place for a full-scale discussion of the normativity of well-being,⁴⁴ but let it suffice to observe that we do take well-being, both our own and others', to be normative. Thus, according to Mill, self-development is not important for the goods it produces, such as the advancement of knowledge or society, but simply as part of what it is to live well as a human being.

Mill takes Wilhelm von Humboldt as a source for his argument. In *The Sphere and Duties of Government*, Humboldt is ultimately concerned with the design and role of government, and he takes the place to start such an inquiry as being an investigation of the worth of the individual, since he takes the point of government to be the protection of individuals' interests. Humboldt traces his idea back to the ancients, finding there an emphasis on virtue and individuality that the Industrial Revolution threatens; he writes: "The ancients devoted their attention more exclusively to the harmonious development of the individual man, as man; the moderns are chiefly solicitous about his comfort, his prosperity, his productiveness. The former looked to virtue; the latter seek for happiness."⁴⁵ The difference between these is subtle; even the ancients believed that virtue was essential to, if not fully constitutive of, happiness. But the notion of "happiness" Humboldt complains of

⁴³ Harry Frankfurt's "On the Usefulness of Final Ends" in Frankfurt (1999) contains a related discussion of the need for us to value certain ends intrinsically in order to live lives of meaningful activity. In this way, "final" ends are simultaneously instrumental (to our living meaningful lives), and "instrumental" ends are final in that they become valued intrinsically. See also Darwall (2002), ch. 4.

⁴⁴ See, e.g., Darwall's *Welfare and Rational Care* (2002) as well as work by Connie Rosati (1995, 1996).

⁴⁵ Humboldt (1854, Introduction).

is different from the eudaimonia that comes from virtue, and this is the substance of his complaint. The problem that both Mill and Humboldt are worried about is the shift in the notion of happiness, a shift that results from the displacement of virtue from a good in itself to a good instrumental in procuring a more material notion of happiness.

Humboldt and Mill each elaborate the idea that it is not only one's ideas that should be one's own, but one's feelings as well. The two can be powerful forces if developed together; the idea is reminiscent of Aristotle's picture of the virtuous person with well-tuned sentiments that fully belong to him in virtue of conforming to reason. Mill writes that "[a] person whose desires and impulses are his own—are the expression of his own nature, as it has been developed and modified by his own culture—is said to have a character. One whose desires and impulses are not his own, has no character, no more than a steam-engine has a character." If character, as opposed to happiness, is the central intrinsic good, as both Mill and Humboldt advocate, then Humboldt's conclusion follows: "The highest ideal, therefore, of the co-existence of human beings, seems to me to consist in a union [of thought and feeling] in which each strives to develop himself from his own inmost nature, and for his own sake." Essential to such development is the making of decisions for oneself. And thus, making decisions for oneself is essential to one's truest well-being, on this view.

What is it to make a decision *for* oneself? The difference between making one's own decisions and letting them be made for one is much like that between steering a boat and drifting in it. In the latter case boat and passenger end up where wind and water take them; in the former, one has some say in where the boat ends up. No one can avoid having to respond to what the world presents, of course, and just as the boat's pilot will sometimes be contending with forces greater than he can control, much of the time the decisions we are given to make do not come about through any doing of our own and take us places we do not plan to go. People often end up living lives they would never have anticipated; this can be by choice, but even then it is by choice given the available alternatives among unexpected obstacles and opportunities. Nevertheless, the difference between being in some measure of control and being in no control at all is discernable.⁴⁶

⁴⁶ To continue the boat metaphor, it does seem to be true that we fare best when we're in control of the boat rather than jostled around by the waves on which it rides (certainly it cuts down on seasickness). So it seems plausible to think that it is indeed part of one's well-being to be engaged with one's life and to make one's own decisions rather than let the world make them for one. Most of us like to be prepared for eventualities; consider the fact that the death of a loved one is much more difficult to adjust to when it is sudden than when it is expected.

Just what is the difference? The analogy between decision making and steering a boat is illustrative. What is the difference between steering the boat and drifting in it? The obvious answer is a kind of engagement. In steering, the pilot has a hand on the tiller and is on the lookout for rocks, waves, and weather. She anticipates the way the boat will move as a result of these things and calculates the moves she needs to make in order to make the ride as comfortable as possible and keep the boat seaworthy. Most of all, however, there is direction. There is a point to all of this activity, whether it is to end up in a particular place or simply to survive and reach shore. A drifter, by contrast, has no direction and hence no need of activity because it is not particularly important where she ends up. If the drifting goes on for too long, she may lose her skill at steering. This is just Mill's fear: someone who lets public opinion and the trends of society guide her life fails to exercise the part of her that truly engages with the world, anticipating events and making decisions to guide herself through them. This is the part of her that makes her an individual,⁴⁷ and without exercise it can wither.⁴⁸

John Smith, then, faces a decision. He could simply drift, which is to say he could let affairs unfold as they will without actually making a decision. The outcome of such avoidance is that he stays in Jamestown; but if this is what happens he has not *made* this decision because he has not taken an active part in engaging with the forces at work around him. Like the drifter, he will lack direction. He will not have done any anticipating of consequences or imagining of lives, and above all, he does no active valuing. Mill would lecture him because he does not exercise his faculties of discernment and judgment. So there is a sense in which Smith *should* make the decision for himself, for if he does not, he is

Of course, when the water is calm, it can be nice just to drift; sometimes we need a break from being in control, and we spend afternoons or whole vacations following our whims and letting the world take us where it will. But this cannot remain a permanent state of affairs; after too much vacation we start to get restless or crabby and long to go back to a directed routine. This brings to mind the experiment that Mrs. March lets her girls conduct for one week in Louisa May Alcott's *Little Women*. The girls complain about their daily chores, and in order to teach them a lesson about the importance of work, their mother forbids any work for a week. The only activities allowed are leisure pursuits. By the end of the week, the house is untidy, the floors are dirty, there are no dishes to use, no food to eat, and the whole family is grouchy. The girls learn the intended lesson, that too much leisure is not good for the soul, and return to their work with renewed vigor. The point, then, is that while it is sometimes good to take a break from exercising control, before too long the need for control reasserts itself.

⁴⁷ I will be discussing what makes a person an individual in more depth in the next chapter.

⁴⁸ Lewis Carroll's scene from *Alice in Wonderland* reminds us: "Alice came to a fork on the road. 'Which road do I take?' she asked. 'Where do you want to go?' asked the Cheshire Cat. 'I don't know,' Alice answered. 'Then,' said the Cat, 'it doesn't matter.'"

not properly engaged with his life. This is not a moral “should,” and as I said, this dissertation is not the place to examine in detail just what sort of “should” it is. Nevertheless, I think it is reasonably clear that there is such a “should” to be invoked here.

Thus, someone who does act as an individual and does engage with his life will find a decision like Smith’s profoundly troubling. Part of what it is to be engaged in the way I have been describing—part of what it is to steer the boat of one’s life—is to be an active valuer, and when two deep values come into profound conflict an agent will feel agonizingly torn. This rules out a coin flip, for as I have discussed, in this kind of situation a coin flip is a mark of indifference unacceptable to an engaged agent. What Smith and others like him need is a way of making their choices of one alternative *over* the other by engaging with reasons.

My problem for the anti-DBR view (including optional reasons variants), then, is that by projecting agents as detached assessors of their own motives, it gives no account of a reasoned way for people like John Smith to make decisions in favor of one alternative over another. Certainly they will not make light of them. They will in fact most likely go for the alternative that they find themselves wanting most or caring most about. Remember, however, that according to the anti-DBR view, this is no reason for the decision over and above the reasons *to* want that alternative; while this wanting may *in fact* shape the agent’s motivation, it is not a normative reason to do so. In these cases, the reasons *to* want the chosen alternative (independent of the actual wanting) did not seem to direct the agent to either option; so despite acting in accordance with what they want, according to the anti-DBR view the agents have actually broken their impasses without the aid of good reasons.

As I said above, this sounds implausible, for it looks as though there is a good reason available, namely that the choice is one that will shape the agent into the person she wants, i.e. can will (all-things-considered), to be. If she has deliberated well, the agent’s “leftover” or “extra” will—the fact that she is engaged with value and wants to adopt a particular ideal or to be a certain kind of person, over and above the reasons *to* do so—is a good reason in part because it has survived scrutiny.⁴⁹ That is to say, perhaps not all reasons *are* independent of desire, despite the anti-DBR view’s claim to the contrary.

⁴⁹ It is important to note that this claim is not the claim that informed desires give reasons; I agree with the anti-DBR view that any seeming normative force a desire has is transparent to the reasons *to* have the desire. But I stand by my way of putting the point because the background assumption is always that the agent is deliberating well, being honest about what the reasons are and the fact that her desire isn’t automatically an important consideration. A good deliberator puts her will to the test when she knows what it is; and if she

What is needed here is a way to understand how something that should be responsive to reasons can at the same time provide them. In the next chapter I will describe and defend the “self-creating” reasons, as I call them, that can play this role.

doesn't, then the deliberation can give rise to it, as the agent discovers and examines relevant considerations (or so I will argue in the next chapter).

In any case, something like this is the motivation for traditional ideal observer accounts. The assumption is that an agent has by hypothesis deliberated well, and what indicates the good or the reasons is what that idealized person wants (for you).

Chapter 4: Care, Identity and Ideals: Breaking the Impasse

We left off the last chapter with the need for a way to understand how something that is responsive to reasons can at the same time provide them. This will be my task over the course of the next two chapters. I see the problem has having two parts: (1) How can the deliberative impasse be broken? This is most prominently a moral-psychological question concerning how someone like Smith is to proceed in deliberation, so it leaves us with a further question: (2) Once the impasse is broken, how can we understand the normativity of the impasse-breaker? This chapter is meant to answer question (1); the next will address question (2).

At this point, let us recall, the problem is as follows. Smith has reached a deliberative impasse because of a conflict between things he cares so deeply about that they form part of his identity: remain in the New World with the woman he loves, or fulfill a dear ambition by accepting the commission to captain an exploratory expedition. Established practical reasons—all of the desire-independent reasons available to him, together with all of the reasons that arise from his current cares¹—leave him with no way of resolving the conflict, so there is a need for new reasons.

Having established in the previous chapter that there are good reasons to avoid an arbitrary decision and to avoid ducking the decision, let me finally come to the problem of showing how a decision such as John Smith's can be resolved in a way that is both non-arbitrary and does justice to the gravity of the situation. He must actively and deliberately change his life's trajectory or orientation, but in a way that is continuous with the current

¹ That is, all of the reasons that arise from his current cares, *however we are to understand such reasons*. For someone who holds the anti-DBR view, reasons that arise from one's own cares must for now be understood as providing the agent with desire-independent, agent-relative reasons for action. Indeed, the literature recognizes that what the agent cares about plays a role in what she has reason to do. For instance, Samuel Scheffler recognizes it when he explains that the correctness of the integrity objection to consequentialism arises from the fact that impersonal theories do not take proper account of the separateness of persons, nor of the natural way in which the consideration of ground projects arises from the agent's point of view. What I have to say in the next chapter will fill in the account of how cares give reasons.

one. He must run a deliberation that is practical in the sense that its input is practical reasons, both of objective value (and duty) and what we might call “personal” value (importance-to-me).

But this deliberation’s outcome is not practical in quite the same way a typical deliberation is. Typically, deliberation results in a judgment concerning what there is most reason to do or to care about, and the decision is yet to be executed. The conclusion of Smith’s deliberation, however, will be not “what *to care* most about” but “what *he wants to care* most about”²—and this is not strictly a discovery; it is, rather, a volitional result of the deliberation. If he is to make his decision actively and for himself, he must come to *make* one of the options more important to him than the other. This is not, of course, directly a decision to act. But it is a conclusion which gives him a reason to act—a reason that breaks the impasse I have been focusing on and makes his decision the right one in the circumstances.³

The task of this chapter is to fill in this brief sketch and show how such a deliberation could work. I will spend considerable time gathering resources in order to give a more detailed analysis of the kind of dilemma John Smith faces. With this analysis in hand, I will be able to give an account of how someone like Smith can break his deliberative impasse. I will then draw on the work of moral psychologists to show how the life-reorientation that this deliberation points to can be accomplished. In the next chapter, I will explain how this way of breaking the impasse is normative, thus resolving the problem I raised in the last chapter.

4.1 *Care and identity*

In order to accomplish the goal of this chapter, it is necessary to understand in detail what makes a problem like Smith’s so difficult. The short answer, as we have seen, is that care is what brings Smith to his impasse, so I begin by providing a detailed discussion of care and its ties to identity.

² Or, as we shall soon see, not “who to be,” but “who he wants to be.”

³ I will have more to say about this correctness and the possibility of being mistaken later, in Chapter 5.

4.1.1 *Preliminaries*

It is readily granted, I think, that value confronts us at every turn. There is more in the world that is worth caring about—there is more value—than any one agent can actively appreciate in one lifetime. But rather than be constantly overwhelmed by it, our interests and talents guide us to engage with some values more actively than others: some love art or music; others love sports; others garden; others are active in political causes; others care for the sick; others conduct scientific research; etc. Most of us engage with quite a few different values at any given moment or over the course of our lifetimes, making them more important to us than other things we acknowledge as *valuable* even though we don't engage with them. That is, we endow some objectively valuable things, but not others, with further personal value. Someone who loves the outdoors and goes camping in the back woods whenever she can might tend to give her money to environmental causes, whereas someone who works for social services may prefer charities that help feed the hungry and house the homeless. It is not that the hiker cannot see the importance of helping out others in need; rather, she has concerns that legitimately take her in other directions. And we take this to be a perfectly reasonable thing to do.

In fact, it is not only reasonable, it is necessary in order to navigate effectively through a sea of value; for without something to anchor our lives, we are adrift, and in a sense there is no one there.⁴ To get a grip on ourselves, we need to direct our activity. To direct activity, we need a direction; we need projects. That is, we need something of importance *to us* so that it stands out from all of the things that are worth pursuing. The need for projects, we might say, stems from Kantian-style considerations of autonomy, worth and dignity. Among these projects, moral considerations are often taken to be of the highest priority. Once we have fulfilled our moral duties, however, there is still a great deal of room for an agent to choose any number of different paths (say, according to different values). Here considerations of value—including particular personal value—are of importance, in addition to considerations of worth and dignity. It is important to note that especially if we acknowledge a plurality of different kinds of value and ways of valuing, the impersonal value of any given project, pursuit, or object is not enough to determine what its

⁴ This was a theme of Isaiah Berlin's; it also shows up in Frankfurt. See also discussion of individuality in Chapter 3.

role should be in a particular person's life. This highlights the legitimate role of desire, disposition, talent, choice, etc. in ordinary life.

We make things stand out by caring about them. As we will see, caring simultaneously endows certain pursuits with what I call "personal value" (which we might gloss as "importance-to-me") and gives the carer an identity or an individuality *as someone*; I am no longer merely a bundle of impulses, desires, appetites and appreciations. I have direction; I have projects in the sense that, in caring about someone or something, I adopt roles for myself and am motivated to fulfill these roles. I am a wife, a daughter, a sister, a teacher, a pianist, etc. I strive to live my life as an environmentally responsible citizen, or an entrepreneur, or to become a famous scientist, or to become a good mother and a good teacher. I orient my life around the cares I have and the ideals I strive to live up to. These cares and ideals, along with the activities I engage in as a result of them, make me the person I am; they constitute my identity in the sense of who I live my life *as*⁵—who the person is that I strive to be.⁶

Although these values are agent-relative (rather than agent-neutral) because they are personal, they can still be understood as desire-independent and would be recognized by the anti-DBR view as giving reasons. Special relationships are one such source of agent-relative but desire-independent reasons. So much can be seen from the fact that being a father gives one reasons to act in certain ways toward one's children that one does not have when it comes to other people, whether one likes it or not. Another kind of agent-relative reason is recognized by the anti-DBR view insofar as it yields indirect reasons: one's greater preference for, interest in, attraction to or care about some pursuits rather than others provides indirect reason to pursue them simply because they will lead to greater amounts of desire-independent, agent-neutral values such as intellectual satisfaction or pleasure. I do not ultimately think this is the only source of normativity in play here; as we will see, care also

⁵ See Kwame Anthony Appiah, *The Ethics of Identity*, chapter 1 (p. 16) and chapter 3 (p. 65ff).

⁶ This notion of personal identity is a notion meant to answer only one of a cluster of questions that are often grouped under the heading 'personal identity,' specifically, the question of "Who am I?" Other questions that often fall under this heading include "What am I?" and "How do I come to persist over time?" It is not entirely clear just how these questions, and the notions of identity that are meant to answer them, relate to one another. (Of course, the "who" and "what" questions presuppose that there is an answer to the persistence question, but there are plenty of details to fill in concerning the nature of the relationship between answers to one of the former and answers to the latter.) Filling in such details is beyond my scope here, but I find a theory like Bratman's, described below, to be an attractive one, partly because it strikes me as blurring to some extent the distinctions between these questions and their answers.

merits respect. I will argue for this claim later. The point for now is simply that caring can endow certain pursuits with the personal value needed for an agent to navigate in the sea of values available to him.

4.1.2 *Details 1: Care*

So much for preliminaries; now for the details. First, just what do I mean by ‘care’ in this context? Care, in this context, is the distinctive attitude defining personal value.⁷ It is an attitude of concern and appreciation that involves both emotional investment and volition, so it takes time to generate and to undo. But because it is responsive to reasons to some extent, with patience we can deliberately both cultivate and enervate care,⁸ and we hope that we manage to care about things worth caring about and not to care about those things that do not merit it.⁹

Care is, at its core, an emotion. It is a kind of constant, underlying emotion that gives rise to particular emotional episodes, desires, and actions. For present purposes, we can distinguish a standing emotion from particular emotional episodes.¹⁰ An emotional episode is a pattern of mental and behavioral events—thoughts, feelings, bodily states, involuntary facial expressions, predispositions to act, perceptual selectivity, direction of attention and imagination, etc.—that constitute what it is to be hopeful, angry, surprised, afraid, and so on at a particular time. An emotion, then, is the underlying and ongoing state (of investment or commitment) composed of these episodes together with the dispositions to experience them, according to the context at hand, because of an underlying vigilance for relevant details. In the case of care, emotional episodes can include excitement in anticipation of a particular event (a research trip, a child coming home from school, a

⁷ The account I give here is most directly drawn from Jaworska (2007) and Helm (2001), but see also Elizabeth Anderson’s (1993).

⁸ I will be saying more about this below.

⁹ Here, I actually mean something stronger than that we aim to avoid caring about things that do not merit it. I think there is actually no particular difficulty in caring about things that do not call for the attitude; my worn old pink rabbit Pete that I’ve had since I was a baby doesn’t particularly merit the care I have for it, but there’s nothing especially alarming, and there is even something endearing, about caring for such an object. Rather, when I say we aim to avoid caring about things that don’t merit care, I here mean that we hope we manage not to care about things that not only don’t merit it, but in fact merit the opposite of care (perhaps contempt, or some form of negatively-tinged indifference).

¹⁰ I follow Jaworska’s discussion (2007, p. 551ff) here.

concert), thoughts turned toward the object, fear for its welfare, the desire to do something nice (prepare a meal for a spouse, dust and protect a precious vase), etc.

Although care is at base an emotion, it does have a rational structure.¹¹ Because of the emotional investment a carer has in the object of care, to some extent the carer himself fares well or poorly according to how the object fares; this is part of what it is to care. And to the extent that these vicissitudes are important to him,¹² this imbues the object with import for the carer. We react with rejoicing when the object of care fares well, frustration when it does not; anger at others who disregard or abuse the object; pride in its successes and disappointment in its failures; desire to help ensure successes and prevent failures; fear when it is endangered and relief when it is spared; grief at its loss, and subsequent nostalgia.¹³ The constitutive emotional episodes, predispositions, and desires imbue the object of care with an importance to the subject (as a result of the subject's emotional vulnerability) that makes it stand apart from other objects in the world. These emotional states all refer to one object that connects them conceptually, and they bring with them rational pressures to have other attitudes in connection with them.¹⁴

In particular, Bennett Helm describes both "tonal" and "transitional" commitments carried by certain emotions.¹⁵ Tonal commitments connect negative and positive emotions; if you are prepared to feel joy at success, you are rationally required to be disappointed at failure (all else held equal). As Helm puts it, "if one experiences a positive emotion in response to something good that has happened..., then, other things being equal, one rationally ought to have experienced the corresponding negative emotion if instead what happened... were something bad." We can test this by receiving consecutive reports about an accident or a medical test. Suppose, for instance, you heard that your best friend was killed in an accident and you felt grief over the loss. Suppose you then came to learn that there was a misidentification and in fact your friend was spared, but you felt no joy over the news; there would be something strange about this (rationally speaking). Or, for example,

¹¹ Here I draw heavily on Jaworska and Helm.

¹² As, for most of us, they are, even though they need not be. (People aren't generally indifferent to how they are feeling.)

¹³ Jaworska (2007, p. 560).

¹⁴ Jaworska (2007, p. 561). See also Elizabeth Anderson (1993), who offers much the same account of care for persons, animals and things, as opposed to states of affairs.

¹⁵ Helm (2001, p. 67-9).

something would be wrong (rationally speaking) if I took pride in my child's accomplishments but would not be disappointed if on some occasion she were to fail. If this were the case (and it could be known), you might question whether I really did take pride in her achievements.¹⁶ Precisely what the "corresponding" negative emotion would be will vary with the details of the situation; often disappointment at not getting what you hoped to get is warranted, but if you don't get it because someone maliciously prevented your getting it, then anger is more appropriate.

Transitional commitments, by contrast, connect forward- and backward-looking emotions. For example, if you fear for a loved one's life, you are rationally required to feel relief to have him back home safely. Or if you hope to receive a grant, you rationally ought to feel satisfaction or disappointment when the recipients are announced. If you do not, one might suspect that although you *seemed* to hope for it, you were in fact indifferent to the possibility of receiving the grant in the first place. Here too the details of the situation make a difference as to exactly which emotion is warranted as a reaction to the anticipated event.

Although care has a rational structure, it is important to note that it is not merely a response to value or even perceived value. Attachments can grow independently of any value; it is possible to care about admittedly worthless things. Additionally, there can be appreciative responses that do not involve the investment or commitment that care engenders. Respect for fellow persons is one such attitude; admiration (which can be directed at both people and things) is another. Furthermore, we can appreciate value without thereby being moved to any action. A sports fan can appreciate the reasons for building a new concert hall, recognizing the worth of such a project, without being moved to contribute to the symphony's fundraising campaign. Care, however, brings with it a tendency toward appropriate action; a music lover will usually be moved to contribute substantially to the fundraiser.

Additionally, care does not fundamentally involve evaluative judgment. First, people who are incapable of such judgment may yet care about things. For example, a young child may try to comfort his upset mother by covering her with Band-aids, or an elderly person whose increasing dementia requires a live-in caretaker may yet care about the cooking she

¹⁶ Note that I might have good reasons to keep my disappointment hidden (e.g., so as not to make her own disappointment worse), but if I did not *feel* the disappointment, something would seem fishy, rationally speaking.

used to take such pride in.¹⁷ Second, we can find ourselves caring about things despite our judgment that they not only do not merit but even militate against such an importance-imbuing attitude. This is the situation for an abused woman who, despite her recognition of good, overriding reasons to leave her husband and subsequent resolve to do so, still loves him.¹⁸

4.1.3 *Details 2: Care and identity*

So far I have been discussing care on its own, but what of the relationship between care and identity that I mentioned above? Michael Bratman's work¹⁹ suggests that care's rational structure is part of a larger weblike structure of interconnected attitudes and episodes that help to constitute a person. He thinks that what sustains an identity is attitudes' continuity over time, together with connections constituted by inter-referential relations among attitudes. (On Bratman's picture, two mental states are linked referentially when they are what they are in virtue of containing references to one another—e.g., a policy refers to a type of action, and an instance of that action is executed as a result of the policy.) Attitudes that represent the agent require the support of such continuities and connections, and these must be the point of the attitude, not just a byproduct.²⁰ That is, I might (say) take a drug in order to rid myself of a desire for the drug, and ridding myself of distracting desires is a standing policy of mine, so that there is a referential connection between a policy and a particular desire. But the point of the policy is not to support my drug habit, nor is the point of my taking the drug in this particular case to support my policy of ridding myself of distracting desires. These are not concerns aimed at organizing my life. As such, they do not contribute to my identity over time. By contrast, an action, such as writing a paper, taken as a result of a plan such as getting published (which may itself be part of a plan to get tenure), does contribute to identity because the connections among policies and actions are intended to support one another, and my agency across time. I think this picture is quite attractive.

¹⁷ Examples borrowed from Jaworska and presented in more detail below.

¹⁸ We might consider Orleana Price, in Barbara Kingsolver's *The Poisonwood Bible*, to be such a person.

¹⁹ Bratman (2000).

²⁰ Jaworska (2007, p.561).

Bratman's work relies on self-reflection to knit these connections. But rather than take Bratman's route through reflection, a consequence of the Frankfurtian framework in the background which seems open to objection on this point, I follow Agnieszka Jaworska in her argument that these referential connections need not be conscious and reflective. The connections are actually more permissive: an emotion is what it is partly in virtue of its reference to a certain object over time, and this need have nothing to do with the subject's realizing that she has such an attitude.

Drawing on the work of Peter Goldie to describe the complex structure of emotions, Jaworska argues (as I mentioned above) that an emotion is a relatively enduring mental state composed of "various interrelated emotional episodes, and dispositions to experience subsequent emotional episodes, unfolding intermittently over time, at least partly in response to the context at hand." These emotional episodes and their objects are linked by conceptual connections, "a kind of conceptual convergence, linking disparate elements of a person's psychology occurring at different points in the history of her mental life."²¹ Many of the connections are referential in the way Bratman's identity-supporting attitudes need to be; they sometimes refer to one another, but more often they simply refer consistently to a particular figure or circumstances (as in grief, when one recalls the lost figure, or imagines what he might say in this situation, or wishes that one could share one's current experience with him). They can thus support the kind of intricate web of referential connections that Bratman points to as constitutive of personality and personhood.

