

**Constitutivism
and Practical Authority**

Euan Russell Allison

University College London

Philosophical Studies MPhil Stud

Declaration

I, Euan Russell Allison, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

I offer a restatement of ‘Kantian constitutivism’ – the position articulated by Korsgaard in her *Sources* (1996) – and argue that it can be defended against recent challenges (e.g. Enoch 2006). Chapter One provides some motivations for pursuing this account by rejecting major competitors. I claim that ‘realist’ interpretations of practical reasons cannot explain why, insofar as we are rational, conclusive reasons will motivate us. But it is difficult, once we accept an alternative ‘desire-based’ account, to accommodate the existence of unconditional reasons (I focus particularly on moral reasons, a subset of such reasons). Chapter Two begins to develop the constitutivist view, which explains *both* how there are unconditional reasons *and* how conclusive reasons will motivate rational agents. This is a desire-based view, but one which denies that the content of our desires is entirely contingent. *If* we care about some things insofar as we are agents, these will ground reasons which apply, necessarily, to all. To vindicate this, I explain why the authority of considerations bearing on constitutive features of agency need not depend upon whether we want to be agents. Chapter Three demonstrates that commitment to certain attitudes *is* constitutive of agency. This issue is pursued via another question which confronts desire-based theories: since desiring something does not by itself make it good, what determines that we should treat certain desires as reasons? Following Korsgaard, I show that agents are committed to valuing the capacity for practical reason itself, and do not need further reasons for doing so. This provides the criterion by which other evaluative attitudes can be selected as reasons, with distinctively moral implications. Finally, I explore an apparent tension between this conclusion and my strategy in Chapter Two, where I suggested we can *disvalue* our agency yet this does *not* undermine the authority of its constitutive standards.

Impact Statement

This research develops an account of practical reasons according to which they are based in *constitutive* standards of action. The impact this work might have within academic philosophy is far reaching. If the account is worth serious consideration, this suggests that a certain methodological approach in the philosophy of action may be called into question. This is because if the account is true, it follows that the metaphysics and normativity of action are not discrete topics of enquiry – we could not settle what an action *is* independently of determining what makes an action *good* (in the sense of supported by reasons). Thus, teaching and research in the philosophy of action which tacitly assumes their discreteness risks confusion. Impact within academic philosophy, both nationally and internationally, could be brought about through publication in journals, as well as presenting at, or organising, conferences and seminar series on related themes. In this way it will be possible to engage with individual teachers and researchers, and so over time affect the practice of the discipline as a whole.

Outside academia, the main benefits I see occurring are in practical ethics, particularly in health and social care contexts. There is much focus in these sectors on the autonomy of patients and clients, and how care provision should be responsive to the demand that their autonomy be respected. The ideas which my research deals with intersect with questions about what it is to act *autonomously* – and thus what it is to *be* autonomous, or at least *capable of* autonomous action. Hence, they are of direct consequence for people working in the aforementioned sectors who want to better understand certain features of human life which place ethical demands on them, and how these interact with their particular professional environment. I think impact is most likely to be brought about here by reaching out to individuals and smaller organisations operating in the sector at a regional and local level. Opening an informal dialogue on these issues could provide a platform for knowledge exchange. This might impact how care professionals conceive their roles, and further academic research could be informed by developing an insight into their experiences of working in a context where concern for autonomy is to the forefront. Both benefits could occur immediately.

Contents

Introduction	6
1 Practical Authority and Realism	7
1.1 The Plan	7
1.2 The Relationship Between Normative and Explanatory Reasons	7
1.3 What I mean by ‘practical authority’ via Williams	12
1.4 Excursus: Epistemic Authority	18
1.5 Realism: A Brief Statement	20
1.6 An Argument Against Realism	22
2 Constitutivism and Its Critics	29
2.1 Constitutivism: An Introduction	29
2.2 The Constitutivist Account of Practical Authority	32
2.3 “Why be an agent?”	36
2.4 Normative Inescapability	36
2.5 Justification Without Agents’ Reasons?	42
3 Practical Authority and Moral Authority	46
3.1 Taking Stock	46
3.2 Some Preliminaries	47
3.3 The Moral Character of the Constitutive Aim of Agency	49
3.4 The Argument	52
3.4.1 Rational Agency	52
3.4.2 The Problem	53
3.4.3 The Solution	56
3.4.4 The Unconditioned Condition of Value	58
3.4.5 Inside Agency	60
3.5 The Value of Agency and Regret about Agency	63
Conclusion	67
Bibliography	68

Introduction

My first task here is to show that two leading accounts of the nature of practical reasons – ‘realism’ and certain ‘desire-based’ views – are inadequate. Neither fully captures what it is to be subject to reasons. This dialectical predicament appears particularly desperate given that the grounds for rejecting either one of these positions provides *prima facie* support for the other. I propose a ‘constitutivist’ theory as a “middle way” through this impasse. I argue that it deals with the problems of the other views without losing anything in the process.

Having put things in this schematic form, we can now turn to what is both interesting and important here by gesturing at how the case will develop. Chapter One describes the phenomenon I call ‘practical authority’. This includes a certain conception of how normative reasons are related to the explanation of action – specifically how they are related to motives. It also includes the idea that certain considerations are relevant to how we should act *irrespective* of how we happen to feel about them (for example, that your action will cause someone pain matters for whether you should do it, whether you think it does or not). In other words, there are unconditional reasons. The task of explaining practical authority is, I think, part of the task of explaining how any consideration makes a claim on our actions.

I will also argue in Chapter One that neither realism nor certain desire-based views explain practical authority. The former because it fails to illuminate the relationship between normative reasons and the explanation of action; the latter because they cannot accommodate unconditional reasons. The alternative is this. We should agree with desire-based views that reasons are necessarily related to antecedent motives – retaining their plausible account of the relationship between normative reasons and the explanation of action. But we should deny that the content of our motives is always contingent – thus allowing for unconditional reasons. We are committed to certain attitudes insofar as we are agents.

Chapter Two defends this account in conditional form – *if* certain attitudes are constitutive of agency (in the above sense) then they give us unconditional reasons. I respond to Enoch’s (2006; 2011) objection that this is false because we might fail to care about being agents. Chapter Three shows there *is* an attitude which is constitutive of agency – following Korsgaard (1996), *valuing the capacity for practical reason*. I discuss what might, therefore, be unconditionally important, and how the conclusion interacts with the possibility of regret about agency.

Practical Authority and Realism

1.1 The Plan

This essay concerns what we, rational agents, and our reasons must be like in order that the latter are authoritative for the former. This chapter articulates the claims that comprise the thesis that reasons are authoritative for us – ‘practical authority’ – and shows they are implicit in practical thought (1.2 and 1.3). This first requires an exploration of the relationship between normative and explanatory reasons (1.2). I also set up a dilemma of sorts, to be resolved in subsequent chapters. I argue that ‘realist’ accounts of reasons cannot explain why conclusive reasons will motivate us insofar as we are rational and so are inadequate accounts of reasons (1.5 and 1.6).¹ This suggests that, contrary to realism, there is a necessary connection between reasons and antecedent motivations. But, and leaving aside other misgivings², at least one prominent view which accepts this (Williams 1981b, 1995) rejects the existence of unconditional reasons (1.3). The existence of unconditional reasons – alongside the claim about how reasons are capable of motivating us – is included in the idea of practical authority. So neither realism nor (seemingly) its contrary account for practical authority.

1.2 The Relationship Between Normative and Explanatory Reasons

Statements with the form “X has a reason to ϕ ” attribute *normative* reasons to agents. Normative reasons are considerations that count in favour of performing certain actions. Our concept of a reason also admits of a separate *explanatory* function. Indeed, we often cite normative reasons for this purpose. If X were to ϕ intentionally, then one way her ϕ -ing can be made intelligible is by articulating the normative reasons she (correctly) believed she had for ϕ -ing.³ But more generally, a distinction between normative and explanatory reasons is appropriate because the considerations which explain action need not also justify them – indeed, actions

¹ The claim that conclusive reasons *will* do this is developed in 1.2.

² See 3.4.2.

³ I am focusing just on the cases of intentional action which are also actions done for reasons. Unless otherwise stated, ‘action’ refers to actions done for reasons, and ‘agency’ refers to the sort of agency that issues in this sort of action.

may not be justified at all, though we can explain them. That I believe the train leaves at 3pm might account for why I arrive at the station at 3pm. It was not – since the train left at 2pm – a *good* reason for doing so (or rather, no reason at all in the normative sense).

How close is the connection between normative and explanatory reasons? Is it exhausted by what I have said – that it is sometimes possible to cite normative reasons in explanations of action? I claim the relationship is closer. Normative reasons are, necessarily, potential explanatory reasons – for any *true* attribution of a normative reason to an agent, it must be possible that the reason could feature in an explanation of her actions. Some work is needed to explicate the sense of possibility this formula appeals to. But we cannot approach this issue without first understanding why there is *any* conceptual connection between normative and explanatory reasons. Consider first the role that the concept of a reason plays in human life, and how it interfaces with certain attitudes and practices that exist therein. That is, why do we have the concept at all? What features of human life supply it with the content it has – content which we feel is significant for how to act? Here is one hypothesis. Human action is subject to a distinctive form of assessment, namely, normative assessment. We think we should recognize certain considerations for the purposes of guiding our actions – those suitable to the task. This requires that we recognize these considerations *qua* considerations of that kind. The preceding goes beyond the idea that we assess outcomes of actions as good or bad, since this need not appeal to whether actions are motivated by considerations which are recognized by the agent as appropriate. Our coming to recognize the appropriateness of considerations in this sense, which is part of the object of normative assessment, can always (in principle) arise from a deliberative process. ‘Reasons’ is the name we give to *what* we must deliberate with in order that we arrive at the right conclusion – at least one kind of failure would be to operate with what we wrongly believe are reasons. Thus, when an agent is motivated by considerations she ought not to be, she acts contrary to what she has *most* reason to do.⁴

If this is correct – the concept of a reason is bound up in this way with our being subject to normative assessment – then we may learn more about reasons by adding detail to our description of normative assessment. Notice, there is an intimate relation between the possibility of normative assessment and the

⁴ This is not to imply that reasons are literally “weighed” according to their strengths, only that there is some kind of mediation between the inputs of deliberation which issues, at the all things considered level, in conclusions about what we should do.

appropriateness of praise and criticism.⁵ Helping ourselves to the hypothesis above, we can say that praising action is only appropriate when there was a reason to act that way; criticism of action is only appropriate when the agent had a reason to act otherwise; and both are unintelligible when there were no reasons for or against the act in question. The condition of the intelligibility of praise and blame is the possibility of normative assessment – of considering the action in light of reasons.

There is also an entailment which goes in the opposite direction. For this claim, all we need is the idea that whenever an agent acts contrary to what she has most reason to do, criticism is necessarily appropriate. So, the appropriateness of a certain kind of criticism is also a condition of the possibility of a form of normative assessment: that an agent had most reason to act otherwise than she did. This has significant implications for our understanding of the concept of a reason. If reasons are the proper inputs for deliberation, we have to assume they are relevant to the determination of what we have most reason to do – the conclusion of deliberation. Since they are independent considerations in favour of particular actions, I see no other way they could be relevant unless they could each, in principle, be part of the case for an action we have most reason to perform. This is not to say they might not, in current circumstances, be defeated – that is, part of the case for an action we do *not* have most reason to perform.⁶ But if some consideration would be defeated in all conceivable circumstances I think it could not be relevant to practical deliberation – it could not count as a reason. Now, if we do not do what we have most reason to do, we are criticisable. But we cannot be criticisable for failing to do what we cannot do. So we must be able to do what we have most reason to do. But all of our reasons could each, in principle, be part of the case for an action we have most reason to perform. So we must be able, in some conceivable circumstance, to do what we have *any reason at all* to do – and for those reasons. I can imagine an objection to this argument. Perhaps in the course of deliberation we learn we could never perform some action for which there is a reason. In which case, all this shows is that the reason could not have been part of a conclusive case for performing an action⁷ – because that action is not one we could ever perform. This seems wrong. If it is conceded that the point

⁵ I am not talking about *blame*. Agents may be criticisable for their actions in my sense, even when they have excuses for so acting that render blame inappropriate.

⁶ And obviously, in this case, we are not criticisable for failing to perform that action.

⁷ I will sometimes say reasons are ‘conclusive’ – by which I mean there is some action we have most reason to perform, the case for performing which may be made up of multiple reasons which need not be independently decisive. My claim is that for any reason at all there must be conceivable circumstances in which it is part of such a case.

of the concept of a reason is to account for what (rightly) goes into our deliberations, then the correct response to learning we could never act on some consideration, and so this could in no circumstance be what we should do all things considered, is to conclude that this consideration was not, after all, a reason.

So, normative reasons are, necessarily, potential explanatory reasons in the sense that we must be able, in some conceivable circumstance, to do what we have *any reason at all* to do – namely, in those circumstances in which those reasons are conclusive. Notice we arrived at this conclusion to account for the appropriateness of criticism for failing to act on conclusive reasons. Part of the object of this kind of normative assessment is the recognition we should act on certain considerations. Thus, the condition that we must be able to do what we have reason to do would not serve its purpose if it only imposed the restriction that we must be able to perform a series of bodily movements that mirrors the actions we have reason to perform. If normative reasons are capable of explaining action – and thus of inviting criticism when we act contrary to conclusive reasons – we must be capable of performing actions *because* we recognize the reasons supporting them.