Moreover, there is evidence from psychology that such emotional connections are necessary for a person to be the person she is.²² Patients with a particular kind of brain lesion retain full reasoning capacities and even "primary" emotions, which involve no higher cognitive processing (e.g. fear, disgust, anger and surprise). They lack, however, the "secondary" emotions that do involve (perhaps inchoate) understanding (e.g. gratitude, jealousy, guilt, hope, and grief). And as it turns out, such patients are unable to coordinate their activity over time. Since they lack no memory or reasoning abilities, it seems that these underlying "secondary" emotions are partially responsible for an agent's normal ability to

²¹ Jaworska (2007, p. 561-2).

²² Jaworska (2007, p. 556-7), citing the work of Antonio Damasio.

function. This is because of the psychological connections and continuities that these emotions support.

From here Jaworska goes on to argue that care is just the kind of emotion capable of supporting a web of interconnections that helps constitute personality. It is not (or not always) reflexive, and it does not presuppose evaluative judgment, neither of which can be assumed to occur in marginal cases. But it does have a rational structure, as I have already elaborated.

Thus, care helps to forge a vast web of rational and referential connections that can support a person's identity over time. On top of this, as Jaworska points out, because caring imbues an object with importance, someone who is capable of grasping that importance may be inspired to further cognitive activity: it can support stable intentions, plans, and policies that guide the activities of an agent, constituting her identity and orienting her life.

What has been said here bears some affinity with Christine Korsgaard's account of the sources of normativity, but I would like to emphasize a difference between her account and the one presented here: namely, that the source of a person's identity—and her personhood—lies in the person's cares, not in her rationally reflective nature. As we will see in Chapter 5, rationally reflective nature does play a part in the self's creating reasons, but it is not doing all of the work. Korsgaard identifies the foundation of her doctrine of practical-identity-as-the-source-of-all-reasons as the need to have an identity in order to be able to live at all:

What is not contingent is that you must be governed by *some* conception of your practical identity. For unless you are committed to some conception of your practical identity, you will lose your grip on yourself as having any reason to do one thing rather than another—and with it, your grip on yourself as having any reason to live and act at all. But *this* reason for conforming to your particular practical identities is not a reason that *springs from* one of those particular practical identities. It is a reason that springs from your humanity itself, from your identity simply as a *human being*, a reflective animal who needs reasons to act and live. And so it is a reason you have only if you treat your humanity as a practical, normative, form of identity, that is, if you value yourself as a human being.²³

This is consistent with the research Jaworska cites on patients who have no “secondary” emotions and as a result cannot live coherent lives, as well as with things I have said and will go on to say. But I might object to Korsgaard's way of conceiving the problem

²³ Korsgaard (1996, p. 120-1).

because, if Jaworska is right (as I think she is), you don't need a *conception* of your identity to be acting out of the cares that make it up. You need an identity—and in this much I agree with Korsgaard²⁴—but you don't need to know what it is or value it consciously in any way. You have it in virtue of caring about anything. I will soon claim that you have to value *something* in order to be a person at all; but you don't have to value yourself directly. Rather, you value yourself indirectly as a result of valuing other things.

Perhaps this is too quick a dismissal of Korsgaard's position; perhaps it is an uncharitable reading of her argument to think that she claims that you must know and value your identity in a conscious way, under some particular description. As I argue later on, children and other marginal agents are able to take a stand and defend themselves because they care about things; but they would be unable to articulate that thought because their rationally reflective nature is still developing or is diminishing, so they do not have what I would call a "conception" of themselves, except implicitly. If this is all Korsgaard means when she claims that a person must (implicitly) value herself in valuing anything at all, then I have no quarrel with her. But it is not always clear that this is what she means. In her response to her critics, for example, she discusses willing as "self-conscious causality," where "to will is not just to be a cause, or even to allow an impulse in me to operate as a cause, but, so to speak, to consciously pick up the reins, and make *myself* the cause of what I do."²⁵ I am not sure that this is what a child is doing when she throws a temper tantrum over wanting to tie her own shoes. Nevertheless, *if* the self-consciousness invoked here is weak enough that children can have it, then I do not disagree. I am not fully confident, however, that Korsgaard uses the term so weakly.

At this point I want to highlight two related aspects of care that make it an important element of a person's identity. First, because care is an underlying emotional state that connects particular emotional episodes and desires, thereby giving rise to actions, it lends an

²⁴ She writes: "[T]o value yourself just as a human being is to have moral identity" (121). I agree with this if it is meant as the claim that any action or attitude that implicitly entails valuing oneself is enough to give the person a certain kind of basic moral standing with others. I believe that Korsgaard means something stronger, however, because she makes claims to the effect that having "moral identity" means taking up the moral point of view that considers and values others, and that this identity is necessary. I agree that the moral identity we have that entails a standing with others and an authority over our own lives is inescapable, and I might with some argument concede that the moral point of view is necessary in some sense (as in, one is morally wrong if one does not take it up), but I would not go so far as to claim that taking the moral point of view is inescapable.

²⁵ Korsgaard (1996, p. 227).

agent a sense of continuity and purpose. We can construct narratives of our lives—not just annals or descriptions—because the objects we care about give us the kind of direction needed to tell a story. Since stories help us to understand events, caring helps us make sense of ourselves.

In fact, David Velleman argues in “Narrative Explanation”²⁶ that the element of a story that uniquely contributes to explanation—going beyond causal connection—is what he terms “emotional cadence.” We understand stories not just, or not wholly, because of the familiar sequences of events they portray, but because of the emotional patterns that we recognize because we have lived them ourselves. Velleman synthesizes a number of others’ ideas on narrative. He draws on Schank’s idea that we understand stories by assimilating them to the familiar, but departs from Schank’s work in arguing that the familiar is really an emotional familiarity, not cognitive familiarity with how things happen. He draws, too, on Mink’s characterization of narrative understanding as literal comprehension (“grasping together”), the idea that narrative form allows the hearer to grasp all at once the complex relationships among events, to understand the events as a completed whole. Velleman departs from this as well, however, because he wants to preserve the link between narrative and explanation which was his initial observation, and Mink’s “grasping together” need not provide explanation, only coherence. Velleman thus emphasizes that emotional aspects provide the further element of explanation by making a narrative into a complete whole, rather than a merely coherent bundle of events.

At the end of the essay, Velleman warns against mistaking the subjective understanding that grows from emotional cadence for objective understanding—i.e. he cautions that arriving at a settled attitude toward events does not entail a true understanding of how they came about (this is a warning mostly about narratives of history, but it has wider application). Real life doesn’t come in convenient narrative packages; rather, we make sense of it by framing narratives that initiate an emotional sequence of tension and resolution. It is a continuous flow of events, but we impose structure by chopping them up into chunks we can handle. Having objects and projects around which to organize these chunks is an essential part of the process.

²⁶ Velleman (2003).

Of course, one agent may have any number of parallel and intertwining stories unfolding at once. But this is the case in a good novel as well: several intertwining threads make the whole. As readers, or as persons living a life, we can focus at times on a single thread, at times on a few together, or at times step back and appreciate how they all fit together to make a whole. Thus, care contributes to identity by providing a basis for understanding ourselves in terms of narratives we construct around the things we care about.²⁷

Above I said that there were two related aspects of care that constitute its contribution to one's identity. The first was the ability to construct narratives. Second and more fundamentally, we need to care about things in order to have identities in the first place. There are two different points to make here. First, in caring, the welfare of the cared-for object becomes part of the carer's welfare, and so the carer identifies with the object in the sense of (to an extent) sharing a fate with it.²⁸ This ties care to a concern for how one's life goes and, hence, to one's identity; concern for the object is indirectly concern for oneself.²⁹

Second, recall that caring about particular things puts needed limits on an agent's deliberative space. We cannot deliberate without limits on the scope of deliberation (as embodied, perhaps, in ideals, about which I will say more later on), for otherwise it is impossible to direct our actions toward any global goals.³⁰ Care narrows the field of eligible

²⁷ See also Alasdair MacIntyre's *After Virtue* (1984), pp. 200-3.

²⁸ (This sense of identifying *with* is different from the sense I mean when I discuss identifying oneself *as* a certain sort of person or leading a certain sort of life.)

²⁹ This may raise the question whether the "self-creating" reasons I will be describing cannot be accommodated in terms of what would make one the happiest or give one the best life. My answer has two parts. First, we must be careful to understand what we mean by being "happy" or having "the best life." Some life-shaping decisions are made *against* one's happiness, as when someone makes a great sacrifice of happiness in order to uphold an overriding value. However, such a sacrifice may be interpreted as a choice of the "best life" (though those with an Aristotelian bent might disagree, if such a life lacks pleasure), which brings me to my next point.

Second, the question of what the best life *is* is a substantive question not only for all of us collectively (i.e., in a generic sense), but for each of us. In the framework I have set for this discussion, I have assumed that certain objective standards have been met, so that anyone who falls into the dilemmas I discuss is well on his way to living a "good life" in the sense meant when we ask "What is the good life?" in a generic way. But even with these basic standards met, the question of who to be is not resolved, because it is different for each person as a result of individual dispositions, preferences, and propensities. Consider, for instance, some historical heroes: Martin Luther, Martin Luther King, Jr., Charles Lindbergh, Sally Ride. These are people who led lives good enough to meet the basic standards of the "good life," and yet their lives were very different. The choices I have in mind are choices not of how to lead a good life, but of which good life to lead.

³⁰ Although my claim here is weaker than Frankfurt's, I draw here on Harry Frankfurt, in parts of "The Necessity of Ideals," "On Caring," and "On the Usefulness of Final Ends," in Frankfurt (1999); and "The Dear

actions by excluding some actions from deliberative consideration and giving an agent something to defend and to benefit. If absolutely every action were up for consideration at every turn, then even the direction of one's will would be available for choice, leaving one no firm basis on which to make any decisions. Thus care functions to set bounds on the realm of deliberative possibilities. These boundaries in turn define the person by defining what actions are within the realm of consideration. Without such identity-conferring boundaries, there is nothing that defines him as a self; he simply reacts to the whim of the moment as aroused by immediate circumstance.³¹

Caring, then, is itself important to us. And as a result, we have good (though not necessarily overriding) reason do what care requires—usually, this is to act for the sake of the things we care about. But because of the tie between welfare and the objects of care, and because of the need to care about something, precisely in acting for the sake of something else, one acts in a self-fulfilling way. But for most of us, it is not enough simply to care about *something*; we have an interest in caring about things that we think really matter, because these are the things that give our lives true meaning by linking them to things beyond ourselves. Note that success is therefore not the only measure of this meaning; a person can spend her life dedicated to work that remains unfinished at the end of it, but looking back she may well be satisfied with her contribution because of its engagement with an important value.³² She may be proud of herself because of the identity she has carved by caring about something valuable independent of her care.

Self,” in Frankfurt (2004). I should note that what I refer to here as ‘care’ is distinct both from what Frankfurt calls ‘love’ and from his ‘care’. For instance, he reserves ‘love’ for particular things we cannot help caring about, making it a specific kind of caring. On the other hand, however, he is unwilling to allow that disappointment over, say, giving up going to a concert to help a friend is an indicator of care, however minor; but I prefer to think that the disappointment is a sign that the music lover *did* care about the concert, just less than he cared about helping the friend. (See esp. Frankfurt’s discussion in “On Caring,” section 6ff.)

Also, see Williams (1982) and Taylor (1985a, 1985b) for influential discussions of the importance of personal projects.

³¹ This should sound familiar from the earlier discussion of individuality.

³² Frankfurt agrees; see (1999, p. 86). Darwall argues convincingly for a similar point in Chapter IV of *Welfare and Rational Care* (2002). Blustein (1991, Ch. 3, p. 45-6) discusses whether something could have personal value without having impersonal value, and seems to agree in a limited way with Loren Lomasky, who asserts that at least some of the time, the answer is no. The “some of the time” seems to be the important qualification here. For the deepest kinds of commitments that guide major life decisions, impersonal value seems to need to underlie the personal value, or else the agent will be unable to find meaning in her life. But Blustein thinks that there are things which can have personal value for us but which we don’t care about in the deepest sense, and these need not imply impersonal value. As Blustein puts it, “A life in which everything we care about is believed to have impersonal value would be a life weighted down by a pervasive sense of seriousness and of the momentousness of our decisions and actions. At the same time,” he goes on, “personal value that is not

4.1.4 *Care and volition*

In the last section we have seen that because it sets boundaries on deliberation, care settles certain typical questions under typical circumstances for an agent. So far I have concentrated on care's emotional and rational aspects, but this last observation brings out a third dimension that is related to both its emotionality and its rationality: care is also volitional.

In general, having a volition is largely a matter of settling a question as to what to do, typically (if not always)³³ on the basis of reasons. Additionally, questions as to what to do are often raised by desires and other motivations, such as those that flow from emotions. So volition has ties to both emotion and reason.

But here we must distinguish between two scopes of the will, which can be thought of as its "agency" and "identity" aspects. The agency aspect is what action theorists such as Frankfurt, Watson, Bratman and Velleman are after. The classic problem of agency is to distinguish "where the agent is" in a sea of immediate or short-term psychic happenings—i.e. to find a distinct attitude (or set of attitudes) that has the authority to "speak for" the agent and guide action. Agency in this sense distinguishes true actions from mere acts by making an act fully the agent's own.

The agency aspect of the will operates locally, directing intention and particular actions (often by committing to the object of one desire among several). Desires and emotions are immediate motivators that pop up in response to particular circumstances;³⁴ they are the psychic states that bring attention to the objects among which the will typically adjudicates to form intentions, thus constituting an agent's agency. Notice, however, that care often settles certain local questions because it defines what has value *to the agent*; hence, local volition serves as an interface between care and desire: local willing is the exercise of agency to adjudicate among the objects of desires, according to the boundaries set in caring.

augmented by the conviction of impersonal value cannot make a person's life *meaningful*' (47). We need the impersonal value for the most important things, the things that make us who we are and around which we organize our lives, but we also need some things that don't necessarily have impersonal value in order to "lighten the load," so to speak. This strikes me as a reasonable position.

³³ For example, Christine Korsgaard argues in her *Locke Lectures* (2002) that the will is identical with practical reason.

³⁴ Frequently as responses to perceived reasons, or as the manifestation of cares.

This suggests that the second scope of the will, what we might call “global volition,” is essentially care. As we have seen, care operates globally, circumscribing the possibilities for action by setting bounds on the local will, and constituting an agent’s identity in the sense of giving her life a structure that she can make sense of.³⁵ So this is “where the agent is” in an identity sense; it is something like a plan of life, to use Rawlsian terms. It is important to note that the boundary-setting here is of a volitional sort, determined by commitments to values and ideals; it is not a matter of circumstance or a failure to consider relevant possibilities. I speak of volition here, and not just desire or mere habit, though these things may conceivably give one’s life structure. This is because of the need to make sense of the structure; *mere* desires or habits may not stand up to reflective scrutiny. Now, certainly, where there is care there is desire and probably habit, but these are not then “mere” desires or habits. They are in line with the care and give a sensible, endorsable structure to an agent’s life. In general, then, volition represents commitment: globally, to a certain path of life or mode of living (typically, according to values one cares about); and locally, to a course of action (which is another way of saying that volition and intention typically go hand in hand).

With this analysis of care, identity and volition in hand, we can now get a better understanding of Smith’s problem.

4.2 *Analysis of the problem*

As I said at the beginning of this chapter, Smith must undertake a kind of deliberation that involves something which goes beyond the deliberation that brought him to his impasse. This original deliberation followed the discovery model described in Chapter 2: it was a search for reasons that are independent of anything Smith wants or is inclined toward. Any attitudes or desires he discovered himself to have over the course of that deliberation were reviewed as transparent to the reasons that support them; that is, the attitudes themselves

³⁵ This idea is related to Frankfurt’s. See note 30. I should further note, however, that this claim need not exclude the possibility that someone could come to recognize new reasons, for instance as put to her in interactions with others. If this were impossible, agents could never change. Since we can, it must be the case that cares do not exclude our ability to take other points of view, from which we may have access to reasons we were unable to see before. Such new reasons may even be the kinds of things that cause an agent to question what she cares about, or whether to care about it. They can therefore be rather disorienting. On the other hand, however, these reasons may take time to be recognized, because of the inertial barrier provided by current cares and concerns. The point here is that without prompting by something or someone, cares typically constrain what an agent will consider as open options for action.

counted for nothing beyond the reasons that made them rational. This kind of thought process, however, did not decide the matter. And Smith has the *further* problem that his cares do not settle anything either. He is at an impasse not simply because desire-independent reasons do not yield a verdict, but because his cares and the reasons they give him do not either.

Cares thus bring Smith to his impasse, and their tie to identity is what makes the problem so difficult (since it is the fact that his identity is at stake that excludes arbitrary picking as a means of resolving the problem). The trouble is that the cares that are normally taken as fixed boundaries for deliberation have themselves come under scrutiny and are therefore unfixed; the local will does not have its usual global boundaries and direction. The question before us, therefore, is what can take place when care does not settle local questions—when what an agent cares about is up for examination. Local volition in situations like this must work without guidance from care: one must exercise agency to operate on identity, the very thing that usually confines its task. Smith must choose a new identity and align his cares in a new way. The difficulty is locating a consideration he can take as settling the question before him, given that: desire-independent reasons do not settle it, the cares that would normally adjudicate among options are themselves up for question, and too much is at stake to settle it by arbitrary means (such as flipping a coin). I will return to the question of locating a question-settling consideration in the next section; right now I want to remain focused on the nature of Smith's problem.

I said above that Smith needs to engage in this deliberation differently than he did before—differently from the way the anti-DBR model would suggest one deliberate. Deliberations on this model are settled by arriving at a view of what there is most reason to care about. Buying a car, I might face a situation in which the model I am to choose is settled once I determine whether I should (or do) care more about gas mileage or safety. Sometimes it may be extremely difficult to determine what reasons dictate, as in the choice Marguerite Blakeney faces between her brother and the Scarlet Pimpernel,³⁶ but nevertheless, if such reasons can be uncovered, they will decide the question for her. Or take questions such as: What should I think about the rights of women or minorities? About what, if anything, should be done about global warming? These too depend on what there is most

³⁶ See Chapter 3 (section 3.2).

reason to care about.³⁷ Smith's deliberation, as I have argued, does not look like this, because his desire-independent reasons do not answer his question as to which life to care more about (and hence which to lead).

A deliberation may also at times be settled by a discovery of what the agent actually cares most about.³⁸ Desire-independent reasons may balance out, but the agent may, after some reflection, be (or become) clear about what is most important to her. Isabel Archer's choice to return to her husband is decided because it is clear what she cares most about: her integrity as a woman of principle. In cases where the agent understands himself less well, we have the stuff of last-minute drama at the movies: a hero realizes what matters most to him, and returns to deliver the happy ending. But again, this is not what Smith does. He realizes already that two things are deeply important to him. He must make one of them *more* important to him. So he must deliberate, but he will not decide his path by discovering what *to* care about, nor by discovering what he *does* care about.

Given everything I've said about care and the source of the problem, however, what could be different now?

My answer begins with the observation that it is genuinely up to Smith what he should do. It is genuinely up to him, not just in the "free will" sense of his being an agent who can be the author of his action, but in the sense that the question of *his* leading his life is before him in an urgent way. Desire-independent reasons permit more than one option, and it is up to him to embrace one of those options, if he is to direct his own life. And as I argued in the last chapter, he has some responsibility to do so.

Because it is genuinely up to him, and because his very identity is at stake, I claim that it is now open to him to consider not "what to do," as in the original deliberation, but "who he wants to be." Now, at first it may appear that there are no more resources for answering "who he wants to be" than there are for answering "what to do." In asking who

³⁷ Sometimes settling what to care about does not fully settle what to do. There may be most reason to care about equality, and thus to end slavery. But even if I am a white Virginian in 1858, who realizes that black people deserve equal rights with white citizens, this opens a number of actions open to me. I could join the legislature and try to abolish slavery through legal means; I could join the underground railroad; I could undertake a campaign to shape public opinion against slavery. Which of these is best supported by reasons may differ from situation to situation.

³⁸ Again, I have only given the briefest of sketches concerning how care gives reasons even on an anti-DBR view of reasons. The full details are forthcoming in Chapter 5.

one wants to be, after all, one essentially asks what reasons there are to care about things. And in the situation described here, these reasons do not decide the matter.

Nevertheless, I contend that “who one wants to be” is a question that signifies quite particularly what is at issue in Smith’s choice. He is at an impasse in which his identity is at stake, and not just his identity on a first-order level as a lover or an explorer, but on a higher-order level, as an agent who can choose and take a stand on who he is to be. After all, as I suggested in Chapter 3 and will argue more fully in Chapter 5, the reason he must make this decision is that it is incumbent upon him that he take up the authority he has to lead his life. Authority over one’s life is a moral notion: it is the basis of the right to autonomy, and it is also the basis of the responsibility one has to oneself to take up the point of view of being in charge, i.e. to live actively by making decisions and embracing courses of action. I will be elaborating on this in Chapter 5. For now, the important point is that this question of taking up authority over one’s own life, which is raised by the kind of situation that Smith finds himself in, engages a self-conception in a way that many deliberative questions don’t. The dilemma has become a question that is as much about him as it is about the alternatives before him.

Why should this be? When does a choice bear on the question of *who* is leading one’s life? I offer this answer: in cases like Smith’s when what he cares about is in question, i.e. when the boundaries of the will have come unfixed. If John Smith’s twin James, who knows all along that he cares more about adventure, were to perform an “ordinary” discovery-model deliberation and determine that there was no fact of the matter as to what he had most desire-independent reason to do, then the choice would be up to him in a genuine sense, but this sense isn’t as strong as it is for John. The sense in which the choice is up to James is the fairly common sense of desire-independent reasons’ being tied: he may do whatever he cares most about, since there is nothing that weighs against his carings’ deciding the matter. But it is up to John in the stronger sense that *his* choice will be wholly a matter of his exercising agency to define for himself a new path. He may do what he cares about, certainly; but he must actively determine what it is that he cares most about. As we shall see, this means that he must do some introspection that goes beyond self-discovery into a robust kind of self-interpretation that raises a new inclination which forms a basis for a sound choice.

This is where the case differs from questions of integrity and moral dilemmas. So before moving on to a discussion of how Smith will break his impasse, I would like to contrast his case with others that are in the same neighborhood, which we have seen once already in Chapter 3.

First of all, determining what one cares most about runs deeper than adjudicating between conflicting desires or preferences. We are accustomed to having to choose between incompatible alternatives, and we do this without thinking much of it; certainly we don't take most decisions to require adjusting our most deeply held values. I can find myself wanting both to attend my niece's basketball game and to attend a banquet in honor of a work colleague, knowing that under the circumstances I cannot do both. So I consider my priorities and decide what makes the most sense for this occasion—perhaps I know that there will be another basketball game that I can attend, so I decide to go to the banquet this time. Or it might be that this game is a championship one that means a great deal to my niece, so I send my regrets to my colleague and go to the game. None of this, though it may cause me some (even great) consternation, regret or disappointment, is the sort of thing that has deep importance for my life or my identity. A life-changing decision is more a matter of examining and perhaps adjusting the deeply held beliefs and cares that make one the person one is.

Second, determining what one cares about involves a permissible adjustment of values that may not be allowed in the presence of a moral dilemma, where the value at stake is not just personal value. How is Marguerite Blakeney to choose between her brother and the Scarlet Pimpernel? Because this is a moral decision involving not just what she values but her obligations as well, choosing to adjust her values in this case does not seem like a viable option. Can she really make sense of her choice by thinking “I choose to live for my family, even at the cost of betraying a hero and the lives he might save”, or, alternatively, “I place the lives of strangers over that of my brother”? Her difficulty is, rather, in understanding where her moral obligation lies. Of course, she *must* make a decision, and she eventually chooses in favor of her brother. And if it were not for later intervening events, she might well have learned to live with the choice; and she might have taken some comfort from the thought that she was not blameworthy for getting into the situation in the first place. But this, though a kind of healing or moving on, is not a reorientation of her life around different values. It is a reactive adjustment to circumstances rather than a proactive

shaping of them. The way for her to deal with the choice is not to explain it as a change in her values, but to say that she did what she had to under circumstances that were out of her control. Very likely, she will harbor a deep regret that she chose as she did—but this would be the case no matter which way she chose.

Third, the choice of cares differs from an integrity-based one. Here I want to distinguish between acting *with* integrity (in the sense of restoring wholeness), which I think Smith does do in making his choice responsibly, and acting *out of* integrity, which I think Smith does not do but Isabel Archer³⁹ does. Someone like Isabel makes a decision in the face of temptation, but in accordance with the values she already has. Isabel knows where her priorities are, and nothing in her circumstances is forcing her to change them, though the circumstances do cause her to reflect on them (as we see in her visit to her cousin Ralph in spite of her husband's forbidding it). Now, this kind of case is more similar to John Smith's than any of the others, in the limited sense that both are choosing who they are in a very conscious way, and both undertake an amount of self-interpretation that reveals quite a bit about themselves. The difference between them is that Isabel has the raw materials for an all-things-considered conclusion, which Smith does not have.

Isabel's temptation to leave her husband presents her with a choice because her misery is a reason to leave him. Despite this, however, she chooses to remain true to her principles and return to her husband. Two (not necessarily exclusive) things might be going on here. First, Isabel might take maintaining her integrity to be her reason. That is, she might take it that above all, she must preserve herself, in this case by acting on her principles. But this seems to be the wrong sort of reason for making a decision; if it is a matter of her obligations to Osmond, then her own character seems not to weigh much in the decision. There is a worry about moral self-indulgence here, that the agent who takes integrity to be a reason is blowing her own importance out of proportion. Principles are principles, and why should it matter that they are *her* principles? It seems odd that something that looks decidable on the basis of the features of the situation (including those of the agent involved) should suddenly be *about* the agent; the issue is what to do, not who the agent is.

But in Isabel's case, like Smith's, it is *her own* principles that are being questioned—as hers, not as principles. The suitor is asking her to reconsider whether it is worth adhering to

³⁹ See Chapter 3 (section 3.3).

principles when doing so will make her unhappy. He is not necessarily questioning the very principle of honoring marriage vows. And so it is a question of personal value: not “Should Isabel, like anyone, adhere to the principle?” but “Should Isabel adhere to the principle because it is her principle?” The question is whether she should hold it as her own principle. In Isabel’s case, like Smith’s, the boundaries against which she normally makes decisions have become unfixed by being questioned. And so it is a question about her, and maintaining integrity is not necessarily the wrong sort of reason to cite.

By framing it as a question of integrity, however, Isabel does not take the possibility of violating her principles seriously. The integrity frame preserves a sense that it is not entirely *up to* her, though it is *about* her in an important way. In framing the thought that it would be a violation of integrity to leave her husband, Isabel restricts her own freedom to choose; she does not face the same kind of self-creating choice that Smith does because she cannot see both of her options as equally worthy.

In putting the point this way, I agree with Jeffrey Blustein, who thinks of integrity as transparent to the objects of an agent’s values. He claims that a person of integrity is not guided by the thought that she is committed to *her* not doing *x*, where the reason is that it would undermine her integrity. Rather, she is committed to something more like “I am committed to *its* not being done through me,” where her reasons for not doing it are simply the reasons not to do it—instead, that is, of its being a sacrifice of *her* integrity.⁴⁰ He takes it that adherence to principles, or the maintenance of integrity, is to be viewed as an expression of the special (we might say, identity-affirming) importance the core values involved have for the person.⁴¹ This avoids some of the classic objections to personal integrity’s being a reason for action.⁴² So while reasons of integrity are agent-relative, involving as they do the thought that the values in question are the *agent’s* values, the force of the reason cited when integrity is invoked is located primarily in the reasons there are *to* value the object. In acting for the sake of her integrity, the agent is actually acting for the sake of the importance the *object* (in this case the principle of adhering to marriage vows) has for her.

⁴⁰ Blustein (1991, p. 86).

⁴¹ Blustein (1991, p. 87).

⁴² Chief among these is the difficulty of whether preserving one’s integrity can be an overriding reason when the integrity one is preserving is a dedication to an immoral project, such as Nazism.

Does this apply to Smith's action as well? Yes, in the qualified sense that his choice is ultimately an expression of the importance to him of accepting his commission. But no, in the sense that he has not already established what is most important to him. Isabel has her priorities set, and is tempted away from them by the suggestion that she has them wrong; but Smith does not have that prior priority to work from. His task is, in part, to *set* his priority; he faces a much more radical question than Isabel does because he cannot frame his question as one of integrity, thus restricting his choice. Smith's choice of who to be is a decision about what self to create; Isabel's is a choice made to preserve the self she has.⁴³

The second thing that might be going on when Isabel refuses to leave her husband on integrity grounds is related to the first. Isabel might find herself unable to think seriously about running off with her American suitor, despite tempting flights of fancy. It is one thing to imagine purely hypothetically how relieving it would be to separate from Osmond permanently, but it is quite another to start packing one's trunk. When it comes time for serious action, sober reflection shows Isabel that escape to the United States is not a serious option for her. Isabel realizes that she would not be herself if she ran away with her American suitor; she would unravel. Although she may enjoy a more pleasant life from day to day, she would have the constant haunting feeling that she had betrayed herself. She simply cannot bring herself to go with him because the ideals she has set for herself define her deliberative space in such a way that she cannot take the possibility seriously. Though tempted, she is not torn. Her commitments do yield a verdict. There is self-discovery here, but not self-creation.

Smith differs from Isabel in this respect as well. He is not going to unravel either way he decides; he therefore has more than one viable path before him. He will have work to do in living with his choice, for both of the commitments between which he is choosing are deep enough to be identity-conferring; but it is not insurmountable work. Even after deciding that he will leave Pocahontas, that choice will not have the ring of unity and freedom that often accompanies a choice that simply feels right;⁴⁴ it will more likely have the

⁴³ Had she left Osmond, she would presumably have made a mistake. This is because it is clear, from the way James tells the story, that Isabel's principles are of primary importance to her. If the occasion had brought them into more serious question, then she would have been facing a choice like Smith's and this case would provide little, if any, contrast to Smith's.

⁴⁴ "The ring of unity and freedom" does not exactly seem apt as a description of Isabel Archer's decision; her choice to return to Osmond is an act of resignation. I do not think this is a reason to believe that she is not acting under volitional necessity, however. Her decision will have the feel of being the right choice if it is

feel of a tragic choice. Once he has made the decision, Smith will very likely feel the relief of having resolved his impasse, but such relief is mixed with the deep regrets of giving up the woman he loves, in a way it would not be were his situation more like Isabel's. (This is not to say that choosing one's integrity excludes all regret; one may certainly regret the circumstances that required a decision. Still, I suspect it does exclude the kind of deep regret of one's own choice that borders on grief, because someone who acts out of integrity will not have to struggle with her contrary emotions, knowing she really could not do other than she did.)

Finally, Smith's case is both like and unlike the story of Nora Helmer in Henrik Ibsen's *A Doll's House*.⁴⁵ In Act One Nora seems to be an innocent, shallow, and flighty character, who must find little ways of rebelling against her husband Torvald (such as buying pretty things and eating macaroons), who clearly dominates her but whom she nevertheless pretends or appears to love. She prides herself on the one rather large secret she has from him: that, rather than her father paying for the trip to Italy that is supposed to have saved Torvald's life, she borrowed money from the sketchy character of Krogstad. In fact, however, we learn that it's worse than that: she forged her father's signature on the IOU, and Krogstad thus has leverage with which to blackmail her. It is a mark of Nora's innocence that she did not know and cannot believe that her forgery is a serious crime despite her good intentions in doing it. She wants to think that her love for her husband and her respect for her dying father could exculpate her from both of her rather shady deeds.

Krogstad's blackmail takes the form of forcing Nora to attempt to get her husband to preserve his job at the bank (of which Torvald has just become manager). Torvald intends to fire Krogstad because of his moral depravity and undue familiarity with Torvald left over from school days. When Nora brings up the subject, Torvald makes a speech which scares Nora: he says that the moral depravity Krogstad exhibited has a tendency to corrupt the whole home atmosphere in which the perpetrator lives. Nora notices that he

indeed true that she could not have done otherwise; sometimes there is something comforting about resigning oneself to the inevitable. So while Isabel's unity and freedom may not exactly "ring" for her, she will nevertheless feel relieved and unified as she travels back to Italy.

⁴⁵ Another I can think of is Orleana Price in Barbara Kingsolver's *The Poisonwood Bible*, who is the wife of Baptist zealot who moved his family to the Congo to be missionaries. The political situation becomes dangerous for white people, but her stubborn husband refuses to leave. Orleana becomes despondent and withdraws to her bed for months. When her youngest daughter is killed by a snakebite, however, she snaps out of her lethargy, leaving her husband on foot with her three other daughters in tow.

puts this first in terms of the *mother's* moral failings infecting the children. Suddenly Nora is afraid of what she may be doing to her own children in keeping her secret and trying to lie about it.

She intends to commit suicide after a Boxing Day masquerade ball. In the meantime, her childhood friend Christine Linde makes a move which soothes Krogstad, who, already having slipped the letter revealing everything into Torvald's mailbox, returns while Torvald is making a scene of denouncing Nora to slip the incriminating IOU into the box as well. His ringing the bell interrupts Torvald's anger at Nora. Upon discovering what has been delivered, Torvald completely reverses his attitude and forgives his wife, reverting to treating her as a "songbird" and "squirrel," or, as Nora then puts it, his doll.

But it is too late. Nora has finally realized the position she is in. She had been expecting a "miracle" when her husband read Krogstad's letter revealing her sin: that Torvald would rise to the occasion and protect her by taking the blame himself. When instead he accuses her and acts as though he will disown her in all but appearance, she realizes that he doesn't know her or love her—that he doesn't respect her. His sudden switch when the IOU is returned—and his first joyful reaction being for himself ("I am saved!")—reveal to her that her husband, like her father, treats her as a doll rather than as a person. Finally recognizing this (though she has been told before, particularly by Mrs. Linde), she cannot tolerate it, and so she leaves him.

Nora's choice is a self-constituting one, much as John Smith's is, and like Smith, her authority is in question; but unlike Smith she faces a situation in which what choice to make on the merits of desire-independent reasons is clear. Nora faces no inner struggle, no conflict of values, as John Smith does. But she does face a crisis of authority much like Smith's. Nora has extremely compelling, will-independent reasons to leave her husband: he does not respect her, and if she stays with him, she is not able to live *her own* life. Until the end of the play she neither clearly appreciates nor has the will to act upon these reasons, which have been there all along, perhaps even dimly recognized (as suggested by her rebellion through her macaroons, her frivolous spending and her secret). Nora's realization is that she has authority over her life and that she can—must—exercise it by taking responsibility for her own life for a change. But she is not a candidate for self-creation the way Smith is because she does not need to exercise her authority in the service of making an

inclination into a reason for her to act. She may face such choices as she builds her independent life for herself, but the decision to leave Torvald is not yet of this kind.

Let me remind us, then, of where we are. In this section I have been focusing on the details of the problem faced by someone in a position like Smith's. The discussions of the contrasts to moral dilemmas, cases of integrity, and cases of taking up the authorship of one's own life were in the service of illustrating quite particularly what is going on in Smith's case: that it is genuinely up to Smith what to do, in the very particular sense that it is by his own authority that he will make his choice. This is because his choice seems radically unconfined,⁴⁶ as the background cares which would normally put boundaries on his deliberation are themselves up for question. He must thus *take up* his authority and take charge of his identity, both as one who cares about things and as one who directs his own life, and find a consideration that he can take as settling the question in a local sense in order to settle it in a global sense.

4.3 *Breaking the impasse*

Now, finally, we can tackle directly the question of finding the consideration that will break Smith's deliberative impasse. How is Smith to determine what he cares about, given that when he arrives at the deliberative impasse, there is (by hypothesis) no fact of the matter as to which option is more important to him and he cannot decide without a reason—that is, given that it is consistent with his reasons *to* care that he choose either way? The answer is that he must, effectively, create a reason in the process of deliberating. It will thus be a “self-creating” reason in two senses: (a) it is created by an existing self, out of reflection on elements that constitute that self; yet (b) it also helps to create and re-constitute the self, for it gives a person reason to reshape those elements in a particular way, thus reorienting his life. Self-creating reasons, as we will see, are dependent on the agent in a way that desire-independent reasons are not.

In the last section I claimed that Smith must undertake a new kind of deliberation in which the question is not so much “what to do” as “who he wants to be,” and as I remarked at the beginning of the chapter, the conclusion will be not “what to care about most” but

⁴⁶ I say “seems” here because in the next chapter I must qualify this claim a bit.

“what *he wants* to care about most”⁴⁷—an expression of his will rather than his reasons. To do this, he must engage in the kind of deliberative self-interpretation that not only assesses and reviews his motivations and commitments, but also engages and shapes them.

The task is now to explain what can go on in such a deliberation, and eventually (in the next chapter) to explain how it comes to have the kind of authority that gives its conclusion the normative status of a reason.

In framing the question before him in the self-regarding terms of “who he wants to be,” Smith is free to view his cares and desires opaquely, no longer scrutinizing them for their rationality. This shift of frame, I contend, is enough to give him a new perspective on the question of which path to choose (i.e., which life to lead). The new deliberative perspective he takes focuses on his attitudes and inclinations with a fresh regard: not as motivations to be examined in light of reasons already discovered, but as expressions of who he is. We might think of it on an analogy to what Thomas Kuhn famously calls “normal” and “revolutionary” science: a discovery-model deliberation is “normal science,” and when she enters “revolutionary” mode, the deliberator shifts her framework for understanding the data, and views new (and old) data in a different light. Normally, a deliberator will view desires as transparent, but when the paradigm shifts, the desires are reinterpreted within the new framework.

We can here follow Charles Taylor⁴⁸ in describing Smith’s process as an attempt at articulating what is most important to him in a very deep sense.⁴⁹ In situations like Smith’s in which the former categories of articulation and evaluation have become inadequate, they become subject to review and revision, and this is precisely because what is of deepest importance to us is not fully fixed. It changes as our lives and cares change, so that one articulation may become out of date. New circumstances then call for a re-articulation of

⁴⁷ It may look strange to say that the answer to the question “who he wants to be” is answered by “what he wants to care about,” but recall from the earlier discussion of care that who one is is constituted by what one cares about.

⁴⁸ “What is Human Agency?”, Taylor (1985b).

⁴⁹ I should note that this deliberative process need not be undertaken in total isolation; certainly a deliberator may wish to deliberate in consultation with others he trusts, as they can sometimes help to articulate thoughts that the deliberator can express only impressionistically and vaguely. In this particular case, for example, Smith owes it to Pocahontas to consult with her, given that his love for her has him seeing her as an individual with cares of her own that merit respect. (Even though the decision in the end rests solely with him, the fact that he does not actually consult her in any serious way is a fault on his part.)

something that is “initially inchoate, or confused, or badly formulated.”⁵⁰ This inchoate quality never fully goes away, even after a new articulation is reached, because it is always open to us to question whether we have fully understood what is most important to us. Taylor observes that the specially problematic character of this sort of process is that

there is...no metalanguage available in which I can assess rival self-interpretations.... On the contrary, the re-evaluation is carried on in the formulae available but with a stance of attention, as it were, to what these formulae are meant to articulate and with a readiness to receive any gestalt shift in our view of the situation, any quite innovative set of categories in which to see our predicament, that might come our way in inspiration.⁵¹

This is the kind of deliberation that Smith must undertake: it is a review of his reasons, and his personal values, but with a kind of indirect attention to what stands behind these considerations.

The self-interpretation process someone like Smith undertakes begins with an honest and accurate assessment of “where he is,” that is, of what he cares about on the level that makes him who he is—of who he identifies himself *as* at the moment in virtue of what he cares about.⁵² Smith, for example, finds himself a lover and an adventurer, a colonist and the king’s subject, an ambassador and a leader, among other things. Upon evaluation he will no doubt see the objective value in all of these things and think he is perfectly right to care about such things. But at this point, objective value is not what matters to Smith; he is concerned with personal value.

We can imagine Smith’s thought process as he deliberates: he rehearses all the reasons he has from an objective point of view;⁵³ he rehearses the importance-to-him of the various goods he would be living with on either path; he rehearses the losses and the depth

⁵⁰ Taylor (1985b), p. 36.

⁵¹ Taylor (1985b), p. 40.

⁵² Note that identifying-as as I am using it here need not be a fully conscious process. People may identify themselves as living out certain roles without realizing that they do so. Such may be the case with Jack the workaholic before we point out to him the toll his workaholism is taking on Jill; until we give him occasion to reflect, he may even consciously *deny* that he places a higher priority on work than he does on his family. But until he comes to see himself more clearly, he will not be able to change. This is why a deliberation of the sort we are examining must begin from an honest and accurate assessment of the agent’s current motivations. This will typically not be a problem for someone like Smith, who arrives at his impasse without any help from outside observers.

⁵³ The objective point of view referred to here is not the same as the explanatory point of view, though both are objective. Here the objectivity is understood in terms of will-independent reasons, which may be either agent-neutral or agent-relative. But either way, the objective perspective is interested in normative rather than explanatory reasons.

of his attachments. He imagines the two different lives he could lead, mentally “trying them on for size,” so to speak, observing how he feels as he projects himself into his possible futures. Captaining his own ship, does he miss Pocahontas so deeply that it’s clear that accepting the commission was the wrong decision? On the other hand, remaining in Jamestown, does his missed opportunity to do what he has always wanted to do continually haunt him as he goes about his daily colonial chores? As he imagines these scenes he observes how he feels embodying the ideals he sees himself as embracing in each scenario. Does the satisfaction of living out the ideal of adventure outweigh the loss of the woman he loves? Or does it feel better instead to live as a pioneer staking a claim in a new world and making a fresh start in his life?

As he answers all of these questions, however, he will also be assessing how he feels about the answers. This is self-exploration and self-interpretation. It is a complex process that involves some amount of self-discovery, but it is not necessarily pure discovery. Self-interpretation, like artistic or literary interpretation, does not generally leave the subject unperturbed; there is no fully objective “way things are,” because interpretation necessarily starts from somewhere and involves a particular point of view which brings its own color to the subject. Over the course of feeling things out, one’s feelings often change, partly because of rational pressures not to have incompatible attitudes, partly because the emotional make-believe reveals new perspectives, and partly because habits and values make us want to see things in some ways rather than others.

Our imaginings never provide us with certainty, of course, and Smith will no doubt imagine many things that do not come to pass, and fail to imagine some that do. His response to his imaginings, however, gives him a sense of his holistic, all-things-considered reaction to his options. Consideration of desire-independent reasons a few at a time, comparing them to one another evaluatively, may bog him down in minutiae, so that he loses the forest for the trees. The imaginative deliberation is a consideration of the whole forest. It is no longer purely reasoning, but something more: Smith is “feeling out” his options in a valuative way by gauging his overall response when he considers each scenario as a whole package. The resulting valuation is not baseless, for it is a kind of response to reasons that are clearly present as he contemplates broadly what each life will be like for him. As Taylor writes:

Our attempts to formulate what we hold important must, like descriptions, strive to be faithful to something. But what they strive to be faithful to is not an independent object with a fixed degree and manner of evidence, but rather a largely inarticulate sense of what is of decisive importance. An articulation of this ‘object’ tends to make it something different from what it was before.

And by the same token a new articulation does not leave its ‘object’ evident or obscure to us in the same manner or degree as before. In the fact of shaping it, it makes it accessible and/or inaccessible in new ways.⁵⁴

The shift from “what to do?” to “who do I want to be?” changes Smith’s access to his reasons, and in doing so, affects the character of what he is assessing. For now it is a quest to define himself, rather than “simply” take action. So he must strive to be faithful to himself, and yet the very purpose of his deliberation is to change who he is.

Suppose that after mulling things over thoroughly in this self-interpretive way, Smith is inclined toward accepting his commission. This inclination is a species of desire that arises as a result of facing the impasse and needing to reconsider the frame of the entire situation; it is something new, which hasn’t been on the table until now, and wouldn’t have been there at all, if it weren’t for these particular circumstances.⁵⁵ This is the outcome of his mulling over his reasons, but it is more a valuation than an *evaluation*. It is not a conclusion that he has most reason to leave Jamestown. It is not a discovery of what he cared most about all along; the considerations he ruminates over are not evidence of his caring, for it is also consistent with these reasons that he care more for Pocahontas and thus choose the other way. Just what does the inclination amount to, then? It is an attraction, a feeling of being drawn in one direction rather than another. It is likely not something Smith could fully explain; if pressed to say why he is attracted to one rather than the other, he might say he is not entirely sure. This is as it should be, if indeed desire-independent reasons have run out.

Phenomenologically, his feeling is probably that something in his thoughts has crystallized. This will manifest itself in his attempted explanation of his choice, which might be that he wants to be the kind of person who lives the life to which it leads. He might say “This is just what I want to do. I can’t explain it,” or “It’s more the person I want to be.” This is not a claim that he will be happier or lead an objectively better life; it is simply a claim about himself.

⁵⁴ Taylor (1985b), p.38.

⁵⁵ (It may still be mysterious how new inclinations arise from this new perspective, but this is because it’s mysterious how inclinations that aren’t the direct result of reasons arise at all.)

The conclusion here is driven less by rational consideration than it is by emotional exploration—his inclination has him seeing the life of adventure as *his* in a way that is independent of any reasoning he can do. It is related to his reasons in that it is by imaginative engagement with them that he comes to see this life in this way, but it is not a cognitive response to these reasons (though it is rational in the sense that it is consistent with these reasons).⁵⁶

Let me pause for a moment to review where we are. Because Smith is at a cognitive and conative impasse, and because he wishes to avoid arbitrariness, he must shift his mode of deliberation in such a way that he can raise in himself a new desire or other motivator that can get him back into action. In the next chapter I will say quite a bit more about why he owes it to himself to do this. For now, I would like to point out that Smith’s imaginative deliberation is constrained in certain ways. He must do his imaginative and emotional exploration by projecting himself into different scenarios and trying to articulate what is most important to him—and *not* by taking a pill that will shift his desires around artificially, or by distorting the situation (say, by associating staying with Pocahontas with having his fingernails pulled out). Reason is involved to the extent that it eliminates certain kinds of imaginings as illegitimate, for the reasons that bring him to his impasse are still in force, but after that it can do no more (or Smith wouldn’t be at an impasse). Within the constraints, the exploration is a matter of imagination and emotion.⁵⁷

Because it arises from an imaginative attempt to answer a question about himself and what he wants, I submit that Smith can take his inclination, which represents “extra” motivation above the reasons *to* be motivated, to express something deep about himself. (In the terms I have been using, he regards the inclination as opaque rather than transparent.) Because it is an inclination and not yet a care, it is not something that yet *constitutes* who he is. He could, conceivably and without contradiction, reject it as not representing him—if, for instance, he slipped back into the discovery-model mode according to which such an inclination is transparent to its object. But given the circumstances under which it arose, and that in these circumstances inclinations are viewed as opaque, there is no reason not to give

⁵⁶ I should mention here that I do not claim that this imaginative exploration is guaranteed to work—that is, there is no guarantee that a new inclination actually will be the result of it. If this deliberation fails, then perhaps Smith is forced into arbitrary picking. But as I argue in detail in the next chapter, he owes it to himself to try his best to find a non-arbitrary way out of his impasse.

⁵⁷ Thanks to Sarah Buss for this line of clarification.

it some credit as expressing himself; and there is every reason to credit an inclination that breaks his impasse. Bestowing that credit amounts to deciding which path he will take.

I contend that Smith's reason—a self-creating reason—to act on this inclination stems not (directly) from reasons to accept the commission, but from his authority to take his inclinations as guides insofar as he is the leader of his life. The thought is that he will do it “because *I* want it” more than it is “because I want *it*.” If someone tried to challenge his decision by complaining that Smith has no *reason* (beyond those already weighed) to want or care about the life he chooses on the basis of his inclination, he might reply that, under the circumstances, it is genuinely up to him how to lead his life, and this can't really be the case unless he gets to count what he wants to care about as a reason.⁵⁸

In choosing on the basis of his inclination, he gives the life of adventure more value to him than it had before, but at the same time he is responding both to the value of accepting his commission *and* to his conception of himself as a chooser who cares about things, which was activated by the shift of frame from “what to do” to “who he wants to be.” Smith has not yet fully brought it about that he cares more about the life of adventure, but he is drawing on his authority and his will to get the process of deepening this care off the ground.

Remember that we have been searching for a consideration Smith can take as settling the question before him so that he has something to act on, agentially speaking, which will in turn settle the identity question that makes his problem so difficult. The inclination that arises from a refocused, imaginative deliberation is such a consideration. And although the normal response to an inclination is to assess it for its rationality, in these circumstances Smith may exercise his authority, as the author of his life, to take the inclination as an expression of himself, and as a result engage his volition, his agency, to act upon it.

And so we have an account of the impasse-breaker which is the source of the reason Smith seeks: it is an inclination toward an identity that seems to the agent to best express

⁵⁸ Thanks to Steve Darwall for this point. Of course, assuming that Smith's choice is genuinely up to him opens my view to challenge by the assertion that it is *not*, after all, up to him how to lead his life. This could mean one of two things. First, it could be a metaphysical point about free will; this I will set aside, as all deliberation takes place under the (possibly illusory) assumption that choice is genuine. The only other way I can think of to understand the objector's point is as a claim about what there is most reason to do. That is, when reasons *do* weigh clearly in favor of some alternative, then there is a sense in which it is not up to the agent what to do. We hear the phrase “It's not up to me” quite often from leaders who must fulfill official roles, or people in situations where morality sets a clear path. But I have set up the situation so that this is not an issue, for I have stipulated that reasons do not favor either alternative.

who he wants to be, reinforced by a conception of self as an authoritative chooser who cares about things. This authority turns inclination into volition, and the agent acts, taking the first steps to reshaping his life.

The astute reader will, of course, have noticed that there is more to discuss here. Although I have given an explanatory account of how the impasse gets broken, I have not yet given a normative one. After all, according to the anti-DBR thesis, an inclination is not itself normative, and even though this inclination arises from a process that lends it a certain amount of weight, we have not yet seen clearly why Smith must defer to the inclination. In the next chapter I will finally turn to this question of normativity: How is this authoritatively-reinforced inclination supposed to give Smith a reason to choose one path over the other? Before I do this, however, I would like to make a digression to show in more detail how it is possible to effect the reorientation of a life involved in acting on the inclination described here.

4.4 *How to reorient a life: ideals*

The problem that faces Smith can also be characterized in terms of what Elizabeth Anderson calls “ideals.” Because ideals are a familiar concept, and because of their ties to both reasons and cares, they will be useful in discussing how to bring about the reorientation of a life. For reasons that will (I hope) become apparent, in what follows I will assume that choosing what one wants to care about amounts to choosing an ideal for oneself. I thus begin this section with a short exposition of the concept.