This last point is important for how we should further unpack our initial formula that it must be possible for normative reasons to explain action. We may begin with a suggestion from Raz, and then consider how his view needs developing to accommodate our findings so far. Raz claims that the physical impossibility of performing an action excludes our having a reason to perform it if the impossibility is “deep” (if not quite *metaphysical*):

It is not the case that I have reason to attend the symposium that was recorded by Plato in the dialogue of that name... but I have reason to go to the concert conducted by Abbado tonight even though I cannot as it is sold out. (2011: 143)

In addition, there are epistemic constraints on what actions can be done for (genuine) reasons. Suppose a meteorite is going to hit my house, and that the movements of the meteorite cannot, even in principle, be predicted – so nobody could know it is going to hit my house. In this case, I cannot leave my house *to avoid the meteorite* (Ibid: 34-35). Even if the bodily movements which would be involved in that action are easily performable – and even performable for some reason, say, to go shopping – the *action* itself is not performable. This is because, as we have seen, when normative reasons explain an agent’s action they do so by highlighting the justification she rightly saw for performing it. But the consideration that the meteorite will hit my house – for which the action is done and which is included in its description – is not something which I could *rightly* take to support

an action, even allowing that it does actually support it. It is not a consideration which I could know applies at all. Thus, it cannot guide my actions in the way normative reasons do when they explain action – which is just to say it is *not* a normative reason.⁸

Do these constraints suffice as an elaboration of the idea that normative reasons are, necessarily, capable of explaining actions? The main omission becomes apparent if we suppose the considerations which support an action are knowable and accessible to me now. If a consideration which justifies my action also explains it, it cannot just be that I know it justifies the action. My awareness of this fact must *motivate* me to perform it. What would we say if I was unmoved by *conclusive* reasons which favour some action? Presumably that I am irrational.⁹ We would not say the considerations in question are not really reasons. Why is this? When conclusive reasons fails to be *actually* explanatory, we are entitled to resist the idea that they were not, after all, reasons because we can point to some condition of the *agent* which is compatible with the *considerations* still being *potentially* explanatory in the circumstance, e.g. it is physically possible to act on them but I am too tired to do so; they are knowable but I am ignorant. Something similar is needed in this case to explain why it is something about me, and not the putative reasons, which prevents the considerations from being explanatory. But of course this cannot be captured by Raz's constraints alone since we suppose I am aware of the considerations and that it is not "deeply" physically impossible for me to act on them – which is enough to satisfy his constraints. The solution is that it is *also* a requirement for considerations to be reasons that, when they are conclusive, they will motivate me insofar as I am *rational*.

This constraint needs some further explanation. It is obviously *not* rational to be motivated by every reason that applies to me on a given occasion. It is not

⁸ Consider cases in which the value-making features of an action are, in an ordinary sense, inaccessible to me in advance of performing it: I know that a dish will be good, because my friend (who is reliable on these matters) told me, but to know *why* it is good I need to try it for myself. Raz's epistemic condition is compatible with the features of the dish which make it good being reasons for me to eat it – they are not in principle unknowable (my friend knows them), so in circumstances in which I know them (I have already tried the dish) I can act on them. In advance of the action, however, they cannot be conclusive reasons for me, since I cannot now act on them (recall that I am criticisable for failing to act on conclusive reasons, but I cannot be criticisable for failing to do what I cannot do). If I now have conclusive reason to try the dish, this may be because I enjoy the surprise element of new experiences (which I know, at least, will be good).

⁹ I am focusing on conclusive reasons because we have said it is because all reasons are potentially conclusive, and acting against conclusive reasons is always criticisable, that all reasons must be potentially explanatory. It is of course rational for agents to be unmoved by defeated reasons. See below.

rational, for example, to be motivated to pursue some end I have reason to pursue, when I now lack means to pursue it. So too, it would be hopelessly inefficient if I were motivated even by reasons which I take to be defeated – since this would act as an obstacle to my doing what I have most reason to do. More generally, the fact that the options which are available to us, and for which we have reasons, are incommensurable, shows it would be pointless to be motivated by all the reasons that apply to us: in the end we cannot do all we have reason to do.¹⁰ This is why our formula requires of our reasons *only* that they would motivate us insofar as we are rational *when they are conclusive*. But why is this a constraint on any consideration being our reason and not simply a fact about how we are responsive to reasons that, in fact, are conclusive on a given occasion? Because, as I argued already, *for each of our reasons* there must be conceivable circumstances in which they are part of the case for an action we ought to perform. If reasons are the proper inputs for deliberation about what we ought to do, then it does not make sense to say something is our reason if it is irrelevant to what we ought to do – which it would be if it were defeated in every conceivable circumstance. But there could only be conceivable circumstances in which some reason is part of the case for an action we have most reason to perform if, in those circumstances, it would motivate us insofar as we are rational. This is because of the appropriateness of criticism for failing to act on conclusive reasons and the *in*appropriateness of criticism for failing to act in a way which we cannot. Hence, it matters to whether anything is our reason that, if it were conclusive, it would motivate us insofar as we are rational.¹¹

1.3 What I mean by ‘practical authority’ via Williams

It follows from the claim that conclusive reasons will motivate us insofar as we are rational that agents who act on conclusive reasons satisfy a certain requirement of

¹⁰ Ulrike Heuer has suggested to me that reasons pertaining to past states of affairs might also belong in this category of reasons it would be irrational to be motivated by. But the right response here is simply to deny that we can now have reasons concerning what we might have done in the past. They cannot explain our actions because of the “deep” physical impossibility of acting on them. We have some reasons to act now because of things which happened in the past, but those reasons pose no special problem.

¹¹ It does not follow that supposing a consideration favours an action, and satisfies the various conditions upon being a reason, it is thereby a reason to pursue any end it might be taken to support. It would not, for example, be a reason to pursue ends that are “deeply” physically impossible to pursue. We do not have reasons to pursue those ends because the considerations which are taken to support them could not explain our pursuing them – nothing could. But these considerations are still be capable of explaining our pursuit of other ends and so are reasons for pursuing them.

rationality.¹² How can it be rationally required that we are motivated by some consideration? This is both a question about the nature of practical rationality and about the nature of reasons. Minimally characterized, reasons are considerations which justify certain actions. Thus, questions like “R is a reason to ϕ , but does it count in favour of ϕ -ing?” and “R is a conclusive reason to ϕ , but should I ϕ ?” are unintelligible. If someone asks these questions we suppose they do not sincerely believe R is a reason to ϕ – ‘reason’ would appear within scare quotes in direct reports of their utterances.¹³ Otherwise, such agents are conceptually confused. We can now add to these claims that an agent who affirms that R is a conclusive reason to ϕ is irrational if she is unmoved by this judgement. What must the considerations that are *reasons* be like such that they are *rationally related* in these ways to our responses? We will not arrive at a satisfactory answer in this section, nor in the remainder of the chapter. But this is what I hope to achieve by the end of the essay. We may, provisionally, refer to the relationship between reasons and rationality which I have described as ‘practical authority’ – though we have not said enough yet to capture all the features of being subject to reasons which I intend this phrase to cover. To see this, I will set out one strategy for explaining this provisional notion of practical authority, exemplified by Williams (1981b, 1995), and consider what it does not account for in practical thought. Namely, Williams’s account excludes the possibility of unconditional reasons. I think unconditional reasons exist – I shall try to explain why – and include this idea in the full notion of practical authority.

We can approach Williams’s view via the “Humean” dogma that motivation always originates in the prior wants, wishes, dispositions, and so on of the agent – ‘desires’ or ‘motives’ to refer to them collectively.¹⁴ *Rational* motivation results from the application of some rational procedure to our antecedent desires. Since all our reasons will, when conclusive, motivate us insofar as we are rational, the motivation for acting on them must also originate in our antecedent desires, and the considerations which constitute our reasons must appeal – via some rational

¹² Unless their motivation is the result of hypnosis or other rationally deviant routes. Thus, I reject the idea that practical rationality is *simply* responding to genuine reasons. I should add that conclusive reasons will only motivate us if we *know* what our reasons are. Thus, having correct beliefs about reasons is itself a requirement of rationality. I leave aside the issue of how this requirement interacts with the requirement that we are motivated by conclusive reasons, given that our beliefs about reasons may be affected by false non-evaluative beliefs which are nonetheless rational (because our evidence is misleading).

¹³ Perhaps what they mean would be better expressed by “Society tells me this is a reason to ϕ , but is society right?”

¹⁴ In calling this thesis a ‘dogma’ I do not mean to prejudice the case against it. The important question is not whether it is a dogma, but rather whether it is a philosophically *useful* one.

connection – to them. Williams also accepts the idea that the content of our desires is ultimately contingent (1995: 36-37).¹⁵ This is true of those we can arrive at by deliberating rationally as well as of our antecedent motives. This is because, if we accept the Humean dogma, the motives which are conclusions of deliberation are partly determined by our antecedent motives which are its premises. It follows that the reasons we have are also ultimately contingent. I shall return to this shortly. But we can witness here the outlines of a dual analysis of the nature of reasons and rational deliberation which explains why conclusive reasons will motivate us insofar as we are rational.

Again, my aim is not to defend Williams's view in its entirety. But his basic insight that there is a necessary connection between reasons and prior motives must, I shall argue later (1.6), be sound. Hence, I must give the shape of a response to criticisms bearing on this point. Recall, when normative reasons explain action, the justification they provide for the action itself accounts for our performing it. This insight is preserved within the current framework if we simply *identify* reasons with motivating desires – those which actually move us to action. The justification for doing something is just *that we do it*. But this view eliminates the possibility of ever acting contrary to reason, and so is unacceptable. It must be possible to judge that we should have acted otherwise than we did and for that judgement, sometimes, to be *correct*. Suppose I correctly judge that I have a reason – it appeals to some desire I would have if I were rational. (We need not specify the rational procedure that is applied to my antecedent desires for our present purposes.) What we must allow is that I can make this judgement when not actually under the influence of the relevant desire – when I am not (yet) rational in this respect. The problem with conceding this emerges if we add an assumption about the process which determines motivating desires: that it essentially consists in desires of various strengths competing with one another. Ulrike Heuer has suggested to me that, on such an account, it would just be “serendipity” if I were motivated in accordance with reason. This is because to allow for the possibility of error, my judgement that I have a reason cannot partake in the process determining motivating desires. The explanation of my action when I act as I have most reason to is always just luck – and never, as I have suggested must be possible, the justification which I believe there is for so acting. The account of reasons is thus inadequate.

The mistake in this argument lies in the assumption about the nature of the process determining motivating desires. We can see this by adopting an

¹⁵ I show in Chapter Two that this is a separable feature of the framework.

instrument of Nagel's (1970: 29), the "motivated" desire. The idea is that, often, the motivating desires that are cited in explanations of action also require explanation – they are themselves motivated. It is hard to conceive of highly specific desires, e.g. to purchase a certain holiday package, as "just existing" in an agent. Nagel took this to support the view that desires are not essential to the explanation of action – that an agent's grasp of *reasons* for both desire and action is primary. But we need not accept this part of the analysis. It is at least consistent to hold onto the insight that some *cognitive* explanation why an agent has certain desires is needed, without accepting that the explanation must appeal to the grasp of reasons that are conceptually unrelated to prior motives. All we need is the idea that some of our most deeply ingrained motives can be highly general and subject to local or even global conflict. Sets of desires like this cannot issue in intelligible behaviour without the mediation of reasoning. In the most basic case, this will involve forms of instrumental thought, i.e. about what I can do to satisfy this desire now. In other cases, things are more complicated. If multiple, and equally strong, wants cannot be jointly satisfied on a given occasion, what am I to do? Williams suggests that changes in desire can intelligibly be brought about by "imagining" what is involved in fulfilling some desire and asking whether we would like it if this were to happen (1981b: 104-5). In this way, the conflict may dissolve.¹⁶ Let us follow Williams in saying that when we apply these forms of thought to our antecedent motives we take a "sound deliberative route" (1995: 35). To the extent we are motivated by the conclusions of such deliberation we are rationally motivated – and when we are, this will be no accident. It will be these conclusions *qua* conclusions of rational deliberation which explain our actions. So, the way Williams avoids the objection above is to identify reasons with these conclusions: what we have most reason to do is just what we desire to do most after such deliberation; what we desire to do at all after such deliberation (though will not be motivated by if we are fully rational), we have *prima facie* reasons to do. It is easy to see that this view retains the idea – rejected by Nagel – that our reasons are necessarily connected to prior motives. This is because our reasons just are considerations directed towards desires we can arrive at by rational deliberation *from* antecedent motives, and thus are partly determined by them.

¹⁶ This would show that one of the conflicting sets of wants did, after all, ground conclusive reasons (see below). It is possible that even after deliberating perfectly there are equally strong reasons to perform incompatible actions and there are no other actions which there are stronger reasons to perform. In this case, I assume it is rational to perform any of the actions that are best supported by reasons. That an agent who finds herself in this position has (on Williams's view) equally strong competing desires does a good job of explaining why acting is often hard in such situations. I thank Ulrike Heuer for bringing this case to my attention.

Now to the larger significance of Williams's view for the argument of this essay. We will approach this circuitously. First, notice how we think something is amiss when a person admits that torturing animals (say) is immoral but engages in such activity uninhibited. We conclude that she is either being insincere, or she is not competent with moral concepts. Why do we do this? Because moral judgements are judgements about reasons. To judge that an action is immoral is to judge we have (or had) a reason to avoid performing it – and often an overriding one. Thus, as a class, moral judgements are only valid if moral considerations are potentially explanatory of action in the way outlined in 1.2. Given this, there are certain responses we might expect from an agent who makes a sincere and competent moral judgement – and so appreciates (in probably quite an inchoate fashion) the conceptual connection between her judgement and action.¹⁷ If she is not sufficiently motivated to perform the moral action, she is at least conflicted over its performance. This conflict manifests itself, characteristically, in the experience of guilt. She may express the wish that she did not have an inclination to perform the immoral action. In our example, we suppose neither of these things happens. We can see that whilst our evidence for rejecting the judgement as insincere or confused is empirical, it points to a connection which is conceptual: an agent cannot sincerely and competently make a certain kind of moral judgement (one that identifies overriding reasons) unless she will act on it insofar as she is rational.

Suppose the agent is trying and failing to deploy moral concepts – rather than being insincere. Perhaps “Torturing animals is immoral” in her mouth means what we express with statements like “Society judges we should not torture animals.” That is, she has confused a normative statement – about what reasons she has – with a purely descriptive one – about what reasons people think we have. Operating with this hypothesis, we can see how the question “Torturing animals is immoral, but why should I not torture them?” might *seem* intelligible from the agent's perspective. But the question is not intelligible when interpreted in the usual way. The agent is certainly confused if she thinks her descriptive statements capture the *normative* phenomena we refer to using the language of ‘reasons’ and ‘morality’. It is clear that the main obstacle to her being motivated by moral considerations is that she does not identify them through the medium of moral judgement to begin with. In this respect, her failure is similar (though more global, and distinctively *incoherent*) than an agent who fails to be motivated by a (conclusive) moral reason because she is ignorant of its existence.

¹⁷ I assume the judgement identifies overriding reasons.