Anderson characterizes ideals as second-order desires that are expressions of deep carings (and not just ordinary or weak desires and valuings).⁵⁹ In particular, she writes that ideals are

conceptions of what kind of person [one] ought to be, what kinds of character, attitudes, concerns, and commitments [one] should have. ...[They] are the objects not merely of desire, but of aspiration. ...[T]o call a self-conception an ideal is not necessarily to endorse it, but to imply that it is a possible object of admiration or condemnation, honor or disdain, and that people who adopt it regard it as worthy. Ideals set the standards of conduct and emotion people expect themselves to satisfy with regard to other people, relationships, and things. ...Ideals give us perspectives from which to articulate and scrutinize the ways we value things. ...An ideal is constitutive of a person's identity if it governs her self-assessments and her

⁵⁹ *Value in Ethics and Economics*, Anderson (1993), pp. 5-8.

responses to her achievement and failure and if she uses it to discipline her desires and frame her choices.

Holding oneself to an ideal involves a certain reflection on the things we care about; to hold an ideal is to care about your carings. An agent might find herself caring about something, but upon reflection realize that it is not something she *wants* to care about (or thinks she should care about), because it is incompatible with an ideal she has set for herself. An obvious example is a religious adherent who is unable to suppress her desire for certain religiously forbidden physical pleasures. Similarly, someone may hold an ideal and berate herself for not living up to it by caring about the things it directs her to care about and acting accordingly; perhaps a workaholic mother upbraids herself for not making more time for her family. If carings constitute a person's identity in the sense of who she *is* at the moment (how she is actually motivated and how she actually identifies herself), ideals characterize the identity a person strives to have. (These identities may, of course, coincide.) Ideals, then, are normative because they are regarded as⁶⁰ connecting to value that goes beyond the merely personal. They are able to give a life meaning because they relate a person to something she can see as greater than herself.⁶¹ Thus, ideals normatively—if not always descriptively—govern cares.

Yet there is nevertheless something quite personal about ideals. As I rehearsed in the first part of this chapter, even within the constraints of morality there is more to value than one can ever actively attach oneself to. There is a whole range of permissible projects, and it is our choosing among them that helps us to carve lives for ourselves. In conceiving of ourselves as striving to live by particular ideals, we conceive of ourselves as connected to intrinsic value. Yet at the same time, the ideals we live by are *ours* in very personal ways, precisely because we are able, in guiding our lives by ideals we choose, to define ourselves as selves.

Of course, we do not always choose our ideals. We may inherit them from social forces; we are clearly influenced by our parents, our education, and the greater culture in

⁶⁰ “Regarded as” is an important qualifier here. It is not inconceivable that someone may hold an “ideal” which is not actually worthy at all. Deciding whether to say that this is no ideal in the technical sense of the term, or whether to say that ideals need not have any actual connection to intrinsic value is an interesting but difficult task, and one I will not undertake here, for it is not necessary for my present purpose. For the task of understanding cases like Smith's, it is safe to restrict ourselves to ideals that are in fact worthy.

⁶¹ We have seen a similar idea already; see note 32 above.

which we live. Forces like these may influence our religious, civic, and character ideals. We may also grow into ideals simply as a matter of contingent availability: we cannot hold as ideals things we never encounter in any meaningful way. Nevertheless, as Anderson writes:

What things a person cares about, as well as how and how much she cares about them, are not solely a function of the social practices, roles, and relationships she participates in. Her character, history, mood, energy, actions, and reading of her predicament play a profound part in influencing what she values, especially in influencing which practices, roles, and relationships she will make her own, how she will interpret, criticize, and change them, and so forth.⁶²

In particular, the process of growing up is partly a process of reflecting on and taking charge of the way one's life will go, and this process often involves reaffirming or rejecting ideals one finds oneself living up to. Such affirmations and rejections are choices of ideals. The process of growing up, insofar as it is the process of choosing ideals, is the process of taking up the authority one has to lead one's own life.

Choices of ideals arise in other situations as well, and such choices all involve taking up the authority one has to lead one's own life. In the next chapter I will say more about the reasons one might invoke in choosing ideals; for now, my task is simply to understand how the reorientation of a life involved in implementing the choice of ideal can be effected.

At this point I would like make a short aside, to point out and resolve an ambiguity in the phrase "adopting an ideal." As we will presently see, there is an easy or weak sense of adopting an ideal, which we might identify as the choice of an ideal. Weak adoption is what an agent does when she decides to change something about herself or the way she lives her life. Following from a successful weak adoption, however, will come the difficult or strong sense of adopting an ideal, when the agent has brought it about that she truly lives by the ideal, i.e. that she has brought her cares (and the habits of behavior, attunement, and emotion they engender) clearly into harmony with it—when, in short, it is part of her identity. In each case, Anderson's description of ideals as setting standards of self-assessment is appropriate. But I contend that only in the strong sense is the ideal truly constitutive of her identity.

With this in mind, let me describe Smith's situation in terms of ideals. After the initial description, I will undertake a discussion of both strong and weak adoption individually.

⁶² Anderson (1993, p. 25).

Metaphorically, we can think of Smith as standing at a fork in his life's road. Right now, having come to see the conflict between two of his important, identity-conferring cares, he stands at the fork. This is "where he is" in the following sense: in caring about each of his options, he identifies with certain objects and hence with certain projects and ideals. (This is "where he is" in an identity sense.)

In particular, at the decision point in the story, Smith finds himself caring on a first-order level both about Pocahontas and about exploration. He has long had the ideal of an explorer-adventurer's life, but he finds that his deep care for Pocahontas embodies a second ideal that involves remaining in Jamestown, having a family, working toward cultural understanding between the colonists and the natives, etc.

Until receiving the news of the royal commission, these ideals were compatible; in living as a colonist and putting down roots in the New World, he could see himself as an adventurer, since the life of the colony was the most adventurous one available to him. Thus, his cares all aligned with the ideal of adventurer. With the commission, however, his cares point him in different directions—out to sea, on the one hand, or firmly in Jamestown, on the other. He now has occasion to reflect and separate cares and ideals that he previously had no reason to sort out, and because of his love for Pocahontas he discovers the possibility of living by a new ideal not previously salient to him. Both of the ideals are already in a sense his, relating as they do to things he cares about. But he must now decide to (weakly) adopt one of these ideals to the exclusion of the other, because they are now incompatible. Weakly adopting one or the other is done, however, with the intention of strongly adopting it. This will involve adjusting what he cares about on the first-order level.

In order to go on, then, Smith must choose one of the branches of the fork; he must put himself somewhere else by exercising agency to change his identity. In the terms I have been using in this section, he will do this by adopting a new ideal, and I will argue in the next chapter that he can choose the ideal non-arbitrarily. In doing so, he does not instantly alter the cares that constitute his identity, however. They would not be cares if they could be so easily shed. Nonetheless, there is something he can do to influence what he cares about, as we will presently see, and this influence amounts to the reorientation that is the subject of this section.

4.4.1 *Strong adoption*

Let me begin by illustrating with an example. Imagine Jack, who finds himself really caring about one thing (perhaps his career), but wanting to care about something else more (perhaps his family). Jack is such a workaholic that we, his friends, might say that his job is “who he is.”⁶³ We still like him, but we can see that things have gotten out of hand lately as he competes with a colleague for some juicy promotion, and it’s taking its toll on his wife Jill, who has been struggling to manage the household chores and the kids’ schedule (not to mention her own career) without his help. Jill has always been a pillar of support, happily climbing the hills of life and making many small sacrifices for Jack, but since we’re friends with her too we can see the signs that she is beginning to crack under the strain. Jack knows that he would be nowhere without her, so he doesn’t want to lose her. But at the moment, as things stand, his career is really the most important thing to him. It is what he cares most about. He lives his life as the guy who lives for his work.

Because Jack knows he would be nowhere without Jill, however, he wants to change this. He wants to care more about his family than he currently does; he doesn’t want his job to be the single central thing in his life anymore. Effecting this change, as I have already indicated, would involve the exercise of agency to change his identity, which is constituted (mostly) by his first-order cares—the things he actually cares about as he currently is. It will take time and patience, but it is possible, as I will argue.

Suppose, then, that Jack decides his family is going to come first: he (weakly) adopts a family-oriented ideal for himself. In order to turn this weak adoption into a strong one, he is going to have to change his long-term pattern of caring most about his career to make more room for family activities. At first, he will still have the patterns of caring that give rise to desires and behavior that draw him toward work, but he can decide on the agential level not to endorse these. So in a sense they’re “him,” but in a sense they’re not. They’re “him” in an identity sense, because that is where his cares (which are distinguishable from immediate desires) as a matter of fact lie right now, but they’re not “him” in an agency sense because he makes an effort not to act on the immediate desires arising from them, rejecting those desires and their derivative motivations. If he succeeds in not acting on these cares,

⁶³ The phrases “who he is” and “where he is” are meant to mean more or less the same thing; the difference in locution only marks a difference in the metaphor being used, and hence which word is appropriate in a particular context.

the hope is that eventually the cares will weaken and he will be less motivated to act according to his workaholic patterns.

To show how this can work, I will draw on work by Bennett Helm and David Velleman. Helm provides us with an account of the dual nature of “import” or personal value as both objective and subjective, with its ties to both reason and emotions. This will help us to see how it is possible to affect the value things have to us. Velleman, then, provides a practical sense of the psychology involved in reorienting one’s life.

Before I come to these accounts, however, let me begin with a sketch of how the process works in real life. Given the emotional quality of care and the attachments and commitments it engenders, letting go of a care in the service of reorienting one’s life is not an easy task. Because cares play a constitutive role in an agent’s identity, making such a change requires giving up a part of oneself, and as grief at the loss of a loved one shows, this is wrenching. But as the analogy to grief also shows, it is possible to heal. The process is slow, but little by little one can occupy oneself with other things, easing the sting of the loss. We tend to think that moving on too quickly shows a lack of depth, both of feeling in the bereaved and of appreciation for the lost person or object. On the other hand, however, it is possible to dwell too much on the loss, thus failing to recognize the worth of one’s other commitments and the rest of one’s own life. A person must not repress grief, but neither should he wallow in it.

So day by day, someone who is healthy takes the feelings of loss seriously, but at the same time he seeks to gain a new perspective that puts the loss in a new place in his life—to learn to make sense of his life without the object of his grief. This happens best, perhaps, by deliberate action. He takes on the mundane tasks of eating, cleaning, and working. Eventually he is healed enough to consider what to do with the traces of the lost person or object. Finally, there is sometimes a moment when he performs a symbolic action of letting go. Such a ritual does not have to mean forgetting his grief or the importance the lost person or thing had in his life, but it does mean that he has made a kind of peace with it and is whole again.

This process is something Bennett Helm⁶⁴ calls “freedom of the heart”—our ability to exercise rational control over what we care about or value (remember that caring and

⁶⁴ *Emotional Reason*, Helm (2001).

valuing are characteristically emotional states). Helm thinks that the distinction philosophers often draw between the cognitive and conative aspects of emotions does not exist; rather, emotions, as “felt evaluations,” involve elements of both cognition and conation and thus blur any distinction there may be. “[E]motions and import are a conceptual package, with neither prior to the other.”

A felt evaluation involves experiencing some aspect of the world as import- or value-laden, since an emotion is the feeling of “things going well or poorly,” according to Helm.⁶⁵ On the one hand, felt evaluations have an affinity for the cognitive in that they are objective and merit responses from us. Felt evaluations display their responsive aspect in the following way: insofar as emotions are feelings that an object has a certain import, you are committed to having other similar feelings in similar circumstances, lest you be mistaken about the import the object has, or fail to do it justice. This indicates an aspect of discovery, in that you feel import impressed upon you. On the other hand, felt evaluations have an affinity for the conative in that they are subjective and constituting. The object would have no import if it weren’t for its connection to things of concern to you. The focus of the felt evaluation⁶⁶ is intelligible as worthy of attention because of the emotional commitments defining your overall pattern of evaluation and response to the object (and your being attuned to these commitments). That is, there is an aspect of invention as well, in that an object could not have import without your responding to it in broadly consistent ways.

Felt evaluations are, according to Helm, a kind of passive assent or commitment to import.⁶⁷ It thus stands in contrast to an evaluative judgment, which is an *active* assent or commitment to import. Helm argues that felt evaluations together with evaluative judgments make up a single evaluative perspective.⁶⁸ They are very similar in that both felt

⁶⁵ In general, the immediate feeling of “things going well or poorly” would be an “emotional episode,” rather than an emotion, as I was distinguishing these terms above. Helm does not distinguish them, however, and since the focus here is different, I think it will not cause confusion to follow his usage in this section, in order not to encumber the language too much.

⁶⁶ The “focus” of a felt evaluation is the object that has import to the subject. This is to be distinguished from the “target” of the felt evaluation, which is the immediate object of emotional response. While the target and the focus of an emotion are often one and the same object, they can come apart. For example, I may be angry at the person who dents my classic car in a parking lot. The target of the emotion is the reckless driver, while its focus is my car.

⁶⁷ Because of this lack of review and judgment, it may be more apt to call them “felt *valuations*.” It is disputable whether Helm would be friendly to this amendment; the “valuation” locution is probably closer to what I want to get at. Nevertheless, I will stick to Helm’s vocabulary in my exposition of his view.

⁶⁸ For further discussion, see Helm (2001, ch. 5).

evaluations and evaluative judgments blur the cognitive-conative divide, since they are commitments to import, which is constituted by and simultaneously serves as a standard for the evaluation. (The main difference between them is activity vs. passivity.)

Evaluative judgments and felt evaluations can come apart, however, and often do. It is not unusual, for example, for someone to say “I am angry [or afraid, or relieved, or...] but I realize I shouldn’t be.” Many times, sincere evaluative judgment corrects felt evaluation. Someone who judges the anger she feels to be inappropriate may make a point of reviewing and rehearsing all the reasons that her anger is misplaced, and this often helps to assuage it. But felt evaluations can also correct judgment by signaling a need for a re-evaluation, if they are persistently recalcitrant.⁶⁹ I know of a case where a man’s wife and daughter convinced him to buy a house he didn’t think they could afford, and he was unable to become comfortable with the decision. His growing orneriness over the few years they lived there gave them occasion to review the purchase. His short-temperedness turned out to be a signal that they had made the wrong choice; as soon as they sold the large house and bought a smaller one, he became his old cheerful self again. Thus, felt evaluations can correct evaluative judgment, and again, evaluative judgments and felt evaluations are not as independent as common accounts would have us believe.

This conclusion has implications for our rational control of our actions and emotions—freedom of the will and heart, as Helm writes. Controlling akratic desires, for instance, requires controlling the way import impresses itself on you—i.e. controlling your evaluative perspective. So to control an akratic desire, one needs to reconceptualize the object of the desire in terms of reasons that support one’s judgment. Normally this will work, he thinks, because both judgments and desires are part of a broad pattern of reasons to which you are attuned.⁷⁰ Freedom of will, then, amounts to “our ability to motivate ourselves directly by virtue of an active, judgmental recognition of and commitment to import and the reasons it provides for action.” And freedom of the heart is our ability to exercise rational control over what we care about or value. It flows from freedom of will; success in changing habits of attunement and behavior, achieved by will, will normally

⁶⁹ See D’Arms and Jacobson’s “The Significance of Recalcitrant Emotion (or, Anti-Quasijudgmentalism)” (2003).

⁷⁰ There are connections here to Aristotle’s account of habituation in the *Nicomachean Ethics*.

succeed in changing the heart (assuming the requisite quality of reasons supporting the change).

This is the sort of idea behind David Velleman's thinking in "Motivation by Ideal."⁷¹ In this paper he suggests that adopting an ideal involves becoming by pretending. That is, coming to strongly adopt an ideal, once you have weakly adopted it (say, by exercising sincere evaluative judgment), involves not just admiring those who embody the ideal; it instead involves a pretense in which you imagine yourself already having achieved the ideal, thus imagining yourself as having and acting from the motivations (and perhaps emotions) of someone who already embodies the ideal. Adopting such motivations will tend to shift your perceptions of the world, making you more attuned to aspects that provide reasons in your new perspective that you would not have noticed or interpreted in the same way before. You are thus able to make sense of your actions in a way different from before, and may in this way facilitate actually matching the adopted ideal.

Velleman illustrates this with the example of a smoker trying to quit. Since the costs of smoking the next cigarette never outweigh the benefits in the short-term, the smoker will take his craving as a reason to have a cigarette as long as he continues to think of himself as a smoker. But knowing that in the long run the costs of smoking outweigh the benefits, the smoker has to motivate himself to do something that is, in the immediate context, irrational. This is where thinking of himself as a non-smoker comes in. Actions that would not have seemed or felt rational from the point of view of the man-as-a-smoker thus become rational from the point of view of the person he wishes to become—the man-as-a-non-smoker, who is the same man, but (as he sees it) improved in virtue of having adopted the new ideal. To rephrase this in other terms I have been using, agents can change their identities by exercising their agency.

Someone like John Smith can go through such a process. Smith ultimately chooses to leave Pocahontas in Jamestown. Perhaps he decides that he wants to care more about exploration than he does about Pocahontas; we could say that he has (re-)adopted the ideal of being an explorer. Having made this decision, he must now take steps to bring his felt evaluations, his cares, into line with it. He must diminish the care he has for Pocahontas in order to make sense of and live peacefully with his decision, even though it is clear that

⁷¹ Velleman (2002).

Pocahontas is very important to him. He thus moves from an assessment of the person he is (what he actually cares about) to the weak adoption of an ideal which will shape the person he becomes. From there, he can effect a reorientation by exercising agency to shape his identity, thus strongly adopting the new ideal.

4.4.2 *Weak adoption*

The story I have just told is what may take place as an agent tries to turn the weak adoption of an ideal into a strong adoption. There is some more to say concerning just what a weak adoption is, however. Smith is not *just* following an inclination, but quite deliberately making a life-leading commitment⁷² to an identity on the basis of that inclination. In outline, my argument is that commitment to the inclination constitutes the weak adoption of an ideal, and since ideals are normative because of the way they work as standards for regulating actions, attitudes, and cares, Smith has a reason to reorient his life according to the ideal he (re-)adopts. Less abstractly, the argument is that because there is good reason to undertake the new deliberation, and raising the new inclination is the *point* of that deliberation, it would be a mistake to ignore the inclination that arises. Insofar as he avoids this mistake, Smith has made a commitment to order his life according to the ideal to which it points. Since it is a genuine commitment, it serves as a normative constraint on further choices; this is part of the function of a commitment. In the next chapter I will argue that the willed commitment Smith makes is normative because it is issued from the perspective of the authority to govern the agent's life; the piece I wish to argue now is that the commitment constitutes the weak adoption of an ideal.

Let us begin with an examination of commitment. In what does a commitment like Smith's consist? On a first pass, it is a certain kind of attachment to a project, person or object. Many commitments are implicitly made in virtue of caring (on the first-order level) about the project. Some are explicitly made *despite* first-order caring; they represent something second-order. Take Jack from the example above. He is a workaholic on the first-order level, but he commits himself to becoming more family-oriented.

⁷² A willed commitment such as Smith's is similar to Helm's "evaluative judgment" in that it involves active assent. Thus weak adoption is the ideal-level analogue of evaluative judgment and strong adoption is the analogue of felt evaluation.

This exercise of agency is the external manifestation of an inner mental state of “commitment.” The real question is what this mental state consists in. A commitment to a project is more than a desire to undertake it. It is something more like a promise to oneself so that one is accountable for failure to make sincere efforts on its behalf. In fact, promising is closely analogous. One can commit oneself to another by making a promise, for part of what it is to make a promise is to make oneself answerable to the other. So the internal commitment one makes in adopting an ideal makes one answerable to oneself for any failure to realize it.⁷³ This is why the commitment amounts to the weak adoption of an ideal: ideals provide the standards of behavior and care that one takes as normative for oneself in making a commitment. The prospect of feeling guilty then tends to invest one emotionally in furthering the project, and it tends to give rise to motivations to take action on the project. Among these motivations are those such as attunement to opportunities to further the project, valuing states of affairs which are conducive to the success of the project, evaluation of different courses of action in light of the ongoing project, and so on.

I say that the willed commitment amounts to the *weak* adoption of the ideal because being committed to an ideal is different from actually living by it, as I observed above. The kind of commitment or emotional investment made here is not a full-blown care, because such care takes time to generate. But it is a commitment to adapt one’s first-order cares to an ideal one adopts. A weak adoption such as Jack the workaholic’s is a promise to himself to become less concerned with work and more concerned with his family. It signifies that he is concerned about what he cares about—that is, that it matters to him on a higher-order level what his first-order cares are. He weakly adopts the ideal in order to change his first-order cares because he wants these cares to be different, i.e., he wants to be a different sort of person. And this wanting is no mere desire, as many of us express when we say with a sigh, “I wish I were a tidier/nicer/more energetic/etc. person.” This wanting is deeper. It is

⁷³ I am aware of the puzzles surrounding obligations to oneself, and what the force of this “answerability to oneself” could be. On the one hand, it seems you can obligate yourself because it is a contradiction of the will to break a promise to oneself, so you must do as you promised on pain of irrationality. On the other hand, however, it seems you can legitimately change your mind—but then in what sense were you obligated? G.A. Cohen provides a brief discussion of this in his response to Korsgaard in *Sources of Normativity* (1996). Without launching into a full discussion of the topic, which would take us far afield, I can suggest that the answerability I mean here has to do with attitudes of blame and guilt, such that unless or until you give up your commitment to do as you promised, you will tend to feel guilty and blame yourself for not living up to your ideal. This does leave it open to you to give up the commitment, but it does seem to require that you do so only with a good reason, and not as the whim of the moment or as a matter of weakness.

a concern arising from the way Jack evaluates the different kinds of person he could be in comparison to the person he realizes he currently is.

This weak adoption of an ideal is the commitment Smith arrives at through the deliberation described above. In deliberating, Smith has begun, through his new inclination, to invest himself in the life of the sailor, but more importantly, he has committed himself to become further invested in it.

We can get a better grasp on the mental state of weak adoption by comparisons to other, similar states. Weak adoption of an ideal is not the same as, for instance, an emotional revelation in response to a sudden traumatic event. For example, suppose that, while he is deliberating, Pocahontas nearly drowns. When the news reaches Smith that she is missing, he is severely shaken and realizes as a result that he could never leave her, a conclusion which is sealed when she is rescued and he is able to hold her in his arms again. This would not be a resolution that grows from deliberation. The triggering event simply provided the occasion for Smith to notice just how deep his love for Pocahontas reached,⁷⁴ and it would be obvious to him that he must stay. It would not make sense for him to go, and indeed, he could not realistically bring himself to do so.

Similarly, Smith's weak adoption differs from what goes on when, for example, someone is convinced by a friend, or his therapist, that he cares about something. This sort of thing does happen to deliberators; we often use others to help us reflect and then resolve to act based on their judgment. In such cases, the friend has done something to help the agent see where his cares really lie; this is a kind of self-exploration, and the reflection can result in a new resolve to act on his care. An agent in such a position is not making a new commitment, however. Moreover, this is not a proper description of Smith's situation. Smith does not discuss his problem with a close friend, who finds it blindingly obvious by his description of the situation what he really wants to do, and tells him so. In such a case, by the friend's lights, only one option makes sense. But again, this is not the situation Smith is in; as we have said, any friend would find no fault with either choice.

Finally, Smith's commitment is not an arbitrary self-launching of the sort that Sartre seems to indicate in his famous story of the young man choosing between caring for his ailing mother and joining the French resistance. This kind of self-launching, I take it, is

⁷⁴ Or, perhaps, the close call made his love deeper.

supposed to be done fully independently of reasons, simply as “plumping for” one of two competing desires, because reasons cannot tell him what to do. I agree that desire-independent reasons can run out; this is what I have been arguing. But this does not mean that a decision in such a situation is fully as arbitrary as flipping a coin or rolling a die. As I outlined above, Smith is trying to bring to articulation something that is as yet an inchoate sense of himself, unlike the young man as Sartre portrays him. Smith is making a choice that is continuous with the paths he has been on, not just in the sense of being a continuation of his story rather than some wholly random act (such as taking a vacation in the Riviera), but in the sense that it is a further articulation of himself and his values. Unlike the young man, Smith’s response grows out of considering the relevant reasons on each side, though it goes beyond those reasons.⁷⁵

I have said that if pressed to explain his decision, Smith might say something like, “I want to do this because that’s the kind of person I want to be.” If someone challenges him by asking why be this kind of person, I see him as having three kinds of responses. (1) He can cite the reasons *to* be that person. This is a strategy we often take, but under the circumstances it does not suffice as a justification because it does not provide a reason to choose one path *over* the other. So even if he says this, he cannot seriously mean it to be the full explanation. (2) He might, then, say he chose it for no reason. This is true if ‘reason’ is meant only in the anti-DBR, desire-independent sense. More likely it means that he doesn’t know how to explain his reason, since it is really based on a feeling arising from a holistic response to the options. So saying “no reason” really amounts to (3), saying he chose it because it felt right.

The point, then, is that Smith’s weak adoption of the explorer’s ideal, this commitment, is not a response dictated or required by his desire-independent reasons. It is not something anyone else could point out to him as being *the* thing to do, and it is not a discovery of what he wanted all along, revealed by his reaction to the evidence. Rather: he has come to it as a result of engaging with the reasons to care, but it is consistent with the relevant considerations that he not choose this particular way. The concluding resolution is not merely a response, but a responsive investment, an articulation of what is most deeply

⁷⁵ Although I disagree with Sartre’s interpretation of his own story (as he also dismisses feeling as a legitimate basis for this choice), I do not ultimately disagree with the point that we are responsible for our choices, and to pretend that we are not is a kind of bad faith. See Charles Taylor’s “What is Human Agency?” (Taylor 1985b).

important to him. This is because his consideration of the relevant reasons is not only rational, but (we might say) phenomenologically valuational, involving imagining life under the different options and projecting emotional responses to them. He is not considering the reasons from afar, metaphorically speaking, but from within his point of view as an agent who must choose a life. And he must come to credit his feelings as far as he can articulate them because there can be no *judgment* to credit. That is, the inclination embodied in the new commitment acquires a distinctive weight by being a deliberatively informed attempt at articulating the personal value of his options.