Finally, suppose the agent who asks “Torturing animals is “immoral”, but why be “moral”?” *does* understand that genuine moral judgements are judgements about reasons, and hence that it can be irrational to be unmoved by our own moral judgements. But if *that* is how moral concepts function, she protests, may she not simply refrain from making moral judgements? Perhaps other people will label her actions ‘moral’ or ‘immoral’. But if they think the standards they invoke in such assessments are *normative* for her – that they give her reasons – they are mistaken. This is because, she claims, there is no guarantee that it is ever irrational for her to ignore “moral” standards – and as we have seen, something is only a reason if, when conclusive, it would motivate us *insofar as we are rational*. Of course, standards can be rationally binding on us – in a conditional sense – if we *want* to conform to them. There would be a *kind* of irrationality involved when agents who have moral action as their goal are indifferent to their moral judgements. But whether irrationality of this kind applies here depends upon whether we want to be moral – and what if we do not? Even if we all do, it would not follow that “moral” standards are normative *as such*. The objection may be expressed like this: the agent acknowledges there are “moral” standards, but she wants to know why they are *her* standards – standards she must take into account in deliberation. Given our claim about how moral concepts function, this comes to the same thing as questioning whether moral standards apply to her. Thus, her indifference to “moral” standards is not trivially irrational because her “moral” judgements are not judgements about moral reasons – *those* she denies having. Her question “Why be “moral”?” is not the offspring of incoherent requests to justify compliance with reasons.

But moral standards *are* normative as such. The suggestion that the skeptical challenge above must be taken seriously is deeply unattractive. This is something of a presupposition of this essay. However we can say a little more in its favour. In ordinary moral practice, people are held accountable no matter their particular aims or desires. But I have argued that criticism of action is only appropriate if the agent had a reason to act otherwise than she did. Thus, our practices of moral criticism cannot be appropriate if moral standards bind us, if at all, only conditionally, as a result of contingent purposes of agents. Practices of moral criticism detach entirely from facts about such purposes. The appropriateness of these practices thus depends upon moral standards being normative *as such*. They must give us reasons to act whether we want to conform with them or not. There are some standards, such as moral standards, about which we can *never* intelligibly say “But why conform with them?”. Some reasons are unconditional. Since practices of moral criticism in which this idea is implicit are both deeply

ingrained in human society and largely held to be appropriate, this must create a presumption against its denial. Let us add it to the group of ideas – central to our being subject to reasons – under the heading ‘practical authority’.¹⁸

Williams denies practical authority. His view is incompatible with the possibility of any unconditional reasons at all. This is, first, because he assumes the content of our desires is ultimately contingent. Therefore, there is nothing inconceivable about someone who does not, and could never rationally, care about acting morally, or whatever else. Further, Williams holds that motivation is always partly determined by antecedent desires. Finally, we have seen that any genuine reason will, when conclusive, motivate us insofar as we are rational. Taken together, these premises entail that it is possible some people might not be subject to moral reasons. For reasons adduced already, I reject this conclusion. Is there an alternative framework that explains practical authority? I will begin with the view that reasons are in no way a function of desire. I will show that this suggestion fails – so we are forced to accept an important aspect of Williams’s view, namely that reasons are conceptually related to antecedent desires. This will motivate my own position which I begin to state in Chapter Two.

1.4 Excursus: Epistemic Authority

Readers will notice that practical reasons do not exhaust the category of normative reasons. For example, there are also epistemic reasons – constituted by evidence that favours having some belief – so mistakes are possible in doxastic as well as practical deliberation.¹⁹ It will help clarify my claims about the relationship between certain normative standards and action by asking whether similar claims can be made about epistemic standards and belief.

Shah claims that deliberation about whether to *believe* p is “transparent” (2003). Agents must regard such deliberation as *the same thing* as deliberation about whether p is *true*. Thus, the aim of truth provides the standard of correctness for such deliberations – we get things right if we form true beliefs. It also “constitutively” governs them – if agents do not take this to be the aim of doxastic

¹⁸ The preceding does not rely on any sharp distinction between moral and non-moral reasons – only on there being some normative standards with the unconditional character described. Standards of this kind may also yield reasons that are apparently non-moral in character (see 3.3 (n64)), or indeed reasons that are not obviously moral or non-moral.

¹⁹ There may be reasons to have certain beliefs that are unrelated to any evidence which supports them, e.g. because having them would be expedient. This is not my topic here since, as we shall see, they cannot figure in *epistemic* reasoning.

deliberation, they lack the conceptual understanding which is needed to engage in such deliberations at all.²⁰ Shah argues that agents (insofar as they engage in doxastic deliberation) immediately recognize that considerations bearing on the truth of *p* settle the question of whether to believe *p*. This does not rule out mistakes about *which* considerations bear on the truth of *p*. But it does rule out the *rational* mistake of failing to recognize the conclusiveness of considerations bearing on the truth of *p* for the question of whether to believe *p*. We may also add, from the recognition that there is conclusive reason to believe *p*, the belief that *p* itself immediately follows – epistemic akrasia is impossible. This is also a product of the claim that insofar as considerations favour having a certain belief, they do so *only* by indicating that the belief is true. So, if an agent is inclined to follow epistemic reasons she acknowledges to be weaker, this is unintelligible: her aim is to ascertain the truth, if indeed she is engaged in epistemic deliberation, and she takes this aim to be better served by following other reasons.²¹ We are forced to say that this person does not recognize the conclusiveness of considerations bearing on the truth of *p* for whether she should believe *p* and so is not engaged in epistemic reasoning. The point again is that in order to make *mistakes* in epistemic reasoning – as epistemic akrasia would be – agents need to engage in such reasoning to begin with (albeit badly). But as Shah correctly observes, someone who does not recognize the conclusiveness of considerations bearing on the truth of *p* for whether to believe *p* is incapable of epistemic reasoning.

The relation of moral reasons – indeed unconditional reasons generally – to action is *weaker* than this relation of epistemic reasons to truth. No clearly defined class of consideration has a monopoly on the considerations which we may regard as bearing on how we should act. By contrast, considerations bearing on the truth of *p* are the only considerations we may regard as relevant to question of whether to believe *p*. We can deliberate practically under the assumption that, given a set of reasons serving competing aims, any one of them might provide sufficient grounds for action. Moreover, this is not qualified by the claim that explicitly moral reasons, when in play, are always recognized as overriding. For example, it is possible to

²⁰ For avoidance of confusion later, Shah does not understand the constitutive relation between belief and truth as having a metaphysical basis – i.e. he does not think mental states necessarily fail to count as belief if they do not aim at truth. I will argue that action has a constitutive aim, but my account does not parallel Shah's views about belief, since I do claim that behaviours which (in the relevant sense) lack this aim fail to count as actions.

²¹ See Raz (2011: 42). Parallel conclusions are avoided in the practical domain because there are practical reasons serving competing aims (more on this below). Thus, it is somewhat intelligible, although irrational nonetheless, for an agent to be motivated by an acknowledged weaker reason which serves a *different* aim than her acknowledged conclusive reason.

ask ourselves whether we should do what is morally required or what is in our ordinary self-interest (i.e. serves some contingent purpose of ours; is not a self-regarding duty). This does, we may think, involve a rational error. We *should* see the consideration that something is morally required as conclusive in practical deliberation when the only countervailing considerations are the ones described. The point is that a failure to comply with this requirement does not lead to the conclusion that someone is not deliberating practically – no more than does simply being ignorant of moral requirements lead to that conclusion. Indeed this failure only counts as a *mistake* because we are deliberating practically. So, it is important to clarify that my thesis about the relationship between moral standards (and probably others) and action does not entail that agents always intend to do what they think these standards demand. That would be false. The phenomenon in need of explanation is just that all agents are *subject* to such reasons, from which it follows that, when conclusive, they will motivate them *insofar as they are rational*. So failures to recognize the significance of explicitly moral considerations, say, are usually just evidence of ignorance or irrationality – rather than evidence that someone is not reasoning at all.²²

1.5 Realism: A Brief Statement

The rest of this chapter focuses on whether the view that reasons are in no way a function of desire – ‘realism’ – can explain why conclusive reasons will motivate us insofar as we are rational. Realism may be motivated by the same concern I expressed about Williams’s account – it cannot accommodate unconditional reasons, but we *do* hold people to have acted wrongly irrespective of contingent facts about their motives. Their response to this difficulty is perhaps the most obvious: we reject Williams’s claim that reasons, and so the facts of when we act wrongly, are conceptually dependent upon desire. That a reason exists (for an agent) is just a fact about the world “outside us”.²³ The explanation why some reasons apply to all agents independently of our individual purposes is just that this world from which stems our reasons is the same world for everyone – and it

²² Whilst I will defend an account of practical reasons according to which they are conceptually related to antecedent desires, I think an analogous claim about epistemic reasons would be false. Epistemic reasons – our evidence – are not dependent upon (e.g.) what we wish were true. I cannot explore why this difference obtains, other than to note that it may have something to do with the other differences between the epistemic and practical domains just discussed.

²³ Reference to the agent may occur in the statement of these facts if her location, competencies, culture, etc., are relevant to her reasons. The crucial point seems to be that these facts are not necessarily linked to how she happens to *think* and *feel* about her place in the world – and it is these, roughly speaking, highly contingent epistemic and motivational tendencies which are irrelevant to the content of her reasons.

contains facts which place like demands on us all. In judging the adequacy of realism as an account of the nature of reasons, we must ask whether it can be reconciled with another dimension of practical authority – which Williams’s account explains – that normative reasons are also, necessarily, potential explanatory reasons.²⁴ My argument *against* realism in 1.6 will be that it cannot provide an explanation why it would ever be irrational for us to be unmoved by reasons – construed as they are by realists – and so cannot explain why normative reasons are also, necessarily, potential explanatory reasons.

It will help to address this objection to a particular philosopher – namely, Raz. Though I will be unable to make the case here, I hope my conclusions generalize. Raz holds that reasons are facts about what is, or would be, valuable – so what we have reason to do is promote or act in otherwise appropriate ways towards things *of value*. These facts are conceptually independent from facts about what we happen *to value* (or facts about our desires). His view is, then, realist in the relevant sense. Nonetheless, Raz accepts that

... normative reasons must be capable of providing an explanation of an action: If that R is a reason to Φ then it must be possible that people Φ for the reason that R and when they do, that explains (is part of an explanation of) their action. (2011: 27)

So, whilst reasons are facts about value that have nothing to do with what we *actually* care about, it is a constraint on some such fact being our reason that it can explain our actions. I have argued that it follows from this that such facts, when they conclusively support an action, would motivate us *insofar as we are rational*.²⁵ This all leaves space for the idea that we do not have reasons to bring about some state of affairs that would – nonetheless – be good when we cannot do so intentionally. This may be, for example, because it is “deeply” physically impossible for us to act on the consideration, or because it is impossible to know that some action could bring about the good outcome. In such cases there is no consideration that favours bringing about that outcome which could explain our action in the way that a normative reason must be capable of – by guiding us towards performing it in light of the justification the consideration provides for doing so. So Raz’s view allows for the possibility that not every fact about value constitutes a reason for us. Thus, it also follows from his view that *even if* every

²⁴ Some views labelled ‘realist’ may reject this requirement.

²⁵ Raz does not mention this specific interpretation of the general requirement (see 1.2). I assume it here.

fact about value could explain our action in the relevant way, the coincidence of all facts about value with reasons would be purely accidental.

1.6 An Argument Against Realism

How could Raz's view that reasons are conceptually independent of desires be consistent with the view that normative reasons are also, necessarily, potential explanatory reasons? Specifically, how is the former view compatible with the claim that conclusive reasons will motivate us insofar as we are rational? Recall that Williams explained this by holding (a) that antecedent motivations provide the starting point, and so partly determine the conclusions of, all rational deliberation, and (b) that reasons necessarily appeal to some motive we can arrive at by deliberating rationally from antecedent desires. It follows that insofar as we deliberate rationally (and also from true non-evaluative beliefs), we will be motivated by whichever of our reasons are conclusive. Raz's view, that reasons are conceptually independent of facts about desires, denies (b). This way of putting things helps clarify why Raz does not think that accepting normative reasons are also, necessarily, potential explanatory reasons leads to Williams's view that reasons are conceptually connected to antecedent motives. Heuer observes, "the answer has to do with his different understanding of our 'rational powers' (2011: 357)". This *must* be so for the following reason. If it were true, in line with Williams's conception of rational powers, that how we can rationally be motivated depends upon our antecedent desires, then the only way it could be necessary that a reason will motivate us insofar as we are rational is if reasons were conceptually linked to antecedent desires. An agent who was unmotivated by some consideration not related to her antecedent desires in the way expressed by (b), would not thereby be irrational. That is because the relevant rational requirement, on this view, is expressed by (a), from which it follows that an agent is rational (in the relevant sense) just if she has desires that are engaged by those considerations which are related, via her rational deliberations, to her antecedent motives. So one way of expressing the grounds for holding that (b) specifies a necessary condition – which Williams also holds to be sufficient – for some consideration to be our reason is this: if (a) is true, considerations which failed to satisfy the condition that (b) appeals to could not explain our actions in the way normative reasons must, because they will not necessarily, when conclusive, motivate us insofar as we are rational. Making this connection between Williams's claim about the nature of reasons and his conception of rational powers allows us to see that although the impossibility of unconditional reasons follows directly from the former – if we do not care (in the relevant sense) about acting morally, say, we

do not have reason to act morally – it is really *the latter* which is driving him to this position. Similarly, by examining Raz’s conception of rational powers we can acquire some insight into why he thinks the realist account of the nature of reasons, rather than (b), suggests itself (i.e. how he thinks it can account for the fact that normative reasons are also, necessarily, potential explanatory reasons).

Raz says the following about the nature of rational powers:

Our rational powers are a general capacity to recognize and respond to facts that make certain responses appropriate, and such facts are reasons because they can be recognized and responded to by our rational powers. (2011: 86)

Raz is drawing upon the idea that normative reasons are also, necessarily, potential explanatory reasons. He is claiming that facts about value are capable of explaining our actions, and hence are normative reasons, if they can be grasped through the exercise of our rational powers. With respect to this subset of facts about value, they could only explain action in the way normative reasons must be capable of if it is possible to recognize the good in question *and* pursue it because it is good. The facts would only be capable of producing such responses in us in virtue of some capacity on our part. Thus, “Reason [is the] universal capacity to recognize reasons, one that in principle enables us to recognize any reason that applies to us, and to respond to it appropriately (Ibid: 86)”. This is, in itself, unhelpful in distinguishing Raz’s conception of rational powers from Williams’s. The preceding statement follows trivially from Williams’s view since he holds that reasons *just are* what we desire to do as the result of rational deliberation from antecedent motives – hence, by reasoning properly (that is one way of exercising our rational powers) we can *in principle* recognize and be motivated by all our reasons.²⁶ So, problems remain about the distinctive features of a conception of rational powers that are needed given a realist account of the nature of reasons.