What if Smith does not go through all of this self-reflective deliberation and simply throws himself into one of his alternatives, *a la* Sartre's young man? If a commitment like the one that constitutes the weak adoption of an ideal gives *some* reason for action, then hasn't Self-launching Smith committed himself to one life rather than the other simply by taking action, and isn't this then reason enough for his choice?

I may concede that such commitment by self-launching amounts to a reason, but I contend that if it is, such a "mere" commitment is not as strong a reason as the weak adoption of an ideal. Self-launching Smith would not be acting as reason-responsively as Deliberative Smith does, because he does not adopt the new life as the embodiment of an ideal he has set for himself; he is not committed to governing himself by its standards. All he has committed to is closing the door on one of his options. Furthermore, the reason constituted by the self-launching commitment is a sort of *post hoc* rationalizing reason: "Well, I'm here now, so I'd better make the most of it." A self-creating reason is not backward-looking in this way, but forward-looking and proceeding on the grounds of justification rather than reaction.

In summary, the weak adoption of an ideal amounts to changing the way one governs oneself on an agential level so as to come to identify oneself that way on an identity level. Having weakly adopted an ideal, someone like Smith must then undertake to bring his actual first-order cares, the ones that constitute his identity, into line with this. I have already indicated how this is possible, but I should add here that it is not inevitable that Smith will manage to reorient his life as he decides to. The decision here does require investment,

because the impersonal value of the choice alone will not be enough to sustain him if he does not manage to adjust his care to his adopted ideal.⁷⁶

4.5 *Conclusion*

Over the course of this chapter I have done two main things. First, I have given an account of care as an emotion constitutive of identity, in the service of showing what makes a predicament like Smith's so difficult. Second, I have given an account of what could break Smith's impasse in order to show how someone in a predicament like Smith's might plausibly pull himself out of it. The analyses I have given here, however, have merely told a psychologically plausible story; they have not treated the question of how the process described may be normative. This is the topic of the next chapter.

⁷⁶ In the next chapter I will discuss in more detail whether and how a choice made on the basis of self-creating reasons can be mistaken.

Chapter 5: Self-Creating Reasons: An Analysis of Normativity in Willing and Identity

Having described the deliberative process that an agent like Smith can go through to break his impasse, it is time to show why this procedure has the kind of normativity necessary to give its outcome the status of a reason.

As we saw in the last chapter, care and its ties to identity are very important because they give the impasse the particular difficult character it has. Because the point of this chapter is to understand the normativity of the way forward that I outlined in the last chapter, and because I identified care as a major source of the problem, I begin here with a discussion of the normative standing of a person that arises from care. This normative status is based, I suggest, in respect-worthiness. Since respect-worthiness is a desire-independent notion, it gives us some traction with the problem posed in Chapter 3 even for those who accept the anti-DBR thesis.

This initial discussion will put us in a better position to understand just why an agent like Smith owes it to himself to undertake the new deliberation described in the last chapter. And this gets us some way toward understanding why the inclination that arises from this deliberation has the particular status it does, namely, status to provide the basis of a life-leading commitment. But just having this status is not enough to understand the normativity of the commitment made on its basis—why one should guide oneself by the results of the exercise. For this, I must explain and invoke a further element of the picture, namely a person's authority to direct his life. It is respect for this authority that completes the account of the normativity of self-creating reasons.

5.1 Care and status

The basic idea of this section is that a person has a certain status in virtue of having an identity, and she has identity in virtue of having cares. So a person has status in virtue of

having cares. That is, cares—not agency, if we reserve the term “agent” for those persons who are fully autonomous and capable of making informed, reflective choices—are what give a person the most basic kind of normative standing. In the end, it will turn out that only full agents can have self-creating reasons, but this is because only agents, with their self-reflective capabilities, can have the kinds of problems that require resolution through self-creating reasons. It is not, at least in the first instance, because of any additional normative authority that agents may have in virtue of their reflective capabilities.¹

In discussing the relationship between cares, identity, and status, there are two related but separate ideas to be distinguished. The first is that the agent’s cares themselves have a kind of standing that requires a certain acknowledgment from the agent even if she ultimately rejects those cares as worthy to constitute part of her identity. She must engage with her cares. The second idea is that the agent has a kind of standing with others insofar as she has cares in the first place.

5.1.1 *The standing of cares with an agent*

What kind of standing could “bare” cares have with an agent? Certainly, such cares tend to assert themselves, usually in the form of particular desires, directing an agent’s attention toward their objects (or toward the means of satisfying them). Such particular desires can be insistent. As we have seen, however, mere desires do not have any normative force in themselves; they often provide no reason whatsoever for action. But their insistent nature sometimes calls for *some* action on the part of the agent, if for no purpose but to extinguish the desire so that the agent may get on with his life. This action may be deliberately distracting himself, or separating himself from the desired object or the means to indulge the desire, in order to learn that he can live without satisfying the desire or in hopes of causing it to atrophy. Or it may be trying to redirect the desire toward a project he can endorse, bringing the desire under clear examination, etc. This is the sort of thing psychotherapy is aimed at: discovering and acknowledging desires for what they are, in order to deal with them more directly and effectively.

Although desires themselves have no normative force, sometimes desires indicate a deeper root cause: they are signs that an agent cares about some object. One may still wish

¹ E.g., competence to make binding contracts, and other legal and medical decisions.

to fight desires that arise from a deep care because one may not endorse the care in light of other cares. But notice that none of the courses of action one could take to fight a desire requires that the agent be alienated from the care that gave rise to it. Calling a psychic element “alien” describes it as something out of the agent’s control and external to her point of view, like a foreign force trying to take her over. Physical addictions and the desires they give rise to may sometimes be such alien forces, for example, but I contend that a genuine care and *its* resultant desires is not. It is not an external force trying to take over an agent, but rather a part of herself begging to be satisfied. In light of the other things an agent cares about, she may not endorse the care and hence may have a strong, rational motivation to get rid of it, but that does not make it any less a part of her or her identity at the moment, and this is what gives it standing with the agent.²

It is important here to understand just what this standing amounts to; it is not necessarily standing to require fulfillment, i.e. that the agent do what the care gives him reasons to do. But because the care constitutes part of who the agent is, it cannot be taken lightly. We might liken the difference between an addiction and a rejected care to the difference between a conquering society’s imposing restrictions on the vanquished, and a society recognizing alarming trends within its own culture and imposing restrictions on itself in order to remedy the problem. The latter may recognize the disapproved trend as growing from their own way of life, and hence as their own, but realize that they wish to frame a different way of life for themselves.

Take the example of the abused woman who knows that for her own safety and peace of mind she must leave her husband, but who must struggle against her love for him. The part of her that loves him is something she feels as an insistence that she remain where she is, and perhaps it even goes so far as to bring her to assent (passively) to the feelings of worthlessness his abuse induces in her. She comes to believe that she could not make it without him and is therefore too afraid to leave. Nevertheless, there comes a point (perhaps a particularly violent incident that lands her in the hospital where a social worker has a

² I should emphasize that the kind of standing I have in mind here bears an analogy to what Darwall has called “recognition respect,” in contrast to “appraisal respect.” According someone the respect of recognition as someone with a fundamental and basic standing is different from respecting him or her as a model to be admired in some way. The point I am making here in no way requires appraisal respect, but merely recognition. (See Darwall (1977).)

chance to talk with her) where she comes to realize that she cannot let this situation continue and resolves to find a way out.

Part of doing this, of course, will be to stop caring so much for the man who abuses her. I insist that this struggle does not alienate her from her care, however; to think so is to misunderstand and trivialize the woman's difficulty.³ This care has been a guiding factor in her life for too long to say that it isn't her. The thought she has is not that some behavior of hers is somehow inexplicable and therefore not "hers." The thought is, rather, that her actions *are* hers, but they are the unhealthy result of untrue beliefs and misdirected feelings about herself, her husband, and her overall situation. Her remaining with him all this time may be inexplicable in the sense that there is no justification for it, but it is not inexplicable in the sense that it is foreign to her, a desire that just swept her off her feet before she could think better of it. Not to recognize the fact of her love would be to repress a vital part of her. It is a part of her she wishes to change, upon reflection, but she must find a way of changing it that accords it at least minimal acknowledgment.⁴

She can acknowledge this part of her without taking that recognition to give her reasons to act as the care directly gives her reasons to (e.g., to fix her husband's meals, to obey his orders to stay home when she planned to go out, etc.). Her care continues to give her reasons to act kindly toward him, but given his treatment of her, these reasons are simply defeated, in much the way pleasure is a reason that is defeated when it is pleasure at causing another's suffering. Such reasons defeated by another aspect of the situation are no longer seen as having their force, though they may provide reasons for entirely different actions. In this case, the woman's care gives her reasons to see a therapist, move to another town, join a support group, make new friends, etc. She no longer lets the care have the place in her life that it once did because it has been defeated by other vitally important concerns.⁵

³ Here I agree with Jaworska in opposing Frankfurt, who claims (in *Reasons of Love* (2004), p.91, among other places) that in deciding to reorient himself around one or another of his cares, thereby making it more important to him than another, it becomes external to him or alienated from him. It devalues the care and hence demeans the agent who cares to treat it as something which no longer belongs to her. Over time, perhaps it won't. But this says nothing about whether it is *her* here and now.

⁴ An apt analogy here is to the difference between slamming the door in the face of a loved one who has, say, a drug problem that has caused a lot of damage to your relationship with her, and treating her kindly enough to assist her to get the help she needs. The point I am trying to make is that care has standing not to have the door slammed in its face; rather, it deserves recognition, if not to fulfill it, then at least to get therapy for it.

⁵ There are many similarities between recovering addicts and people who are "recovering" from misdirected cares; very little of what I say here is not also applicable to the case of a recovering addict. But this is because a recovering addict has emotional patterns very like cares as well as a physical addiction to overcome. There *is*,

Consider two other examples. A man whose wife died of breast cancer while still in her forties may be mired in grief and therefore be closed off from experiencing new love. Perhaps he meets another kind, caring woman and they gradually grow close, but his clinging to his unresolved grief inhibits him from fully giving himself to his new relationship. Things come to a head over some relatively minor incident, and the man realizes that he must let go of his grief in order not to lose the good thing right in front of him. He may then perform a symbolic act of closure, perhaps by scattering his wife's ashes or taking off his wedding ring.⁶ Or consider a young, silly girl who fancied herself in love until her beloved publicly snubbed her. She may have a hard time getting over her disappointment, but eventually resolve to let go and mark the resolution by burning a number of small keepsakes she had up to this point treasured.⁷ In these cases the need to complete the emotional cadence provides a reason for these characters to do what they do. It is not that they are alienated from the care, but that they need to give it a different place in their lives (possibly to the point of eventually extinguishing it altogether). So care can have standing as a part of oneself, and insofar as it does, it requires acknowledgment, though sometimes the reasons it gives may be defeated by further considerations.

5.1.2 *The standing of cares with others*

Now, if cares merit a certain recognition from the agent whose cares they are, do they also merit a certain kind of recognition from others? I now want to return to the second idea I mentioned above: not only do cares have a standing with the agent whose cares they are, but that the agent has a kind of standing with others insofar as she has cares in the first place. Here I draw on the paper by Jaworska cited earlier called "Caring and

however, a difference between feeling the force of the physical aspect of the addiction and feeling the emotional patterns that tend to accompany it. In fact, one of the most important aspects of breaking a physical addiction is to manage the emotional patterns that feed into it. These emotional patterns are not alien to the addict, in exactly the same way as the carer's emotional patterns are not alien to her, and that is why both need to recognize their emotional situations for what they are before they have any chance of recovering. While the behavior caused by the physical addiction may not "belong" to the addict, behaviors that brought him to become an addict and to continue the addiction do. And this is common to the "recovering" carer as well. The difference I mean to point out here is the difference between these emotional patterns and the behaviors caused by a physical addiction.

⁶ Something like this, with gender roles reversed, is essentially half of the plot of the film *Sliding Doors*.

⁷ Here the reader may recognize a rough sketch of Harriet's plight in Jane Austen's *Emma*.

Internality.”⁸ By insisting on including the marginal cases of the very young and the increasingly demented as instances of genuine, deep caring behavior, Jaworska makes a case that carings, rather than agency, are a fundamental—if not *the* fundamental—part of making a person who she is, and in particular carings play a fundamental role in making a person worthy of respect. Caring is at the root of an agent’s identity, and it is the identity of a person, rather than the agency, that gives the agent standing to be treated respectfully.⁹

Jaworska argues, and I think she is quite right about this, that we tend to think that ignoring or restricting caring behavior inflicts a wrong on the one who cares (all else held equal), even if he is not a “full” agent able to make authoritative claims on his own behalf. Because carings run deeper than cravings, we tend to believe that not to take someone’s caring into account in how we treat him is damaging, stunting, or disrespectful in a way that restricting appetitive behavior is not. This is because restricting appetitive behavior can be in the person’s own interests, despite temporary frustration, but caring behavior is an expression of identity and thus, typically (but not always), an expression of those interests. Caring behavior stems from a much more important source than appetite, and so although it can still be in a person’s interests to restrict it,¹⁰ such restriction must be undertaken with more caution and greater reason.

Jaworska gives examples that I think make a strong case. One in particular concerns the elderly Mrs. Rogoff, who in her prime was a successful businesswoman and popular cook and hostess. Now in her eighties, she has developed dementia and requires a live-in housekeeper, Fran, to take care of her. Occasionally, Mrs. Rogoff will become alert and jealous as Fran prepares dinner, and will insist strongly on taking over the cooking. It seems clear to Fran that Mrs. Rogoff’s self-image as culinary expert suffers when she recognizes that Fran has taken over in the kitchen. At such times, Fran tries to alleviate this effect by assigning Mrs. Rogoff small helpful tasks so that she feels as though she has some control and can make a significant contribution to the meal. One might ask whether Fran is simply

⁸ Jaworska (2007).

⁹ If what I have said above about the distinction between identity and agency senses of “where the person is” is correct, then I suspect Jaworska tends to run these two senses together. It would be an interesting project to sort out the relationship between them, but I will not attempt this here.

¹⁰ The abused woman who still loves her husband, for instance, has a care that works *against* her greater interests. Nevertheless, because caring is an attitude that attaches one’s own welfare to the welfare of the object, *most* of the time, one’s interests are expressed by caring behavior. And even in the case of the abused woman, *some* interest is satisfied by the caring behavior, even though it is not a long-term interest.

acting out of kindness rather than respect, as I have suggested. I will return to this question below.

For now, consider the observation that third parties react censoriously when someone flagrantly violates another's interests by ignoring what she cares about. To take Jaworska's example again: Suppose Mrs. Rogoff has a new caretaker, Jen, who doesn't allow her to participate in kitchen work, construing her desire to do so as "frivolous nonsense." Jen's supervisor or Mrs. Rogoff's family might well censure Jen for such unfeeling behavior. Why? Jen seems not to be reacting properly to Mrs. Rogoff, and not just because she is entrusted with the care of the older woman, but because *anyone* should react kindly to the wishes of someone else *because they are that person's wishes*—if not to indulge them, then at least to take them seriously as important in themselves. There is something inherently weighty about our carings, and it does not have to do with our endorsement of them; for an abused wife doesn't endorse her love for her husband, but as we have seen, that doesn't diminish its importance for her as something she must take seriously as a part of herself.¹¹

Given that Mrs. Rogoff is not a full-blown agent, one lesson to take from the example is that inherent worth does not derive purely from rationality. It derives from an aspect of emotional life instead: care. For most adults, this kind of care will often be coupled with self-reflection and endorsement; this is certainly true of John Smith as I have been telling his story and describing his situation. Smith's deliberation is nothing *but* self-reflection, and that reflection is what results, as I claim, in his self-creating reason to leave Pocahontas and take up the king's commission. It is therefore easy to think that reflection is the important psychic element here, and indeed this is reflected in the literature on agency, especially in the work of Frankfurt, Bratman, and Velleman.

But as Jaworska's examples of marginal agents show, it is important to keep in mind that the care—rather than the reflection—is doing the work.¹² This way of framing the idea represents a departure also from Korsgaard, who bases her ideas about practical identity on

¹¹ As we will shortly see, there is a distinction to be made among the objective value of carings, the personal importance of carings to the person who cares, and the objective value of third parties paying due regard to these carings, because of their personal importance of these carings to the person who cares. My discussion so far has emphasized this third type of value in particular, not only with respect to true third parties, but for a single agent reflecting on her own carings. I will have more to say on the first kind of value below; the second kind was discussed in chapter 4. (Thanks to Elizabeth Anderson for prompting this clarification.)

¹² When it comes to self-creating reasons, however, reflection does play some role as well, since part of having a self-creating reason is adopting an ideal, which requires some reflective structure.

the premise that we are reflective beings and cannot *but* be reflective beings.¹³ This is the premise that Jaworska questions. A child covering his sad and frustrated mother with Band-Aids isn't doing this because he finds that upon reflection he endorses this motive in himself or that he must do this in order to remain who he is. But something in the rest of us recognizes that to stifle this behavior would be to mess with who he is, because he is acting out of his care for his mother. This suggests that it is not the reflective structure that is doing the work here, but that a practical identity *is*.

Jaworska's paper thus represents a turning away from the customary assumption in the literature that reflection is an essential part of the source of reasons. Her point is that caring, which need not be reflective (as evidenced by marginal cases of agency), is also a source of reasons. We take the cares of the very young and the demented to provide us with reasons for action, even though we do not say that they are fully rational agents the way a normal adult is. The reason for this is that hindering, ignoring, squelching, or in other ways not taking seriously the carings of marginal agents seems to be disrespecting them in a most egregious way—or at least, not giving them a kind of engaged recognition, acknowledgement, or standing that they deserve in virtue of their caring.¹⁴ Perhaps this is precisely because we take our carings to be what makes us not just human but *individual* humans most fundamentally; thus carings are most closely tied to our inherent interests. Carings, as we have seen, seem invariably to be where the person “is” (in an identity sense) among the psychic happenings that occur within her. And so it is important to recognize them if we are to appropriately recognize and interact with the agents who care, that is, if we are to respect them properly.

One might think our responses to others' caring could be grounded simply in concern rather than respect. In seeing that others have cares analogous to ours, we perceive the importance of their having their cares satisfied as analogous to ours. A sympathetic or empathetic connection, then, would have us “taking on” that importance on behalf of others. We see it as a good thing that their cares be satisfied. In addition, in seeing an object

¹³ See *Sources of Normativity* (1996, pp. 112-115).

¹⁴ The language of respect here unfortunately suggests a kind of distance which is somewhat opposed to the phenomenon I have in mind. While authority inspires respect by making use of a psychological distance between the agents, I would like to note that caring behavior can simultaneously inspire something that is *like* respect insofar as it is a recognition of standing, but it is a recognition that engages with the carer and tends to reduce psychological distance.

as important to me and thus meriting careful treatment, a concerned person might tend to see the object of my care as having value because of its value to me. I value a raggedy pink rabbit my grandmother made me when I was two years old, and you may see that as reason not to treat the rabbit roughly—because of the importance the rabbit has for me, it is in my interest that it be treated nicely. In being concerned about me, then, you are concerned about what is in my interest, and since it is in my interest to protect the rabbit, you are protective of the rabbit.

In contrast, respect requires that the respecifier leave a space for the respected to pursue her own projects *simply because they are her projects*—i.e., because they are an expression of her identity (whether or not they will benefit her or are in her interests). Note that there is an important contrast between what is in a person’s interest (i.e. for her benefit) and what her interests are (i.e. what she herself cares about). Respect leaves a space for the latter even when they do not clearly have value. Consider the following contrast. We often restrict the behavior of people who are not full agents because it is in their interests that we do so, that is, we do it out of concern for them: we stop children from playing with dangerous objects or in hazardous environments, for example. But there are some behaviors we must think twice about before we restrict them. Even when in a hurry, a parent will often let a child tie her own shoes (which can be a very slow process), not only because it is important for the child to learn this skill, but also because it is clearly important to her to “do it *myself*.” Because carings like this run deeper than cravings, we tend to believe that not to take someone’s caring into account in how we treat her is damaging, stunting, or disrespectful in a way that restricting appetitive behavior is not.¹⁵ Someone has standing in virtue of having identity-defining cares—i.e., in virtue of being a person.

Now, the objector might think that in giving this response I have not quite been fair to the worry.¹⁶ I seem to have painted the question as a matter of *either* respecting carers as beings with a certain standing, *or* being concerned with their welfare construed as a state of affairs. But couldn’t concern be about something other than good states of affairs? Couldn’t we be concerned that someone get the respect they deserve? Certainly; but then it

¹⁵ This is not to say that we shouldn’t try to redirect misdirected cares in ourselves and others; but it does require that our approach to doing so be subject to constraint by respect. This might be expressed in the thought that we are to respect the person, even if we do not respect the project. See Seana Shiffrin (2000) on a closely related point, especially section II.A and pp. 247-8.

¹⁶ Thanks to Sarah Buss for pressing me on this point, and also for suggesting the second of my responses.

seems to me that one of two things is going on. The first is that we are concerned that there be a state of affairs in which the person is respected, in which case my point stands. The second is that we are concerned not to hurt that person in a particular way—not to be unkind. Like respect, kindness seems to be something that obligates *me* because of something about *her*. In what way? The “pain” of having one’s cares ignored is not the same as physical pain. How are we to understand what kind of “hurt” we are inflicting on a person by not taking her cares seriously? It seems to me that this can only be understood by invoking recognition respect, putting her cares in an analogy to our own. If we think our cares matter, we must also think hers do. And so respect seems essential to understanding the reason not to be unkind here.

The point might be further illustrated by considering the contrast between how we would treat Mrs. Rogoff and how we would treat an Alzheimer’s patient in the final stages of the disease, when cognitive function has deteriorated to the point that even cares are no longer present, and the person is again a creature of appetite. Mrs. Rogoff clearly has interests that are stronger and deeper than mere appetites, and any kindness we do her will be intermingled with an attitude of respect for her as a person. But treating the end-stage Alzheimer’s patient kindly will not, it seems, involve this same sort of recognition, for without cares it is no longer clear that there is a *person* there, in the fullest moral sense.

Jaworska’s argument is that even when a person’s caring behavior is futile or (I would add) perhaps, within limits, destructive, we must have good reasons to hinder such behavior.¹⁷ Respect has nothing to do with judgment of those projects’ worth; it gives us a reason to let people count blades of grass if this is what they want.¹⁸ And (as I argue) it requires that people be allowed to commit to lives they come to want after imaginative deliberation, *because* it is the path of life they would commit themselves to (and so, in that sense, the life they care most about having) as leaders of lives.

The source of reasons we seek cannot *just* be the kind of concern just described, therefore, for such concern is contingent on your seeing my caring about something (e.g. the

¹⁷ I should emphasize that such reasons can be available, for respect is one source of reasons among many. Still, sometimes people (I have teenagers specifically in mind here) should be allowed to pursue projects that others can see are futile or against their interests or otherwise of negative worth, simply to accord them the respect that they deserve (and perhaps in hopes that they will learn something). Such cases are limited, of course, for there are projects too dangerous to allow even out of respect.

¹⁸ This argument parallels Darwall’s in (2001).

rabbit) as somehow valuable. Something stronger is required for the point I want to make. Whether or not my care or its object is a good that merits your sympathy, concern, or valuing, the point I want to make is about respect rather than concern. The self-creating reason I am after here is one that stems from something about my own intrinsic worth, not from the state of affairs concerning my feelings about the rabbit and whether they are valuable or not.

To see this, note the role that censure plays in our interactions with one another. There is no denying the fact that others' emotions often move us. People are inclined to react in defense of what they care about, and when such a reaction is directed at me, it *strikes me as though* there are reasons to comply with my critic's wishes. That is, censure directed at me tends to create in me a felt evaluation that I have done something wrong. If that evaluation fits with an established pattern of valuations and evaluative judgments, then I can see it as warranted and thus as giving me reasons. In the case of censure, the target¹⁹ of both my and my censor's emotion is my behavior in this episode; the focus is the worth of the person being defended. Thus, in feeling the emotion (e.g. anger on the part of the defender and guilt on my part), I implicitly assent to the import of both its target and its focus. That is, if in censuring me, the agent whose cares I have violated (or a third party) can get me to feel that my behavior is bad, it will be because I join in the assent to the import of my behavior *in light of* the import of the person and her cares. Since import gives me reasons, I have reasons to take the person and her cares seriously and behave more respectfully toward her. I may ultimately judge that these reasons are not overriding, but in accepting them as reasons I accord the person appropriate recognition. That is to say, I recognize her intrinsic worth.²⁰

For example, suppose I am playing with my rabbit Pete and I decide that my sister's teddy bear Elmer is Pete's friend and they should play together. My sister, however, is playing her own make-believe and starts to cry when I grab Elmer. Our mother comes over, angry at me for snatching my sister's toys. Feeling the force of her anger, I start to cry myself. I give the teddy bear back, not (or not just) because it's good for my sister to have her own toy, but because my mother's anger has taught me a lesson: my sister's toys are hers and I must respect that. Of course, my four-year-old self doesn't learn the lesson quite this

¹⁹ See note 66 in Chapter 4 for an explanation of 'target' and 'focus.'

²⁰ Cf. Darwall's (2006) discussion of reactive attitudes (chapters 4-7).

cognitively, but the point still stands: what I am made to feel has nothing to do with value or concern, but rather, respect.