Williams’s view that reasons are conceptually connected to antecedent motives is driven by the claim that these motives are the starting points for deliberation, and so partly determine its conclusions. Thus, Heuer observes that Williams’s conclusion about the nature of reasons might be avoided if “the starting points of deliberation typically aren’t motives at all... and even when they are, those motives are not beyond rational scrutiny (2011: 358).” If antecedent motives do not necessarily constrain how we can rationally be motivated, then the claim that

²⁶ The force of the ‘in principle’ qualifier is that we could only recognize all our reasons given optimal epistemic conditions.

conclusive reasons will motivate us insofar as we are rational does not force the conclusion that reasons are conceptually linked to antecedent motives. On the contrary, if it is now conceded that these motives can be *obstacles* (rather than starting points) to reason, then the realist view that reasons do not depend (conceptually) upon what we happen to care about seems to be demanded, since irrational motives could not play a role in determining reasons. It also provides an account of how unconditional reasons are possible. If we do not want to be moral, say, or only want to be moral because it serves some other purpose of ours, this would not show that we do not have reason to act morally for its own sake. Rather, the fact we (all) have moral reasons shows that our desires are irrational.

To reemphasize, the realist account of reasons is motivated by a conception of rational powers that, unlike Williams's, allows that we can acquire motives to perform actions (those supported by reasons) that are unrelated to motives we already possess. Is a conception of rational powers which permits this tenable? Williams's account provides a simple explanation why an agent will be motivated by conclusive reasons insofar as she is rational – what she has most reason to do is what she would be motivated to do after subjecting her prior motives to forms of instrumental reasoning. By identifying her conclusive reasons, we identify considerations she is already committed to acting on.²⁷ How does realism explain the connection between reasons and motivation? Suppose reasons are facts about the world that are conceptually independent of desires and that an agent becomes aware of a conclusive reason she has. What guarantees that this fact will motivate her insofar as she is rational? Raz might avail himself of Scanlon's thought that "The practical import of [a reflective state] lies in its distinctive content as a judgement or belief about reasons (2014: 57)." If we can recognize that some fact is a reason, then we can also be motivated accordingly because in understanding considerations *as reasons* we understand them as considerations that direct us to *do* something. There is no space for us to recognize the one thing but not the other – and so a failure to be moved appropriately by reasons either shows we are irrational or we are not reasoning practically to begin with.

There is an argument against realism which has roughly the following form (e.g. Dreier 2015; Dasgupta 2017). We can agree with Scanlon about a feature of

²⁷ Ulrike Heuer has suggested to me that this may be a *problem* for the view. Given that our desires are relatively transparent to us, and if reasons are grounded in desires, is it not hard to explain why we are so fallible when it comes to ascertaining our reasons? I hope I said enough in 1.3 to explain how the fact it is an *idealization* of our desires which grounds reasons creates some space between us and at least some of our reasons.

practical reasoning: that it is unintelligible to recognize there are conclusive reasons to perform some action, but to wonder whether we should do it, since these judgements are identical. If we recognize we should *do* something, we will be motivated to do it on pain of irrationality.²⁸ The objection is, though, that this is something which an account of the nature of reasons should *explain*. What is it in the nature of facts which are reasons which means that when we think a fact is a fact of that kind, we are sometimes rationally compelled to act on it?²⁹ Scanlon, on the other hand, is relying upon this connection between reasons and how we ought to be motivated in order to explain why certain facts will sometimes motivate us insofar as we are rational – they do so because they *are reasons*. I do not think much progress is likely here. The objectors are probably right insofar as it would be *nice* if we could say more about why some facts rather than others are reasons – that is, amongst other things, why they are connected in a certain way to motivation. But these are insufficient grounds on which to reject realism. It may just be that, as many realists accept, the concept of a reason is irreducible. If this is the case, we can do nothing more than affirm that some facts are reasons, and that humans are disposed to be moved by thought about reasons.

My own objection takes a less direct route to challenging realism on its compatibility with the claim that conclusive reasons will motivate us insofar as we are rational. We must consider first how, within the realist framework, agents could be rationally compelled to recognize certain facts as reasons at all. For example, what is wrong with an agent who consciously refrains from making moral judgements – in the sense of judgements about moral *reasons* – because she denies the normativity of the relevant considerations? (The argument can be generalized to an agent's denial, when she has all the non-evaluative facts, of the normativity of any consideration which is actually normative for her.)³⁰ It will help to first consider Williams's response. Recall there is a complication about whether this agent *is* actually wrong given Williams's approach. But let us assume she is wrong in what she says insofar as there are moral reasons *for her*. This must be because the relevant considerations are directed towards desires she would have if she subjected her present motives to a process of rational deliberation. When this agent denies the normativity of the relevant considerations (for herself) she is making the following kind of mistake: she is wondering why she should care about

²⁸ We may still fail to do it, and this may or may not be a rational failure. But motivation is at least a necessary step towards doing what we are rationally required to do.

²⁹ Williams would answer this by claiming that the facts in question concern something the agent is committed to caring about in virtue of her existing motives.

³⁰ I am not yet *arguing* that the skeptic should accept she goes wrong with respect to any class of considerations *in particular*.

something which she is already committed to caring about given her present desires. To be clear, she might wonder whether she ought to care about what she currently cares about. That would simply involve, on Williams's view, subjecting them to the process of instrumental deliberation by which we determine our reasons. The source of confusion here is that the consideration whose normativity the agent questions is by hypothesis something which that process, after which no normative questions remain, would find to be normative.

The story realists tell here cannot, like Williams's, appeal to the agent's desires – they claim these have nothing to do with whether the denial that certain considerations are normative is true. This point is connected to the realist's conception of rational powers. Since reasons are not grounded in motives, and since conclusive reasons will motivate us insofar as we are rational, *rational* motivation cannot be constrained by what we happen to want. It must allow us to go "outside ourselves" to determine what the reasons are, and it is this recognition which will motivate us. Thus, the natural thing for realists to say about someone who rejects the normativity of genuinely normative considerations is just that they have failed to exercise their rational powers (whatever sensitivities and competencies this ultimately involves) in such a way that they lack knowledge of their reasons.