The normativity of care, then, lies in respect merited because of the identity-conferring status of care. It is identity, which a person has in virtue of caring about things, that gives her a fundamental status by meriting basic recognition respect,²¹ and this status will be contrasted with authoritative and agential status in the coming sections.

5.2 *Identity and authority*

I can now show how undertaking the deliberation outlined in the last chapter is a matter of respect. This will lead into a discussion of a particular kind of identity that a person can have, and it is this identity, the identity as authority over one's own life, that provides the normativity sought in this chapter.

In undertaking the deliberation under discussion, the agent shows respect both for the person he is now and for the authority he has to direct his life. This section's argument is roughly as follows. Someone like Smith has a good reason to try an impasse-breaking deliberation on two counts: (1) because each part of his identity stems from cares, it is respectful to try to adjudicate on the basis of something that can *justify* the choice; and (2) because part of what you are is a chooser with the authority to direct your life, you have a responsibility to care about choosing, and hence you owe it to yourself to choose.

5.2.1 *Why the deliberation is a matter of respect, part 1: substantive identity*

In Chapter 4 I gave an analysis of Smith's dilemma in terms of care: it is cares that bring Smith to his impasse, and it is cares' tie to identity that makes the problem so difficult. We can now add that the difficulty does not reside merely in the agony of having to choose between two "parts" of oneself, but also (and perhaps more pressingly) because each of the identities in question merits respect. Smith is both a lover and an explorer, and (in keeping with what I said above) it is important that he recognize these facts and give due recognition

²¹ It may be asked whether this criterion for meriting recognition respect excludes newborn infants, who have not yet formed attachments that we could call cares, from meriting recognition respect and hence from having moral standing in their own right. My intuitions on this point do not clearly militate against accepting the consequence that infants cannot have *this* kind of moral standing, since such a claim does not carry any implication regarding other kinds of moral worth than newborns certainly do have. Even if this kind of response proves untenable, however, I suspect that arguments for the moral standing of newborn babies could be constructed by extrapolating from proposed solutions to a similar obstacle encountered in metaphysical theories of personal identity. (Thanks to Elizabeth Anderson for bringing this to my attention.)

to each of these aspects of his life. Choosing one over the other seems, at first, not to do this.

Under the circumstances, unfortunately, he must choose. Remember, however, that the recognition respect Smith owes his lover identity—not to mention his lover herself—does not require that he act on the reasons this identity gives him to act. It does, however, demand that he choose on the basis of some consideration that justifies his choice—almost as if the lover in him were another person to whom the choice must be justified, just as Pocahontas herself is.²²

One might wish to argue here that the respect Smith owes to each of his identities gives him a reason to flip a coin—not out of indifference, but out of fairness, given the fact that at the impasse he has no reason to choose one over the other. That is, it may seem as though he has a reason to flip a coin, so as not to show preference to the explorer side of him, *out of respect* for the part of him that cares for Pocahontas. Yet expressing fairness in the sense of treating all parties equally seems like the wrong concern under the circumstances, since under the circumstances not everyone *can* be treated equally. There is no way to be impartial here.²³ But each party can still be treated fairly in the sense of being treated with respect—which in this case means providing a reason that can justify the action even to the party it does not favor. Now, it may be the case that even after an earnest attempt to break the impasse and find a reason, no impasse-breaker appears. Then perhaps the only thing to do is flip a coin. But the point is that someone like Smith must not do it without at least trying to give a better reason first.

By way of analogy, consider a case in which you are the person who doesn't get a promotion because of a coin flip. Do you think the decision procedure was fair? You might, on the one hand; after all, what else could your boss do? And at least it wasn't because of any quality that makes you inferior, you might rationalize to yourself. On the other hand, however, you might still be upset because you know it *could* just as well have been you, and there seems to be something patently unfair about its not being you even

²² In the film, Smith does not as a matter of fact consult Pocahontas. While this is understandable, given that the very fact that he is willing to consider leaving her may be hurtful, it seems as though he clearly owes it to her to deliberate together if their relationship is based on reciprocal love. After all, if she loves him, she must also understand what the royal commission means to him, and thus understand the need for deliberation, if not also endorsing the outcome despite her hurt.

²³ (Moreover, part of the problem is specifically that he must *not* be neutral, because this expresses a kind of indifference that, given the stakes, isn't there.)

though the choice was made in the fairest way possible. Is such a feeling warranted? It might be that you are upset about the *situation's* being “unfair,” rather than that you believe you were treated unfairly. Then it would not be appropriate to feel as though there was something unfair about the way you were treated. But if this is the case, it might be more satisfying to have your boss cite a reason as justifying her choice. In particular, it might make you feel more respected. Alternatively, you might be unable to believe that there really was no difference between the candidates, and anyone who flipped a coin was not fully considering you—so you feel less respected than if your boss gave a reason.²⁴

(Consider the same situation from the point of view of the person who actually was promoted. While she may well feel as though her worthiness for the promotion was recognized, she may yet feel uneasy because it was something arbitrary that put her in her fortunate position. This again might be because it is difficult to believe that there is nothing to differentiate the two candidates, and a coin flip represents a failure of diligent consideration. She, too, might feel more respected if there were a factor that justified her promotion over yours, lest she feel something analogous to survivor's guilt.)

If the boss cites some inclination toward the person chosen, just a feeling that she would be better than you, in conjunction with her authority to make the decision, you have to accept it precisely because of the authority. The content of the reason, based as it is on an “arbitrary” inclination on the boss's part, may not be satisfying; but (assuming you trust the boss to have given it careful and fair consideration, free of prejudice against your membership in some irrelevant category) given the authority, you have no grounds for protest.

It may seem as though there is nothing new being said here; after all, I claimed in the last chapter that cares are what make this problem of Smith's so difficult. But there I implicitly pinned the problem on the emotional difficulty of letting go of something that is quite dear to the agent. Here, however, I've brought out another dimension to the problem: there is a normative element to the cares that demands some gravity when making this decision—in particular, respect for the identities one has.

²⁴ Thanks to Steve Darwall for this alternative.

5.2.2 *Why the deliberation is a matter of respect, part 2: authority*

So far, I have claimed that the normativity of Smith's first-order cares gives him a reason to engage in the deliberation described in Chapter 4. But there is, I think, a further consideration in play here. As I argued in Chapter 3, Smith has good reason to choose, a reason along the lines that Mill outlines in *On Liberty*. In particular, in addition to his being a lover and a sailor, another part of who Smith is is a chooser, one with the authority to direct his life along the path he sees fit. He has a good reason to make the decision before him actively because of the value of leading a full life—but also, and more importantly for current purposes, because he owes himself respect as the director of his life. As I argued in the last chapter, the question is about him in a very particular way, and as a result it engages his self-conception as the leader of his life. This is an identity of his that becomes salient to him as part of the unfortunate situation he is in, and insofar as he cares about being in charge,²⁵ he will respect himself by attempting to find a reason and making the choice.

The point here is that it can matter to an agent specifically *that* she choose, as well as what she chooses. A worry, however, is whether the authority I'm talking about here can plausibly be something we can construe ourselves as caring about in the deep, identity-conferring sense I have been using; hence, we might worry that all this talk of the normativity of identity won't do the work I need it to do when it comes to an agent's identity as authority over her life. Now, it may seem as though it is some kind of failing of an agent, who has authority in virtue of being an individual person with normative standing, not to care about that authority at least implicitly, if not also explicitly. And as we shall see, I agree that it is. So this worry might not seem pressing.

But consider the following. Having an identity as the authority over one's life does not seem to be the same as the substantive identity as a lover or explorer (or mother, teacher, doctor, etc.). After all, directing one's life does not seem like the kind of project that has values one can promote independently; in this it is similar to the project of "pursuing happiness." One pursues happiness precisely by pursuing other projects one values; similarly, one leads one's life by making decisions about what to do in particular situations—by exercising agency, one might say.

²⁵ I will be saying more about this, claiming that someone like Smith will care about being in charge because he has substantive, first-order cares he wishes to defend and preserve.

Of course, at times one also leads one's life by choosing among ideals by which to guide it. We do see that people value "taking charge" of their lives in crucial moments. Most of the time we don't have to take up this kind of identity in an active way, because (like substantive identities) it typically just operates in the background. But when it comes to a direct question about where one's life is going, it becomes more pressing to *take up* this authority as its own kind of identity—that is, to act from within the perspective of the author of one's life, rather than from the perspective of lover, explorer, teacher, doctor, father, etc. This is just the kind of situation that Smith faces.

Still, is leading one's life something one can *care* about, as I described care in the last chapter? Is directing one's life the kind of project that can define an identity? In the last chapter I described a procedure by which the authority part of Smith found something to grab onto—a consideration he could take as breaking the impasse and settling the question of what he wanted, hence what he would do. As I put it there, the authoritative part of him chooses an identity by exercising agency to act on a particular inclination. This suggests that the authority in question here is more closely akin to the agency scope of the will rather than the identity scope, as I made that distinction in the last chapter. And if this is the case, then the normativity of one's authority over one's life is not the same as the normativity of one's identity, as I have just described it. The kind of respect demanded by authority can't be the same as the kind of respect demanded by identity.

Yet they are closely intertwined. To show this, I would like to discuss an argument Stephen Darwall makes in a paper called "Because I Want It."²⁶ Darwall observes that because desires do not seem to give us reasons from within our own practical perspectives (which is a way of saying that desires do not give us reasons on the anti-DBR view),²⁷ once all the rational considerations are out on the table an agent may nevertheless simply want something more than is warranted by the reasons *to* want it. Under the right conditions, he then argues, this want can actually turn out to be reason-giving for another agent. Similarly, as I claim, an agent at a deliberative crossroads may take into account all the reasons on the table, and then decide that "this is who I want to be" in a sense that goes beyond the reasons *to* be that person. And (again, as I claim) this is reason-giving for that agent (as well as

²⁶ Darwall (2001).

²⁷ See my discussion of this in Ch. 2.

others). This may seem “arbitrary,” but it is arbitrariness resulting from aspects of our psychology beyond our control, arbitrariness in a sense which no longer threatens to fracture the self because the agent’s will provides the needed resolution between the open options.²⁸

Darwall’s argument is fairly simple. It begins with the idea that a *claim* an agent makes on another agent is an expression of will that purports to be reason-giving. The reasons the claim purports to give are agent-relative and independent of the value of any state of affairs; and claims purport to merit consideration from the addressee, thereby implicitly recognizing the addressee’s authority and capacity to recognize, consider and act on the reasons they purport to give. This last point is crucial, according to Darwall. Agents are bound to recognize only the authority that they are committed to in making claims themselves, and this authority is precisely what makes it the case that an agent’s will can give another agent a reason. In short, the idea is that if an agent wants her own will to give reasons to others, she must recognize the reciprocal possibility that another person’s will can give *her* reasons. This is easiest to see with an example.

Suppose that both Liz and Ben desire to eat a particular apple, and each forms the end of getting it. From their own perspectives, their desires do not make an appearance among their reasons for wanting the apple; they are instead focused on its juiciness, nutritional content, etc. Suppose that their sympathetic tendencies are overridden for the moment, and under the pull of their respective desires they see the other’s having the apple as having no value. So if Ben addresses Liz and says she should let him have the apple because he wants it, he purports to give her a reason to surrender the apple. Note that this reason is different from a claim that he should have it because his desire for it will make him enjoy it more than she would. The example is constructed so that this is not the case, but even if it were, this greater satisfaction is not what Ben cites in making his claim to Liz. The reason he purports to give stems only from his will to have the apple, not from an appeal to there being more value in his having the apple.

Of course, Ben’s way of valuing the possible states of affairs on the table will strike Liz as incorrect, since her valuation is exactly opposite of his. This very fact, however, might give her pause. Up until now Liz has had no reason to challenge the appearance of value her

²⁸ Here I wish to remind the reader about Rosati’s three-way distinction among kinds of arbitrariness (see discussion in Chapter 3.4). One of the reasons the inclination-based exercise of the will is not arbitrary in the vicious sense is that it is seen as a form of self-expression. Acting on such an inclination thus will not fracture the self, as threatened by vicious arbitrariness.

desire has bestowed on her having the apple. Now that Ben has voiced his desire, however, Liz can see that Ben's desire causes him to see things differently—and Liz can find no explanation of why either her desires or Ben's should furnish the correct appearances, because desires do not have correctness conditions the way, say, beliefs do. That is, no one's desire seems to get something right in a way that the other's doesn't. Therefore, insofar as she credits her own claim to the apple, Liz must credit Ben's as well. In doing so, she recognizes both herself and Ben as what Rawls called "self-originating sources of valid claims." That is, their wills are legitimate sources of reasons independent of any features of the world that give them value. (With this mutual recognition, they may decide that the most appropriate resolution is to divide the apple and share it.)

One lesson to take from this is that agents typically do not notice these will-based reasons from within their own perspectives because they are backgrounded until challenged by an opposing will. Agents are typically focused on the features of the object that give them reasons, and until challenged they do not notice the roles that their own cares, and their authority to direct their own lives, play in giving them reasons.²⁹

Additionally, however, I'd like to bring out two things going on in the apple scenario. First, there is the fully conscious, agential willing on the basis of a desire. This is closely analogous to what I described in the last chapter as the result of Smith's new deliberation, and I will be returning to this topic. Second, there is the demand for respect implicit in the explicit demand for the apple. Darwall locates the legitimacy of this implicit demand in each agent's respect-worthy status as free and rational. But given what I said above about examples like Jaworska's, I believe that something more may be going on here.

Children and mentally handicapped adults are not generally considered to be autonomous agents, in the sense that we do not fully credit their competence to make important decisions. They probably cannot make the specific kind of claims Darwall has in mind in the apple scenario, if the claim explicitly purports to be reason-giving for the addressee. This kind of explicitness requires an amount of self-reflection and consciousness that such marginal agents are not capable of. Nevertheless, they seem capable of making demands that have the kind of authority that requires respect. Consider a child's throwing a temper tantrum at being prevented from doing something she really wants to do, such as tie

²⁹ This suggests that "desire-based" reasons such as those under discussion may be more common than we think. This is an idea to explore in further work.

her shoes: “I want to do it *myself!* Mom-*myyy!*” Isn’t this a demand for recognition? It is not explicitly so; the child is focused on practicing her new skill. But in caring about and thus vehemently defending her tying her own shoes, the child is making an implicit demand for recognition as someone who is in charge of her life. In caring explicitly about her project, the child also cares implicitly about herself and—might we not say?—her authority over her own small sphere of life. She could not formulate such a thought, and especially because she is not yet a full agent, we might override her demand; but it is nevertheless a demand addressed to a particular person (we might certainly feel it that way), and probably one that is made with the kind of authority Darwall is getting at (“You! Recognize me!”), though again, this is not reflective or conscious.³⁰

Now, if the child’s demand gets its normativity from her implicit caring about her authority over her life (and not anything to do with freedom and rationality), why not think that this can be behind the implicit demand for recognition in the explicit demand for the apple as well? I believe that we should. And if we do, then at least part of what is being respected here is the person’s identity as the author of her life, and not just her autonomous agency in virtue of her freedom and rationality.

The upshot of this argument is that one’s identity as the author of one’s life *is* something one can care about in an identity-conferring way, despite its having a different character from substantive identities that arise from caring about projects and people and objects. This kind of care is merely implicit in many of our practices of making claims on one another, as both the apple and the temper tantrum examples show. Furthermore, as John Smith’s case demonstrates, authoring one’s life is something that one can *explicitly* care about as well, and it can in that way become a substitute substantive identity when the very question is which things to care most about, or which substantive identity to have. Therefore, I contend, Smith owes it to himself to make his choice actively because his identity as the author of his life demands respect insofar as he cares about it. The normativity in play here has the same roots as the normativity of substantive identities.

A natural question at this point is whether Smith *must* care about his authority. It is not obvious that he must. After all, there are identities that seem to eschew one’s own authority, such as dedicating oneself to the will of God or one’s husband, or adopting a

³⁰ The case of Mrs. Rogoff can be seen as similar to this one.

classical Stoic or Buddhist lifestyle which eschews care altogether. But we must be careful here. I agree that it is surely possible to fail to care about your authority. And this seems like a failing as a human being. Part of the explanation of this failure, as Mill demonstrates, is that not taking up this authority suggests that the agent is missing out on an important human value. But as I've said before, we cannot dedicate ourselves to all that is of value in the world, and it seems as though there is nothing contradictory about not being compelled to take up this authority. (It is beyond the scope of this project to consider just how compelling this particular value is.)

So there is a failing here that goes beyond a failure to value something which is intrinsically valuable. In having authority over one's life, which an agent has in virtue of being a person with standing, an agent also has a certain responsibility to care about that authority: to take it seriously and act out of respect for it. (Consider the second-personal counterpart to this: because you have authority over your life, I have a responsibility to take you seriously and act respectfully toward you.) This is the basic thought behind various arguments that one cannot sell oneself into slavery.³¹

But this kind of failing is not always what is going on when someone takes up a project that seems to eschew care for one's own authority. I think it is important to recognize that even an identity such as devoting oneself to the will of God or to some great cause such that one would say "my life is not my own" is not necessarily a full abdication of an agent's responsibility to herself. Such a project is need not be the extreme kind of project in which a person abdicates all rights to make decisions for herself. There is still room here for an agent to respect herself, and to reject her dedication to the cause upon further reflection. When she says "my life is not my own," this need not be a literal giving away of authority, but rather an expression of her passion for her project.

Nevertheless, in making the temper tantrum argument, I am committed to the position that in caring about anything substantive, I must also implicitly care about guiding my life by those things, and hence about my own authority over my life (which, it should be noted, is different from my life itself, which I may be willing to sacrifice). The monk and the crusader can certainly fit this bill, so long as they can make intelligible claims on their own

³¹ E.g., arguments made by Rousseau, Kant, and more recently, Thomas Hill (1973).

behalf in defense of these projects.³² It seems, however, that a counterexample can be found in a particular species of deferential wife, discussed by Thomas Hill (1973) and Andrea Westlund (2003), or other characters who seem incapable of justifying their deference, so deeply does it run. The deferential wife apparently does not even implicitly care about her own authority; any reasons she gives when questioned about her deference seem to trace back only to her husband's interests and never to hers. Yet she cares very deeply for her husband, and the project of supporting him in every way is what defines her. But defines her as what, exactly? Certainly not as an autonomous individual who is the author of her life. She seems defined by the very fact that she is not defining herself. With being a person comes a certain sort of authority over one's life, and with that comes a certain amount of responsibility to lead that life, which the deferential wife clearly is not taking up. In this, then, she seems to be a failure, and in fact Hill argues that she is morally culpable for her lack of self-respect. I will add that the deferential wife, as described by Hill and Westlund, seems not to have a certain capability that even the temper-tantrum child has: the ability to take a stand on her own behalf (however inarticulate that stand may be in the child's case). In the deferential wife, her care for her project doesn't seem to translate into care for herself, so thoroughly has she subordinated herself, and this seems at best an unusual misfiring of moral psychology, and at worst a culpable lack of self-respect.³³

5.3 *Identity, authority, and agency*

In this chapter I ultimately want to argue two things: first, that Smith has a respect-based reason to make his choice actively, i.e. to direct his life; and second, that the reason for which he makes it is also grounded in the normativity of respect. The first, as we have seen,

³² These claims need not fully succeed, of course, but they must at least be a legitimate attempt at playing the game of justification. For an illuminating and detailed discussion of this kind of case, see Andrea Westlund's discussion of self-deprecation in her (2003).

³³ I am not ultimately convinced that there can be a character such as the deferential wife, given that one implication of my temper tantrum argument seems to be that substantive caring involves an implicit caring about one's authority over one's life. It is very difficult to imagine someone who has enough conception of herself to be able to function as a separate human being, doing things on another's behalf, but no conception of herself as having authority and autonomy to direct her life *as* an individual self. That is, it is hard to imagine someone who would not eventually respond to pestering by a friend with something like "Just leave me alone! This is just what I want, all right?"—which the wife does not do, as her exasperated response is "What can I say? I just want him to have what he wants; that's all." (The emphasis in my rendering is on "I", but in the Hill/Westlund case, it is on "him.") I say that this is difficult to imagine, but it is enough for my argument that if such a person did exist, she would be going wrong in some way by not taking up the authorship of her life, and so I will not pursue this here.

is grounded in the fact that he lives up to his responsibility to lead his life by caring about doing so, and this care requires respect. I will argue that the second is grounded in the demand for recognition implicit in willing a course of action on the basis of his authority. The contrast is seen in the following two statements: “You’re the one in charge; it falls on you to make this decision” vs. “You’re the one in charge; I’ll respect your decision.” I have now completed the argument for the first point; it is time to turn to the second.

Recall that above I suggested that the respect demanded by identity and the respect demanded by agency were closely intertwined, though distinguishable. The arguments given so far show how they are closely intertwined: the apple example illustrates the point that the authority implicit in agency is ultimately based on the authority that comes from identity (in particular, identity as one who cares about directing one’s life). To see how the normativity of agency works, we must think about how the two kinds of respect just mentioned are distinguishable.

The main features which distinguish agential authority from identity authority are the same features that distinguish autonomous agents from persons who are not autonomous agents. Most particularly, there is a reflective element missing in non- or marginal agents. Full agents are capable of understanding the motives at work within themselves, standing back from these elements, and deciding freely and rationally among the options available.³⁴ Such agents are capable of consciously and explicitly valuing their authority over their lives, and as a result they are able to make the kind of claims Darwall designates, claims that explicitly purport to give another person a reason to act simply because of the claim,³⁵ and that simultaneously implicitly respect the other person’s standing to consider and evaluate the claim.

Because they are capable of reflection, full agents are capable of reviewing and choosing the ideals by which they guide their lives and govern what they care about. This is something that marginal agents cannot do. For example, compare the temper-tantrum child and John Smith. First, Smith has a problem in the first place precisely because he is capable

³⁴ As I understand it, this is the concept of agency used by such philosophers as Darwall, Scanlon, Korsgaard, and Frankfurt, among others. Furthermore, it is, presumably, the concept of agency associated with judgments of competence in the legal and medical realms. (See, for instance, Beauchamp and Childress’ *Principles of Biomedical Ethics* (2001).)

³⁵ That is: in the full knowledge that the claim is meant to give the addressee a reason. This is to be contrasted with cases like the temper tantrum, in which there is a claim being made, but it is not made with the direct, conscious intention of giving a reason.

of reflection on what he cares about. Second, as a result, he has a way of framing his problem that the child does not have.³⁶ For Smith, it is not merely a matter of deciding between the objects of conflicting desires, as it would be if he were incapable of reflection, but how to understand, compare, and articulate the underlying values of his options.

Furthermore, Smith's substantive cares aren't functioning in the normal way right now, precisely because they are what is at issue. He cannot make a choice from within the perspective of a substantive care, as the child does. Thus, Smith's caring about leading a life must be explicit in a way a small child's could not be.

An implication of this is that not everyone can have a self-creating reason. But I think this accords with widespread intuitions about competence and autonomy. A child, for instance, cannot make a serious decision about who she wants to be, or what ideals are to guide her life. And since she cannot, she also cannot make who she wants to be into a reason for action for herself. For this reason, we typically do not think a child has *full* authority to lead her life in the same way mature people, with their reflective capabilities, have such authority.

Smith is a full agent who can choose from the perspective of authoring his life. What he chooses will be a matter of agency and reflection—willing on the basis of something that goes beyond the reasons there are *to* will it, as in Darwall's example. Darwall's story focuses on how one agent's will can give another agent a reason for action. The question that concerns us, however, is how an agent's will can give *the agent* a reason for action. There are no direct *claims* being made, in Darwall's technical sense, but there is an expression of will that purports to be justifying. Consider the fact that Smith could (and should!) address a claim to Pocahontas with the same content as the reason he gives himself.³⁷ If his will is capable of giving her a reason, why shouldn't it also give him a

³⁶ We might compare Smith and the child to Charles Taylor's strong and weak evaluators (see Taylor 1985a, 1985b).

³⁷ To see this, consider how the situation looks from Pocahontas' perspective; she has a stake in the matter too. She will, of course, want Smith to remain with her, and will reasonably do her best to persuade him to stay. But if, after deliberating both together with Pocahontas and on his own, he can make clear to her that accepting the commission is in the end most important to him, she may feel any number of warranted and unwarranted emotions, but ultimately if she is able to think clearly about it she will see that tying him down would be a terrible betrayal, and anger at him *for making the decision he makes* would not be warranted. If he does not consult her at all, of course, she would be justified in being angry for that further reason.

reason?³⁸ He may have no particular grounds for thinking that the inclination on which he bases his willing gets things right, precisely in situations like the present one where reasons don't settle the matter. But given that there is no question of getting it right, there is also no question of getting something *wrong*, either, and hence no reason not to will that he act on the inclination.³⁹

And so if John Smith wills, on the basis of his authority to lead his life as he wishes to, that he trust his inclination to accept his commission as representing *him*, he makes that inclination into a reason for himself, much the way Liz makes her desire for the apple into a reason for Ben when she makes the claim on him that she have it. In neither Smith's nor Liz's case is the desire itself normative until an authority with the status to make claims gets behind it. Notice, however, that in Smith's case it is not just the willing that does the job; it is also partly the fact that he is free to see the inclination as being his in a deeply important way. Seeing it in this way, he can see himself as willing not on the basis of just any desire, but one that articulates what is of deepest importance to him, a substantive identity that is worth respecting. This last point bears further discussion, and I turn to it in the next section.

5.4 *Constraints on authority*

I have shown that what Smith wills is authoritative, so that he has a reason to do as he wills under the circumstances. But there is a further dimension here that bears exploration because of the case I made in the last chapter that Smith is able to see the inclination as expressing *him* in an important way. This way could not be a full-fledged identity, for that could only be based on a care, and this is as yet a mere inclination. Yet the inclination is something more than a "mere" inclination, such as (perhaps) Liz's desire for the apple was.

Recall from above that one aspect of Smith's reason to undertake the further deliberation was the fact that each of his cares required a certain amount of respect. I made an analogy there to a boss deciding between two equally qualified candidates for a job, and suggested that although they will abide by the decision simply in virtue of the boss's

³⁸ (This claim is to be understood within the context, according to which he does not have reason to struggle against what he really wants, since both of the options he faces are perfectly permissible, rationally and morally speaking.)

³⁹ This way of putting it is not entirely accurate, though it suffices for the present point; below I will discuss the sense in which one's desire might get things "wrong."

authority to choose, each of them would feel more respected if they could view the boss as having chosen on the basis of some consideration, rather than arbitrarily. With Smith we have a similar situation: although his authority to choose lends normativity to whatever choice he makes, he will be able to respect himself more if he can see himself as choosing on some basis. This is why he needed something that could break the impasse.

A point I would like to bring out here is that although the normativity of his choice issues from the authority he has to make the decision, that authority must work within some constraints. Smith cannot choose on the basis of any old desire. The inclination he sees as breaking the impasse could not, for instance, be the desire to get out of the problem entirely by taking some third path, like joining a band of pirates, or doing something that deliberately causes Pocahontas to hate him, or committing suicide. It must, instead, be one that he can credit as resulting from the earnest attempt to articulate a self-interpretation that is true to who he already is, even as it reinterprets and shapes that identity. Respect for who he is—for his substantive identities as both a lover and a leader—demands that much.

In the last chapter, I argued that “this is who I want to be” has the feature of being true to something important to him that simply is not available from the discovery-model deliberation undertaken from the anti-DBR perspective, where “what to do” was not a question of what he wanted, but what he had reason *to* want. Thus, a self-creating reason really has two features: first, it is created by authority; second, however, that authority is constrained by the necessity to be true to the self it is creating. Many possible ways of making the choice will be articulations of value; that is why there is a problem in the first place. The question is whether they are articulations of *personal* value appropriate to the person deliberating. The freedom we have in the choice of ideals is not, then, a full, radical kind of freedom. The inclination Smith follows is not, therefore, any old inclination.

Talk of appropriateness suggests a challenge. Can we be wrong about the ideals we adopt? On a first pass, I can think of three different senses in which the answer to this question might be positive. The first is that the agent cannot manage to effect the reorientation of her life that she commits to (perhaps because her will is too weak and she gives in too easily to temptation; perhaps some addictions are like this). The second is that an agent chooses a bad ideal for her, i.e. one that is not sustainable because it requires behavior contradictory to that required by other ideals she will not give up (as in, perhaps, a woman who tries to be both a mother and a powerful executive, but does not really manage

to be a good mother because she delegates all her familial duties to her husband or nanny). The third is that an agent chooses an “ideal” representing something immoral or inadmirable along some other dimension (e.g., becoming a terrific scam artist or the lazy “dude” in *The Big Lebowski*⁴⁰). Though, as my brief examples show, these ways of failing need not overlap, it is also possible that they do. For example, an agent may try to adopt an ideal that does not fit her and be unable to effect the strong adoption because of an inconsistency with something she is unwilling or unable to give up. Moreover, the inconsistency may be between moral commitments and aspects of her project.

Each of these ways of going wrong does seem to be the result of a failure of some sort: to introspect clearly, to consider all relevant reasons fully and vividly, to reason consistently, or to bring emotions into line with judgments. Although these are important kinds of failures, I take it that these do not get at the thrust of the objection. The spirit of the challenge here is something more like this: Can one be wrong about the kind of person one wants to be, *even* when one is able to strongly adopt an ideal that is perfectly worthy in the sense that it is attached to a morally permissible project of genuine value?—that is, can one be wrong even when there is no failure of rationality or self-control?

I admit that it is difficult for me to make sense of the question here. What would it mean to be wrong about this, if one is in fact successful? Could an impartial observer blame me (morally, rationally, or otherwise) for choosing the ideal I did? I am not sure on what grounds. Could I be wrong about what I really want? It seems possible to strive to live by an ideal and, once successful, find that it is hollow and unsatisfying. But it seems to me that this kind of case falls under one of the failings I mentioned first: a failure of introspection or reflection on the actual value of the project (either intrinsic or personal). Even leaving this aside, however, it seems to me that if I am successful in reorienting my life, the question of whether I was correct would not arise in any serious way, for comparisons between opportunities not taken are speculative at best, and will always be colored by the perspective of the opportunity actually taken.

Perhaps regret is a sign that I got something wrong. Yet regret is somewhat tricky. We can regret situations that force us to choose among incompatible options, which may or may not be the same thing as regretting the incompatibility of the options. This kind of regret is no doubt possible for the person who has successfully effected a reorientation of

⁴⁰ Thanks to Matthew Slater for this example.

her life toward a worthy project. Anyone may regret circumstances over which he has no control. But this is not, I suspect, the kind of regret that answers the question as to whether one can be wrong about what one really wants. The necessary kind of regret has been called “agent regret” in the literature, and it is a regret concerned with oneself or one’s choices, things over which one has some control.

Monika Betzler gives an attractive account of agent regret in her “Sources of Practical Conflicts and Reasons for Regret.”⁴¹ Betzler’s argument is that the particular kind of agent regret she wants to isolate is a response to reasons not acted on. It is the result of feeling lingering reasons *from* commitments,⁴² the reasons *for* which have faded or been outweighed. That is, Betzler distinguishes two categories of reasons: first, there are reasons to have certain commitments; and second, there are the reasons that stem from these commitments, reasons we have only if we have the commitment. Betzler thinks that the second kind of reasons can take on a life of their own and continue to have a partial grip on us once we have come to recognize them, even if the commitment which brought us to recognize them has diminished in importance. The idea is that regret responds to reasons that are no longer in force, reasons which represent commitments not (or no longer) acted on.

Regret of this sort may well be possible for someone like John Smith, even if he is successful in acting on his choice and reorienting his life. Yet this does not necessarily mean that his decision was mistaken; it is only a sign that his attachments were deep and still retain some value to him. That is, regret is the inevitable predicament of a person who must make decisions and close off opportunities.

The only sense I can make of being wrong about a self-creating reason, therefore, is the sense in which interpretations and articulations of who we are can be more and less adequate, more and less rich. This is difficult to understand clearly because self-interpretation is a give-and-take process; interpretation shapes object and object shapes interpretation. This idea appears in Charles Taylor’s “What is Human Agency?” when he discusses self-interpretation and experience, for example. On the one hand, according to Taylor, reflection “can sometimes win through to a new way of seeing our predicament,” but

⁴¹ Betzler (2004).

⁴² “Ideals” would fit just as well as “commitments” here. Recall the relationship between commitment and adopting ideals in Chapter 4.

on the other hand, “certain descriptions of experience are unacceptable or incomprehensible to some people because of the nature of their experience.”⁴³ One could be wrong about who one wants to be in the sense that, try as one might, one cannot make one’s experiences fit the ideal one is attempting to live by. This also fits with Bennett Helm’s argument, discussed in the last chapter, that evaluative judgments and felt evaluations are sides of the same coin, shaping one another as the agent goes about living his life and interpreting his experiences. But again, this seems to be one of the first two kinds of failure enumerated above.

So it is not clear to me that one *can* be wrong about who one wants to be, as I have here understood this challenge; it is clear, however, that one may need to refine the self-conception one is trying to articulate in acting on a self-creating reason. The person one is, then, normatively constrains the person one may become.

5.5 *Conclusion*

The normativity of ideals is complex and tricky because ideals involve several kinds of value and normativity at once. First, insofar as ideals do connect to agent-neutral intrinsic value, they involve an amount of normativity that is independent of an agent. Being a good pastry chef, for instance, involves certain things that the agent has no say over. Yet there is nothing that agent-neutrally requires her to adopt the ideal of a good pastry chef. We have freedom to adopt some ideals to the exclusion of others because there is more that is of worth than we can ever dedicate ourselves to in an active and meaningful sense. It is only once an agent has adopted an ideal that this kind of normativity really governs her life—or perhaps the better way to put it is that an agent governs her life by the normativity of an ideal only once she has adopted it.

Because of this, however, ideals involve a second level of value: an agent-relative level. Barbara has more reason to become a good pastry chef than she has to become a breeder of orchids because of her particular talents in the culinary realm, or perhaps simply because she likes cooking better than horticulture and would be happier as a pastry chef. Nevertheless, Barbara *need* not become a pastry chef. No doubt her talents and dispositions suit her to plenty of other occupations as well. So the agent-relative value of adopting a

⁴³ Taylor (1985b, p. 37).

particular ideal also does not require anything very specific of the agent, because there may be any number of suitable interpretations or articulations of her life. But the talents and dispositions at the core of these interpretations do serve to constrain what she may reasonably make of her life. The choice of ideals, then, involves a good deal of freedom, but it is not unconstrained freedom.

The argument of this chapter has been intended to show that despite the quite broad freedom we may have to choose our ideals, the choice need not be radically arbitrary. I have argued, in particular, that Smith has a non-arbitrary reason to adopt the ideal of explorer, even though this non-arbitrary reason was also not given by any desire-independent reasons that he had. Instead, it arose as an inclination which was the outcome of an attempt to articulate his inchoate sense of himself. Even assuming that an inclination is not in itself normative, as the anti-DBR view holds, I have argued that Smith has the authority to make such an inclination normative by adopting an ideal on its basis. Furthermore, I have been making the case that the normativity involved here derives from the particular kind of self-respect demonstrated in an agent's taking seriously the authority he has to direct his life.

In the language of self-creating reasons, the conclusion is as follows. Because respect is a (desire-independent) normative notion, the normativity of a self-creating reason derives from the respect an agent accords himself in choosing on the basis of his authority to lead his life, i.e. to make his life what he wants to make it. Choosing on this basis requires that the agent take certain inclinations—the ones that arise in the attempt to articulate his vague sense of what is important to him, in response to the kind of deliberative impasse discussed here—to express who he is, and that he govern himself by ideals to which they point.

And so a self-creating reason is both identity-based and identity-shaping: it is based on the agent's identifying himself as an authority over his life, and it shapes the substantive identity he comes to adopt. It is important to be clear here: a self-creating reason is thus a reason whose content is an inclination (which often takes the form of "this is who I want to be"), and whose normativity derives from the respect I show my authority by governing myself by the ideal I adopt on the basis of the inclination. I thus create a reason for myself, out of elements that constitute my identity; and this reason is self-creating in the further sense that I shape myself according to the ideal it leads me to adopt.

Chapter 6: Conclusion

If the arguments of the preceding chapters are right, I have shown that the detached, discovery-model picture of deliberation presented by the anti-DBR view is incomplete, and that some deliberation is rightly about the agent in a deep way. As a result, not all deliberation involves an agent holding herself to some external standard of behavior, such as that given by basic models of consequentialism, deontology, or even, perhaps, virtue ethics. In this concluding chapter, I would like to gesture at some of the broader implications of this view, sketching ideas that may be developed in further work.

6.1 *What about the DBR view?*

Throughout this project I have taken the anti-DBR position as a fixed background, defining the problem of choosing ideals starkly against this view. This is all well and good for someone in the anti-DBR camp, one might say, but what's the use of worrying about all of these complications concerning self-respect if you thought desires could be reasons all along? If I am right, what I've done is reinterpret the normativity of a particular kind of desire, locating it not in the mere wanting of something, but in a more substantial aspect of our lives.

Having seen how this argument runs with the anti-DBR thesis as a target, we can see that the normative structure I've suggested is plausible even for those who already thought desires could provide reasons. I have argued that the kinds of desires that have ties to a person's identity provide a particular sort of reason that is not captured by the simple claim that all desires provide reasons insofar as they are desires. Some desires have a further normative structure even if you accept that all desires are normative.

To reiterate: the ties to respect and authority are what make a self-creating reason a reason. Ordinary desires do "belong" to the agent in the sense that they need not be alien forces that undermine her agency, but they do not necessarily merit the acknowledgment and

the respect-worthy status that cares do, because not all desires have the ties to identity that merit respect. The inclination that forms the basis of the self-creating reason, by contrast, has a special status because the agent's exercise of her authority to interpret the inclination as tied to identity merits respect. This does justice to intuitions on both sides of the DBR question, satisfying intuitions that desires can have some normative significance while also satisfying intuitions concerning the reason-based structure of deliberation when it criticizes desires.

Here one may worry that the line between what counts as an ordinary or "mere" desire and what counts as a genuine expression of identity may not be fully crisp, but I do not think this is a threat to my defense of self-creating reasons because it reflects a real ambiguity in our psychological and volitional life. In some situations, one may not be sure that she has assessed the situation correctly. Part of the value of my project may thus lie in demonstrating the value self-creating reasons have for helping out the epistemologically limited deliberators we are. In real life we never really know that the reasons balance out; there is only so far regular deliberation will take us. By arguing under the assumption that will-independent reasons are tied, I have shown that will-*dependent*, self-creating reasons have normative authority, with the implication that they can be legitimate even without the certainty that regular reasons have run out.

What I have argued here does not, then, fully resolve what is at issue between the DBR and anti-DBR camps. It has allowed us to see the pressures in both directions, however. Because self-creating reasons are neither purely desire-based nor purely desire-independent, I suggest, the issue loses some of its original bite in cases like Smith's.¹

6.2 *Are self-creating reasons merely tiebreakers?*

Because of the path I have taken to build my case for the existence of self-creating reasons, one might think that they can serve only as tiebreakers, when desire-independent reasons have run out. This suggests a peculiar category of reasons that are reasons sometimes but not always, which would pose a problem for self-creating reasons.

On the surface, it seems as though situations of self-creation arise only rarely. As the case of John Smith illustrates, they require a serious question of taking up the authority one

¹ Thanks to Louis Loeb for this point.

has to direct one's life, and they require a set of genuinely rationally permitted options. In fact, if we recall Raz's "basic belief"²—to the effect that the *usual* situation of an agent is one in which multiple options are eligible—I venture to speculate that unless much of our everyday action is arbitrary, we need reasons to decide among several open alternatives. And if, further, we recall that self-creating reasons appear when the volitional boundaries formed by care become unfixed, I venture to speculate that one implication of the basic belief may be that such volitional uncertainties may be more common than we would ordinarily think. Self-creating reasons are a natural candidate (though not necessarily the only one) for filling in the gaps left by desire-independent reasons; there is no reason to think that self-creation is not an ongoing and everyday kind of process. We must author our lives every day.

It would take more argument than I have space for here to construct a detailed defense of this speculation, but here are a few thoughts. The John Smith example is a dramatic case, but given the nature of care and its gradual ebb and flow, many "smaller" decisions we make will have a tendency to shape our identities, especially as they contribute to forming habits. Examples of this include such everyday acts as the activities that maintain friendships and family relationships, the development of regular work and leisure patterns, decisions that affect the way one balances work and family life, patterns of diet and exercise, choices of hobby, etc.³ The things an agent cares about are not under a dramatic threat in any given instance, but they may be under the small threat that they will slip and fade if they are not maintained. To the extent to which this is so—and I am not claiming that it is a very great extent—self-creating reasons may well enter deliberation: I may understand a given action or desire as expressing who I want to be, and act at least in part for that reason, thus actively directing my life. Given that impersonal, desire-independent reasons frequently underdetermine what is to be done, it seems as though we *need* a second sort of "personal" reason that can help us navigate non-arbitrarily among a vast number of intrinsically worthy options.

In any case, because of the tie to self-respect I do not think self-creating reasons are legitimate considerations only when all desire-independent reasons result in an impasse. In fact, I believe there are cases in which desire-independent reasons clearly or even

² Discussed in Chapter 2.

³ Thanks to Berislav Marusic for prompting me to think about this point.

overwhelmingly favor action, but it is a self-creating reason which actually prompts that action—and this is entirely appropriate, normatively speaking. That is, it can be appropriate to act on a self-creating reason even when desire-independent reasons also decide the matter.

In this vein, I would like to revisit Jean Valjean’s problem, which I mentioned in my introduction. Here I will use the Broadway musical version of *Les Misérables*⁴ because it illustrates the problem quite conveniently. Valjean learns that Inspector Javert has apprehended a man whom he believes to be Valjean. Recall that the real Valjean broke a harsh and largely unjust parole and has been in hiding for years now, living under the assumed name of M. Madeleine. Beginning with the silver given to him by a kind bishop in a life-changing act of forgiveness, “Madeleine” has become a famously morally outstanding citizen, built a prosperous factory which provides work for many otherwise unemployed villagers, and become the mayor of his town. Upon hearing Javert’s triumphant news, Valjean realizes the terrible choice he faces: shall he protect his prosperity and the livelihood it provides for people less fortunate, or should he give up everything he has worked for, turn himself in, and be returned to the galleys for the sake of preventing the injustice of condemning an innocent man? He sings (forgive the long quote):

*He thinks that man is me
He knew him at a glance!
That stranger he has found
This man could be my chance!
Why should I save his hide?
Why should I right this wrong
When I have come so far
And struggled for so long?*

*If I speak, I am condemned.
If I stay silent, I am damned!*

*I am the master of hundreds of workers.
They all look to me.
Can I abandon them?
How would they live
If I am not free?*

*If I speak, I am condemned.
If I stay silent, I am damned!*

*Who am I?
Can I condemn this man to slavery?*

⁴ Boubil and Schönberg (1980).

*Pretend I do not see his agony?
This innocent who bears my face
Who goes to judgment in my place—*

*Who am I?
Can I conceal myself forevermore?
Pretend I'm not the man I was before?
And must my name until I die
Be no more than an alibi?
Must I lie?*

*How can I ever face my fellow men?
How can I ever face myself again?
My soul belongs to God, I know
I made that bargain long ago
He gave me hope when hope was gone
He gave me strength to journey on.*

*Who am I? Who am I?
I'm Jean Valjean!*

In his question, Valjean sees his problem as a matter of who he is, not only in the sense of the outward identity by which the world knows him, but also in the sense of his character. An act of mercy long ago put him on the path to his present powerful and beneficent position, and on the one hand he is drawn to protect everything that has grown from that one act—not just his own prosperity, but that of his workers and townspeople. Yet on the other hand, the very spirit of that merciful act, a spirit in which Valjean has since directed his life, commands him to reveal himself, to redeem the innocent victim of Javert's unforgiving zealotry. He struggles, but decides that although the loss of his factory will be bad, he will not be someone who can let an innocent man be damned to the brutalizing fate Valjean himself knows all too well. He makes his choice in the second-to-last stanza, knowing it is the right thing to do. He cannot face himself if he lets Javert prosecute his victim, and so Valjean chooses to own himself as the man he once was, even knowing that everything Javert has known him to be since then will not cancel the requirement of “justice” imposed by the self-righteous constable.

Given his opposed commitments, either choice Valjean makes would leave its mark on his life; either choice would shape him—deeply, not just by external circumstances—in a new way, certainly from his own internal perspective (if not also from a third person's). Either way he decides, he will be affirming one responsibility and betraying another. It seems that, morally speaking, the choice is clear. Yet Valjean chooses to surrender not (at

least not entirely) on the basis of its being the morally right thing to do, but on the basis of who he will be—to himself, if no one else.

The salient feature here is that Valjean is not making a judgment on the basis of something external to him, though external reasons clearly decide the matter. The problem for Valjean is not to figure out what his moral duty is, as it was for Marguerite Blakeney, but to steel himself to do it. By understanding it as a question about himself and not just his deeds, he gets more traction on this; the self-reflecting frame brings into stark relief exactly what is at stake in his choice: whether he will be morally strong and own his responsibility rather than be weak and shun it. Moreover, the scope of this decision is not “just this once,” and Valjean recognizes this; he clearly sees himself as choosing a path that defines his life, just as his pledge to the bishop did long ago. This very pledge is now called into question, as part of Valjean’s question to himself is whether to continue to honor it and embrace morality as one of his ideals, at the expense not just of his material prosperity, unfortunately, but also something quite morally significant: his beneficence toward his workers.

If this is right, we should understand the boundaries of his will to have become unfixed in the same way John Smith’s have, making the dilemma a problem that is about Valjean even though it is also a moral problem (“even though,” because moral problems are typically *not* about the agent in this way). “Who am I?” is thus a perfectly legitimate frame for the dilemma. And this lends a normative legitimacy to his deciding on the basis of a self-creating reason and not entirely on the merits of the case.

So I believe that self-creation can be present even when self-creating reasons are not needed as tiebreakers. I would like to emphasize, however, that there is a difference between a self-creating reason and certain other cases where self-respect makes an important contribution to the considerations on the table. It is certainly true that some of the considerations of self-respect that give rise to self-creating reasons are present in many important deliberations (take the Nora Helmer case discussed in Chapter 4); we often do not need to notice or rely on them because desire-independent reasons often clearly decide the matter for us, or because it is clear which of the alternatives is more important and there is no need to invent the path that is right for us. For the latter instance, Smith’s twin James, whose cares do settle the question for him when the original deliberation did not, can also see himself as leading his life according to the values he has chosen; his following his cares is a matter of self-respect because his identity as an adventurer is worth respecting. Such is

also the case in situations like Isabel Archer's where one acts to preserve one's integrity. But in these cases the question of whether the agent will actively lead his life is not necessarily a seriously urgent question, as it is when self-creating reasons are involved; in cases like James', responding to reasons in the usual way *just is* leading his life actively. Both self-creating cases and ordinary self-respect cases involve respect for identity and authority, but only in self-creating cases is there a specific respect for the agent's identity as an authority over his life when his ideals have come into question.

6.3 *Self-creation and immoral projects*

Jean Valjean's problem as I've just interpreted it may prompt the further question whether we must respect cares—and authoritative decisions to act on them—even when they are for bad or immoral things. What if Valjean had instead chosen to let Chapmathieu (the man accused by Javert) be condemned to the galleys? What about the sincerely committed Nazi or sadist?

Here I contend that the identity-expressing care and the self-creating reason it gives rise to do constitute reasons, but like any reason, they are outweighed, overridden, or otherwise defeated when they are for immoral things. Moral considerations, in particular, very often⁵ defeat other reasons. An agent who acts on a self-creating reason which is clearly defeated by other considerations is doing something incorrectly—perhaps failing to deliberate correctly, e.g. failing to see that this particular question is not properly about her and her life, just as some questions are not properly about what is prudent or what is most enjoyable. The fact that some questions are not about prudence or enjoyment does not cancel any reason-giving power provided by these considerations; the fact that some questions are not about the deliberating agent similarly does not cancel the reason-giving power provided by self-creating reasons. But they may well be defeated, as pleasure is defeated when it is pleasure at someone else's pain. And so, had Valjean chosen instead to remain in his life as factory owner and mayor, we might well sympathize with the choice while nevertheless judging it as morally, and hence overall, incorrect.

⁵ If not always, as is commonly (though not universally) believed.

6.4 *Final thoughts*

Over the course of this thesis I have attempted to pinpoint the factors that make some practical dilemmas so important to an agent that they become self-defining. This question is the precursor to a deeper one: how does, or should, an agent finally tip the balance of reasons in favor of some alternative? I have argued that sometimes, there is something more to say about what can take place, deliberately speaking, after reasons seem to give out.

To bring the project to a close, I would like to visit briefly a thesis of Thomas Nagel's in "The Fragmentation of Value." Nagel is a value pluralist. He believes that there are fundamentally different types of value, stemming from different (incomparable) sources, and thus most appreciated from various different points of view which any human can adopt. He contends that most practical conflicts arise when two of the fundamental types of value speak in favor of competing alternatives. Where does this leave us agents? "When faced with conflicting and incommensurable claims we still have to do something.... And the fact that action must be unitary seems to imply that unless justification is also unitary, nothing can be either right or wrong and all decisions under conflict are arbitrary."⁶ But of course we should not accept such a conclusion, if for no other reason than our own experience; it does not *feel* as though all of our decisions in cases of conflict are just arbitrary.

Nagel takes note of two pitfalls he wishes to avoid. I agree with him that we should shun both "romantic defeatism" and "exclusionary overrationalization."⁷ The former gives up altogether on any sort of systematic understanding because any such understanding leaves many problems unsolved. The latter deliberately ignores any considerations that cannot, as Nagel puts it, "be brought within the scope of a general system admitting explicitly defensible conclusions." The notion of self-creating reasons I have presented here avoids both of these problems, and although it goes further than Nagel's own proposal, it is in a similar spirit. Nagel's own proposal is, briefly, "that there can be good judgment without total justification, either explicit or implicit." What he means by this is worth quoting at length:

The fact that one cannot say why a certain decision is the correct one, given a particular balance of conflicting reasons, does not mean that the claim to correctness is meaningless. Provided one has taken the process of practical justification as far as it will go in the course of arriving at the conflict, one may be

⁶ Nagel (1979/1991), p. 134.

⁷ Nagel (1979/1991), p. 137.

able to proceed without further [rational] justification, but without irrationality either. What makes this possible is *judgment*—essentially the faculty Aristotle described as practical wisdom, which reveals itself over time in individual decisions rather than in the enunciation of general principles. It will not always yield a solution: there are true practical dilemmas that have no solution, and there are also conflicts so complex that judgment cannot operate confidently. But in many cases it can be relied on to take up the slack that remains beyond the limits of explicit rational argument.⁸

He goes on to note that

the position does not imply that we should abandon the search for more and better reasons and more critical insight in the domain of practical decision. It is just that our capacity to resolve conflicts in particular cases may extend beyond our capacity to enunciate general principles that explain those resolutions.

Bracketing some of the details, as I said, I believe Nagel's basic idea is right.⁹ We can take the process of rational justification—in the discovery model's sense—as far as it will go and then keep going without becoming irrational. This is a surprising claim: even after all desire-independent reasons are accounted for, a decision that breaks a tie or tips a balance need not be arbitrary in a threatening sense. There is still room for *choice* rather than picking.

An adequate understanding of Nagel's faculty of judgment would need to explain how this could be so. Yet as he presents it, Nagel leaves much to be desired; I suspect that he is much closer to romantic defeatism than he thinks. As Michael Weber writes:

Some maintain that without an all-encompassing perspective there is no choice but to resort to picking: There is nothing to do but simply pick one perspective and act according to its dictates, according to what it regards as best. The choice between perspectives—and hence choice in general—is thus fundamentally nonrational. According to this picking model, choice can never fully take into account all the perspectives and hence all the values at stake. This seems to me a dire conclusion.¹⁰

In this much Weber and Nagel are in agreement. But Weber rejects Nagel's proposal that the inarticulate faculty of judgment can make these calls in a rational but sub-understood way;

it seems to me that giving up on justification altogether is just as dire as giving up on somehow taking into account the variety of perspectives, leaving it merely a matter of picking between them. A theory which claims that our choices cannot be

⁸ Nagel (1979/1991), p. 134-5.

⁹ Of course, I differ on some of the details. First, I don't think the faculty is fully inarticulate. Second, my view is not fully in the spirit of Aristotle's, as Nagel's is; Aristotle's faculty of practical wisdom differs, as I understand it, from the quasi-Kantian thesis I offer here, for the volitional and authoritative talk I employ would be alien to Aristotle.

¹⁰ Weber (2004), p. 86.

defended and rational choice is just a matter of ‘seeing it right’ is deeply unsatisfying—and deeply subject to abuse.¹¹

I agree with him about this much. And I agree with Weber’s underlying idea that the art of decision making is an art of navigation and compromise among perspectives, both temporal and valuational. But his suggestion (discussed in Chapter 2) that satisficing is rational because it is rational to compromise among perspectives amounts to the claim that it is rational because we have no other option. I think there is more to say. A choice of perspectives is rational because of the way we create the reasons to take it up, thus creating our selves—i.e. because of self-creating reasons.

To put the point another way: first-person experience with dilemmas such as those under discussion suggests there is often more to say about the process of judgment than Nagel suggests. My argument has been that purely *rational* justification involves making use of a certain desire-independent understanding of reasons, and that once these reasons have given out, conflicts of commitments to ideals land us in the realm of volitional, self-creating reasons. “Judgment,” then, consists partly in a process of bringing rational and volitional selves into harmony.

In light of all I have said, I would like to propose an alternative metaphor for the process of deliberation: not weighing, but navigation. A journey involves a great deal of freedom to choose your path, but you must still be responsive to the features of the landscape: the rocky shores, the islands, the mountains, the rivers and lakes. Instead of discovering a single correct decision by looking at scales, we agents create our journeys, and there may be different “correct” routes through the terrain, which can change over time (both by engineering and by natural processes).

We are creatures who look for reasons, there is no doubt about that. But in this thesis I move away from the anti-DBR tendency to think that all reason-seeking involves detachment and a scientific kind of objectivity. By arguing that agents who actively lead their lives are tangled up in the decisions they make and the deliberating they do, I have tried to give voice to a deep sense that agents are not *just* practical reasoners, but *livers of lives*. A narrow focus on practical reason as the salient feature of agency tends to efface the

¹¹ Weber (2004), p. 86.

individual, but it is important to remember that the classic question of ethics is “How should one live?”—a question that was for a long time seen as asking how a person should *be*, and not just what he should do.

Nevertheless, we *are* creatures who look for reasons. What I have done here is try to explicate how some choices agents face when reasons seem to have run out can be defended: an agent may appeal to the fact that her decision makes her into the person she wants to be.

Appendix: Further Examples

1. Eric O'Neill and Robert Hanssen, partly as told in the movie "Breach," but partly as heard from O'Neill in NPR interviews on "Weekend Edition" and "Fresh Air."¹

Eric O'Neill was an FBI operative in the Special Surveillance Group as an investigative specialist (a "ghost"). He was tasked to work as an assistant to Robert Hanssen, whom the FBI suspected of trading secrets to the Soviet Union and Russia over a period of twenty-two years. His job was to help the Bureau catch Hanssen in the act so that they could be certain their case against him would be airtight. Hanssen was extremely savvy at reading people, and also extremely controlling and looming as a boss, dragging O'Neill and his wife Juliana to church and surprising them with unexpected visits, trying to convert Juliana, and requiring a great deal of loyalty from O'Neill to the detriment of his marriage. The strain eventually came to fruition as O'Neill procured the information necessary for the FBI to arrest Hanssen in the act. They caught him immediately after making a drop at Foxstone Park near his home in Vienna, Virginia.

The movie opens with a scene that makes clear that before the experience with Hanssen, O'Neill's ambition was to become a special agent doing undercover work. O'Neill admits in an interview with Rebecca Roberts that the experience of catching Hanssen gave him a bit of a bug for undercover work. His role in the operation would have been his ticket to fulfilling this ambition, but instead he decided to withdraw from the Bureau and become a lawyer. "A lot went into that decision; it was probably one of the hardest I've ever made," he admits. After the operation was over, O'Neill found himself evaluating his life and what he wanted to do with it. "I really needed to patch up my marriage, I really owed that to Juliana, my wife, who really had been put through the wringer through this process." He eventually made the decision to stick with Juliana, having realized that if he continued on the

¹ O'Neill and Ray (2007).

path to becoming a special agent doing undercover work, it would be “more of the same and more dangerous.”

After Hanssen’s arrest, then, O’Neill finds himself at a fork in the road, torn between two things that mean a great deal to him: his ambition and his wife. According to his assessment of the situation, he would have to give up one or the other, for to continue with both would leave him fractured and unable to give himself wholly to either one. Obviously, such a decision is not one he could make lightly, because it is the choice of one life or another. But it is also truly a crossroads, because it is not clearly a case of integrity and it is not a moral dilemma. O’Neill is in a position to choose the ideal by which to live his life, and as he arrives at the crossroads there is no prior reason to choose one ideal over the other. He must invent the reason.

2. An example loosely adapted from *The Namesake*.²

Gogol Ganguli is the son of Indian immigrants to New York. His parents’ marriage was arranged, and his mother was never entirely happy living in the United States. Gogol and his sister grew up with their feet in two cultures: the American world all around them and the small community of Bengali immigrants to which their parents clung. As tradition dictated, Gogol was given two names: Gogol, after his father’s favorite Russian author, and Nikhil, his “good name” given by his grandmother in India. While still a young child, he chose to go by “Gogol.” This proved to be a choice that gave him enormous grief as he grew up; the kids in school relentlessly teased him. He spent his youth resenting his father’s choice of a name and the fact of being different from his classmates because of his Indian heritage. This resentment, along with the distance of his father’s affection, created a rift between them, and as soon as he left home Gogol chose to be known by his good name—Nikhil, or Nik for short.

One day Nik brings his white American girlfriend home to meet his parents. After dinner, Nik’s father calls him away from the gathering and tells Nik the story of how he got his name. It is the story of the train wreck that caused him to decide, after recovering from his injuries, to travel the world and see it for himself instead of always reading about it in

² Lahiri (2003).

books. The book he was reading at the time of the wreck was by the Russian author Gogol. Somehow, though he does not fully understand why, this story moves Nik.

As the story really goes, Mr. Ganguli dies suddenly of a heart attack while spending a semester teaching in Ohio. This is a turning point for Nik/Gogol, who realizes just what his father meant to him, finally reads the work of the man for whom he was named, and comes to understand his father better posthumously. He embraces his Indian heritage and ends up marrying another Bengali-American, thus managing to find a way to live in both cultures.

But we can imagine a different story, according to which Nik's embrace of his heritage does not come as the result of an emotional revelation upon the death of his father. Imagine that after the dinner with his family and the American girlfriend, in which his father tells him the story of the train wreck, Nik comes to a better understanding of his father which opens up for him a path he was not previously considering, namely, embracing his Bengali heritage.

Imagine, then, that after meeting the American girlfriend Nik's parents begin a campaign to find him an Indian wife. Suppose they find and introduce him to a woman with whom he gets along well and to whom he is attracted despite himself. He may come to a point where he must choose between her and the American, which amounts to choosing whether to embrace his Bengali heritage or to reject it and live more as an American. At such a point he must choose an identity. Nik's options are not confined to just one or the other, of course, as they have been in my other examples. He has the option to find a middle ground, as most second-generation immigrants do, synthesizing his two competing identities. But the process by which he does this will look very much like what I have said Smith goes through as he deliberates about which of his options he cares about most. The case is similar in that an option opened to him by something he discovers within himself, something he comes to care about, gives him pause, and a reason to reconsider the life he was envisioning for himself. He must then choose how to define himself.

3. Robert E. Lee

General Robert E. Lee was descended from Virginia gentry, the son of a distinguished Revolutionary War general. His father died when he was twelve years old, and because the family was deeply in debt, he grew up in the homes of relatives. Nevertheless, he was a well-bred and proper child, and excelled in his studies. After completing grade school, he

followed his family's military tradition and attended West Point. His academic record was outstanding, and his sense of propriety was such that his behavioral record has been unmatched in all the years since his attendance. Biographer Michael Fellman³ emphasizes Lee's perpetual struggle for "self-control"; it was a struggle in which Lee quite clearly prevailed. After West Point he embarked on a distinguished military career.

Through the late 1850s the political climate between North and South became tense and Southern states began to talk of secession. Lee was reluctant to take sides on the matter; Fellman paints him as having values agreeable to pieces of both sides. "...Lee was a Unionist as well as a proud Virginian until secession forced him to choose between these two loyalties. At one level, this was an agonizing decision, but at another it proved to be a natural extension of his class, sectional, and racial identity."⁴ He was a staunch Unionist as things heated up in 1859-60, but when it came down to it he decided he couldn't fight against his family and friends and the land he loved. He was a thoroughgoing Virginian, embodying the genteel values of his time and culture.

Lee writes in a letter to his son Rooney (Fitzhugh): "Things look very alarming...I prize the Union very highly, & know of no personal sacrifice that I would not make to preserve it, *save that of honour*." Later, to cousin Annette Carter: "If the Union is dissolved, I shall return to Virginia & share the fortune of my people." Later still, to sister Ann Marshall: "With all my devotion to the Union and the feeling of loyalty and duty of an American citizen, I have not been able to make up my mind to raise my hand against *my relatives, my children, my home*." He became a reluctant secessionist. According to Fellman, however, Lee backed into the decision; it was "reluctant and only partially conscious."⁵

Lee's story is an imperfect example of the kind of case under study in this thesis. There are two reasons for this. The first is that Lee's decision was less than consciously self-constituting and active. Still, this is contingent; it would not be hard to imagine a man only slightly different from the historical Lee engaging with the dilemma and actively making the choice. Second, from our historical perspective there was a clearly morally correct choice as to which side of the Civil War to choose, and so it may seem as though a decision like Lee's should not be about his identity in a normative sense, even if as a matter of descriptive fact it

³ *The Making of Robert E. Lee* (2000).

⁴ Fellman (2000, p. 83).

⁵ Fellman (2000, p. 86, 89).

was. Here again, however, it is possible to imagine a situation in which the morally correct choice of a side is less than clear, even from the perspective of history. In such a case, one man's decision as to which side he is on may well be a matter of identity not only descriptively, but normatively as well.

The case as it actually happened, however, does have a feature that I wish to point out. Lee's attitudes changed as the war progressed, and he became more and more vitriolic against the North. For example, before the war he thought slavery a necessary evil, rather than positive good. But his position on many political and racial issues hardened over the course of the war; he became very resentful and vituperative against the North, and began to articulate quite racist views. In particular, Lee had a prejudice against black Union soldiers. He thought that being a soldier was something far too dignified for a person of (as he thought) inferior race to do. He and U.S. Grant had a long negotiation via correspondence concerning prisoner exchange, and Lee ended up doing his army a disservice by being stubborn about exchanging escaped slaves who had made it north and joined the Union army—Grant would settle for no less than full POW exchange, and had Lee accepted the terms, he could have replenished his forces to some extent. But he wouldn't.

This hardening of attitude vindicates my contention that a decision as momentous as this requires making the chosen option more important than the other. Lee's drift from moderation and being torn between the two sides into resentment of the North and staunchly racist views shows that he effected a full reorientation of his life, even if his shift in view was gradual and not particularly conscious. After his choice, he made the life of the South, with all of its flaws, fully his own.

4. Sartre's young man

In "The Humanism of Existentialism,"⁶ Sartre tells the story of a young man whose father had caused his mother great grief in supporting the Nazis when they invaded France and whose brother had been killed by the Germans. The young man had a strong desire to avenge the brother and join the Free French Forces in England in an effort to free his country from the invaders. On the other hand, however, he was all his mother had left, and she was deeply upset by the "half-treason" of her husband and the death of her elder son. If he left her, the young man knew she might fall into despair. As Sartre tells the story, staying

⁶ Sartre (2000).

in France with his mother was a sure success, but one which would affect only her; leaving to join the resistance was much less certain but could potentially have much wider effects.

Sartre draws certain lessons from this, which I discuss in the dissertation. But as a part of this appendix, the story itself serves as an alternative example to John Smith. The young man is caught in a deliberative impasse much like Smith's, though the two options are supported by somewhat different reasons. It is very difficult to tell what the young man has most reason to do; in fact, it is plausible to claim that his will-independent reasons do not admit of an all-things-considered judgment. His choice must be made by a decision to dedicate himself to one ideal or another: live as a loving and supportive son (perhaps doing what he can to support the resistance from within France), or live fighting for the freedom of his country and the memory of his brother.

References

- Alcott, Louisa May. 1983. *Little women*. New York: Scholastic.
- Anderson, Elizabeth. 1993. *Value in ethics and economics*. Cambridge, MA: Harvard UP.
- Appiah, Kwame Anthony. 2005. *The ethics of identity*. Princeton: Princeton UP.
- Austen, Jane, Richard Cronin, and Dorothy McMillan. 2005. *Emma*. Cambridge edition of the works of Jane Austen. New York : Cambridge University Press.
- Beauchamp, Tom L. and James F. Childress. *Principles of biomedical ethics*. 5th ed. New York: Oxford UP.
- Betzler, Monika. 2004. Sources of practical conflicts and reasons for regret. In *Practical conflicts: New philosophical essays*, ed. Peter Baumann, 197-222. Cambridge: Cambridge UP.
- Blustein, Jeffrey. 1991. *Care and commitment: Taking the personal point of view*. New York: Oxford UP.
- Boubil, Alan and Michael Schönberg. 1980. *Les Miserables*. Based on the novel by Victor Hugo. English lyrics by Herbert Kretzmer.
- Bratman, Michael E. 2000. Reflection, planning, and temporally extended agency. *Philosophical Review* 109 (1) (January): 35-61.
- Broome, John. 2000. Normative requirements. In *Normativity*, ed. Jonathan Dancy, 78-99. Cambridge: Blackwell.
- Carroll, Lewis. *Alice's adventures in Wonderland*. Available online at: <http://www.cs.cmu.edu/People/rgs/alice-table.html> .
- Chang, Ruth. Forthcoming. Voluntarist reasons. In *Practical Reason and Action*, ed. David Sobel and Stephen Wall. Cambridge: Cambridge UP.
- Coen, Ethan, Jeff Bridges, John Goodman, Julianne Moore, Steve Buscemi, David Huddleston, Working Title Films, Universal Studios Home Video (Firm), and Focus Features. 2005. *The big Lebowski*. Vol. 1 videodisc (118 min.). Universal City, CA : Universal Studios Home Entertainment : Focus Features.

- Cohen, G.A. 1996. Reason, humanity, and the moral law. In *The sources of normativity*, Christine Korsgaard. New York: Cambridge.
- Cooper, Chris, Ryan Phillippe, Laura Linney, Caroline Dhavernas, Gary Cole, Dennis Haysbert, Kathleen Quinlan, et al. 2007. *Breach*. Vol. 1 videodisc (111 min.). Universal City, Calif. : Universal Studios Home Entertainment.
- Dancy, Jonathan. 2004. Enticing reasons. In *Reason and value: Themes from the moral philosophy of Joseph Raz*, ed. R. Jay Wallace, 91-118. Oxford: Clarendon.
- . 2004. *Ethics without principles*. Oxford: Clarendon.
- . 2003. What do reasons do? *Southern Journal of Philosophy* 41 (Suppl.): 95-113.
- . 2000a. *Practical reality*. New York: Oxford UP.
- . 2000b. Should we pass the buck? *Philosophy: The journal of the Royal Institute of Philosophy* 47 (Suppl.): 159-173.
- D'Arms, Justin, and Daniel Jacobson. 2003. The significance of recalcitrant emotion (or, anti-quasijudgmentalism). *Philosophy: The Journal of the Royal Institute of Philosophy* 52, (Suppl.): 127-45.
- Darwall, Stephen. 2006. *The second-person standpoint: Morality, respect and accountability*. Cambridge, MA: Harvard UP.
- . 2002. *Welfare and rational care*. Princeton: Princeton UP.
- . 2001. Because I want it. *Social Philosophy and Policy* 18 (2): 129-153.
- . 1983. *Impartial reason*. Ithaca, NY: Cornell UP.
- . 1977. Two kinds of respect. *Ethics* 88 (1) (October): 36-49.
- Davidson, Donald. 2001. Actions, reasons, and causes. In *Essays on actions and events*, 2nd ed., 3-19. Oxford: Clarendon.
- Dewey, John. 2002. *Human Nature and Conduct*. New York: Prometheus.
- Dreier, James. 2004. Why ethical satisficing makes sense and rational satisficing doesn't. In *Satisficing and maximizing: Moral theorists on practical reason*, ed. Michael Byron, 131-154. Cambridge: Cambridge UP.
- Fellman, Michael. 2000. *The making of Robert E. Lee*. New York : Random House.
- Frankfurt, Harry G. 2004. *The reasons of love*. Princeton: Princeton UP.

- . 1999. *Necessity, volition, and love*. New York: Cambridge UP.
- Gert, Joshua. 2004. *Brute rationality: Normativity and human action*. Cambridge: Cambridge UP.
- Gibbard, Allan. 1990. *Wise choices, apt feelings*. Cambridge, MA: Harvard UP.
- Gowans, Christopher. 1996. Moral theory, moral dilemmas, and moral responsibilities. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 199-215. New York: Oxford UP.
- Hampton, Jean E. 1998. *The authority of reason*. Ed. Richard Healey. New York: Cambridge UP.
- Harman, Gilbert and Judith Jarvis Thomson. 1996. *Moral relativism and moral objectivity*. Cambridge, MA: Blackwell.
- Helm, Bennett. 2001. *Emotional reason: Deliberation, motivation, and the nature of value*. New York: Cambridge UP.
- Hill Jr., Thomas E. 1973. Servility and self-respect. *Monist: An International Quarterly Journal of General Philosophical Inquiry* 57 (January): 87-104.
- Howitt, Peter, Sydney Pollack, Philippa Braithwaite, William Horberg, Gwyneth Paltrow, John Hannah, John Lynch, et al. 1998. *Sliding doors*. Vol. 1 videodisc (99 min.). Hollywood, Calif. : Paramount.
- Hugo, Victor. 2005. *Les Miserables*. New York: Pocket Books.
- Humboldt, Wilhelm von. 1854. Introduction. In *The sphere and duties of government*. London: John Chapman. Available online at:
http://oll.libertyfund.org/?option=com_staticxt&staticfile=show.php%3Ftitle=589 .
- Ibsen, Henrik, and Nicholas Rudall. 1999. *A doll's house*. Chicago: Ivan R. Dee.
- James, Henry. 1940. *The portrait of a lady*. New York: The Modern Library.
- Jaworska, Agnieszka. 2007. Caring and internality. *Philosophy and Phenomenological Research* 74 (3): 529-68.
- Kant, Immanuel. 1997. *Critique of practical reason*. Trans. Mary Gregor. New York: Cambridge UP.
- Kingsolver, Barbara. 1998. *The poisonwood bible : A novel*. New York: Harper Flamingo.
- Korsgaard, Christine. 1996. *The sources of normativity*. New York: Cambridge.
- . *Locke lectures*. Available online at:
<http://www.people.fas.harvard.edu/~korsgaard/#Publications> .

- Lahiri, Jhumpa. 2003. *The namesake*. Boston : Houghton Mifflin.
- MacIntyre, Alasdair. 1984. *After virtue*. Second ed. Notre Dame: Univ of Notre Dame Press.
- Malick, Terrence, Colin Farrell, Christopher Plummer, Christian Bale, August Schellenberg, Wes Studi, Q'Orianka Kilcher, and New Line Home Entertainment (Firm). 2006. *The new world*. Vol. 1 videodisc (ca. 135 min.). Burbank, Calif. : New Line Home Entertainment.
- Marcus, Ruth Barcan. 1996. More about moral dilemmas. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 23-35. New York: Oxford UP.
- . 1980. Moral dilemmas and consistency. *Journal of Philosophy* 77 (March): 121-136.
- Mason, H. E. 1996. Responsibilities and principles: Reflections on the sources of moral dilemmas. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 216-235. New York: Oxford UP.
- Mill, John Stuart. 1986. *On liberty*. Buffalo: Prometheus.
- Mothersill, Mary. 1996. The moral dilemmas debate. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 66-85. New York: Oxford UP.
- Nagel, Thomas. 1979 (Canto edition 1991). Fragmentation of value. In *Mortal questions*, 128-141. New York: Cambridge UP.
- . 1970. *The possibility of altruism*. Oxford: Clarendon.
- Narveson, Jan. 2004. Maxificing: Life on a budget; or, if you would maximize, then satisfice! In *Satisficing and maximizing: Moral theorists on practical reason*, ed. Michael Byron, 59-70. Cambridge: Cambridge UP.
- Nozick, Robert. 1981. *Philosophical explanations*. Cambridge, MA: Belknap.
- O'Neill, Eric and Billy Ray. 2007a. "'Breach' Captures an FBI Traitor in the Act." By Rebecca Roberts. *Weekend Edition Sunday*. NPR, February 4, 2007.
- . 2007b. "Eric O'Neill and Billy Ray Discuss 'Breach'." By Teri Gross. *Fresh Air*. NPR, January 31.
- Orczy, Emmuska. 2002. *The scarlet pimpernel*. New York : Modern Library.
- Parfit, Derek. 2001. Reasons and rationality. In *Exploring practical philosophy: From action to values*, ed. Dan Dgonsson, Jonas Josefsson, Bjorn Petersson and Toni Ronnow-Rasmussen, 17-39. Burlington, VT: Ashgate.

- Parfit, Derek, and John Broome. 1997. Reasons and motivation. *Aristotelian Society: Supplementary Volume*, 71 (Suppl.): 98-146.
- Pettit, Philip and Michael Smith. 1990. Backgrounding desire. *Philosophical Review* 99 (4) (October): 565-92.
- Railton, Peter. 1996. The diversity of moral dilemma. In *Moral dilemmas and moral theory*, ed. H.E. Mason, 140-166. New York: Oxford UP.
- Raz, Joseph. 1999. *Engaging reason: On the theory of value and action*. Oxford: Oxford UP.
- . 1990. *Practical reason and norms*. Princeton: Princeton UP.
- Richardson, Henry S. 2004. Satisficing: Not good enough. In *Satisficing and maximizing: Moral theorists on practical reason*, ed. Michael Byron, 106-130. Cambridge: Cambridge UP.
- Rosati, Connie S. 1996. Internalism and the good for a person. *Ethics* 106 (2): 297-326.
- . 1995. Persons, perspectives and full information accounts of the good. *Ethics* 105 (2): 296-325.
- . 1989. Self-invention and the good (ethical theory). Ph. D. diss., University of Michigan.
- Sartre, Jean-Paul. 2000. The humanism of existentialism. In *Essays in existentialism*, ed. Wade Baskin, 31-62. New York: Citadel.
- Scanlon, T. M. 1998. *What we owe to each other*. Cambridge, MA: Harvard UP.
- Schmidtz, David. 2004. Satisficing as a humanly rational strategy. In *Satisficing and maximizing: Moral theorists on practical reason*, ed. Michael Byron, 30-58. Cambridge: Cambridge UP.
- Shiffrin, Seana Valentine. 2000. Paternalism, unconscionability doctrine, and accommodation. *Philosophy and Public Affairs* 29 (3) (Summer): 205-50.
- Slote, Michael. 1989. *Beyond optimizing: A study of rational choice*. Cambridge, MA: Harvard UP.
- Smith, Michael. 1994. *The moral problem*. Cambridge: Blackwell.
- Taylor, Charles. 1985a. The diversity of goods. In *Philosophy and the human sciences*, 230-247. Cambridge: Cambridge UP.
- . 1985b. What is human agency? In *Human agency and language*, 15-44. Cambridge: Cambridge UP.
- Ullmann-Margalit, Edna, and Sidney Morgenbesser. 1977. Picking and choosing. *Social Research* 44 (4) (Winter): 757-86.

- Velleman, J. David. 2003. Narrative explanation. *Philosophical Review* 112 (1) (January): 1-25.
- . 2002. Motivation by ideal. *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action* 5 (2) (May): 90-104.
- . 2000. The possibility of practical reason. In *The possibility of practical reason*, 170-199. Oxford: Clarendon.
- Weber, Michael. 2004. A new defense of satisficing. In *Satisficing and maximizing: Moral theorists on practical reason.*, ed. Michael Byron, 77-105. Cambridge: Cambridge UP.
- Westlund, Andrea C. 2003. Selflessness and responsibility for self: Is deference compatible with autonomy? *Philosophical Review* 112 (4) (October): 483-523.
- Williams, Bernard. 1982. *Moral luck*. New York: Cambridge UP.
- . 1985. *Ethics and the limits of philosophy*. Cambridge, MA. : Harvard University Press.
- Wolf, Susan. 1982. Moral saints. *Journal of philosophy* 79 (August): 419-439.