But how exactly should we understand this lack of knowledge? In many cases, ignorance about the existence of moral reasons is the right explanation why an agent is not motivated to act morally. The crucial feature of these cases is, however, that the agent is ignorant of the reason because she is ignorant of other salient features of the case (e.g. that her action hurts someone, or that it falls under an act type which she knows is generally prohibited). Simple ignorance is also how we would diagnose the mistake of someone who attempts to question the normativity of morality, but who really has no idea in what morality consists (e.g. that, if common sense is right, lying, stealing, and assault are *prima facie* immoral). Now, the first scenario is clearly not the position of an agent who denies the normativity of morality since it is the agent's acceptance of moral reasons generally which makes ignorance plausible as an explanation of her failure to be motivated to act morally in this case. As to the second, we can simply stipulate that the knowledge that is lacking here is not lacking in our example of an agent who denies the normativity of morality. In fact, I propose we go further and stipulate that she knows all the morally significant features of some concrete case (performing the action will cause bodily harm to someone, it will go against her

expressed wishes – however we want to cash this out³¹) and that she denies *their* normativity. If I am right about the resources that are available to explain mistakes in identifying reasons given the realist’s conception of rational powers, then all we can say here is that what the agent misses *just is* what she denies: that these considerations are reasons. But that – and this is the heart of the objection – mischaracterizes her error. Her problem is not that she lacks knowledge – she already knows everything that could make a difference to whether she has moral reasons. Rather, her failure to be motivated by these considerations is itself a *rational* failure.³² We are confused as to how someone can know all this about the case, and yet be unmoved. That she does not know the considerations are *reasons* does nothing to remove this confusion. On the contrary, the idea that such knowledge would alter her motivation seems repugnant – what *should* motivate her are the considerations she already acknowledges.³³

Realists cannot do justice to this insight. Of course, they can say it is irrational for someone to be unmotivated by these considerations just in the sense that their motives do not line up with the facts about how we should be motivated. But, because these facts are themselves not grounded in motives, they then have to explain how they motivate by appealing to the nature of our judgements about them – that they are judgements about reasons. Thus, what is *really* irrational for realists is failing to be motivated appropriately by judgements about reasons. Since this does not apply in the case at hand, the more salient explanation of the agent’s error is that she does not believe she has certain reasons. Perhaps someone might reply that an agent could not, as I have suggested, know all the morally significant features of a case and yet fail to see she has moral reasons. But this does not address the problem. If that claim were true, we assume it is so only contingently, absent some further argument. The realist still places all the weight of the explanation why it is sometimes rationally required to be motivated by certain considerations on the idea that we correctly judge them to be reasons. What I have argued is that the connection between the presence of certain considerations, e.g. those we sometimes call ‘moral’, and the rationality of being motivated by them is necessary, so does not depend on any such contingency.

³¹ So long as it does follow conceptually from our description that there are reasons to act one way or another, e.g. the action would violate her right. I thank Ulrike Heuer for reminding me of this.

³² Notice, Williams *can* say this in those cases where he agrees the agent has moral reasons. See n29.

³³ This point bears similarities to Smith’s discussion of “moral fetishism” (1994: 71-76).

On these grounds I reject realism. In short, realists can only make sense of the idea that normative reasons are also, necessarily, potential explanatory reasons – and so, when conclusive, will motivate us insofar as we are rational – is by relying on the idea that rational agents correctly judge certain considerations to count in favour of certain actions. But this idea is ill equipped to explain why failures to be moved by certain considerations *as such* (i.e. those which are reasons), and not just failures to be moved by certain modes of thought about them, is itself sometimes irrational. So, realism cannot in the end explain why conclusive reasons will motivate us insofar as we are rational, and thus cannot explain how we have reasons at all. Much less why some reasons apply to all agents.

Constitutivism and Its Critics

I have argued that neither realism nor certain ‘desire-based’ views explain practical authority. The former because it fails to account for why conclusive reasons will motivate us insofar as we are rational. The latter because they cannot accommodate unconditional reasons. ‘Constitutivism’ offers an elegant solution to these difficulties, explaining both features of being subject to reasons. (Although, at this stage, the explanation will only be conditional in form: *if* there is a constitutive aim of agency, this explains practical authority.) I will first describe constitutivism and clarify it in view of some initial worries. In 2.2 I will outline the constitutivist account of practical authority. I will explore an objection to this explanation and respond in the remainder of the chapter.

2.1 Constitutivism: An Introduction

‘Constitutivism’ is the view that practical reasons are grounded in constitutive standards of action – the features of an event which make it an action. The idea that action has “constitutive standards” may be clarified with an analogy from Korsgaard (1999: 14-15). We can praise or criticize many things about a house. For example, we can praise its beautiful hanging baskets and ornaments – and say, “This is a great house!”. But a house which conforms or fails to conform to the standards appealed to in this judgement is not any *more or less of a house* as a result. A house which did not have beautiful hanging baskets or ornaments could still be as much of a house as any other. But *if* the concept house is governed by constitutive standards, there are also standards of another kind. A house which fails to conform to them is less of a house as a result. That is, if houses have constitutive standards, there is something an object must be or do otherwise it is not a house. If sheltering inhabitants, say, is *constitutive* of houses, then when we say a house is good because it does this well, the sense of ‘good’ merely indicates that it meets the criteria for being a house. To be clear, the claim is *not* that if we want a good house it must shelter its inhabitants, but if we just want any old house it need not. Rather, there is one sense of being a “good” or “bad” house (amongst others), where this is simply a measure of the extent to which it counts as a house at all. That is, a measure of how well it does the thing, which if it did not do *at all* –

as opposed to doing it very badly – the concept house would cease to apply. The claim that action has constitutive standards is structurally identical: an action must conform to them just in virtue of being an action – and to the degree it does, it is more of an action.

The second step is to argue that *if* there were constitutive standards of action, they would be a source of reasons. If there are standards that actions must conform to insofar as they are actions, then agents cannot – at least not coherently – form intentions to act in ways contrary to these standards. The possibility of forming intentions itself presupposes that intentions are, to some extent, formed in accordance with them. Since, at the very least, what it is to be an agent is to form and implement intentions, it follows that someone to whom these standards are alien is not *an agent* capable of intentional action. Thus, we can equally speak of constitutive standards *of agency*. So far, this is just to say they provide a special kind of evaluative measure for actions – just as the standard of sheltering inhabitants might do for houses. The difference is that forming intentions is a self-conscious activity. We can always, in principle, plan our actions reflectively in response to standards we believe apply to us.³⁴ Because this is so, we say our actions are done for reasons – which may be genuine reasons if the standards to which our actions respond really apply to them. Thus, if there were a constitutive aim of action, yielding its constitutive standards, it would be a source of reasons. Insofar as agents modified their actions in light of it, they would act for reasons, just as when they respond to any other standard which they believe, rightly or wrongly, applies to them. But crucially, if we could show that such standards exist, their normativity could not be in doubt. Their genuine applicability to action is already contained in the idea that they are constitutive of it. Compare how other standards we have discussed *may* give rise to reasons. If I intend to build a house, and there are constitutive standards of houses, then I must build something that that conforms to these standards, otherwise my intention will not be effective. Building a house would, minimally, *just be* to build a structure that conforms to these standards. But do constitutive standards of houses give me reasons to build something which conforms to them? No – I have reasons to do so only if I have a reason to build a house, and this has not been settled. But with the constitutive

³⁴ The claim is *not* that actions which are attributable to our agency in this strong sense must result from deliberation. We often recognize (or think we recognize) reasons immediately. But even when we act on such judgements without reflecting further upon how these considerations justify our action, we *could* have done so. This is distinctive of rational action.

aim of action there is no further question whether they give me reasons on some occasion. They apply, necessarily, in the formation of every intention.

I will now consider some objections.³⁵ We might think: “Surely you mean that *good* actions are done for reasons which derive from certain aims (good ones), and so actions which are not done for those reasons are not good actions. But it is still an action, right?”. This is wrong: constitutivists claim there is a sense of ‘good’ such that ‘an action’ and ‘a good action’ *mean the same thing*, since the concept is misapplied to something which does not conform to the relevant standards. This should be qualified. Bad actions, which have the constitutive aim of action but fail to conform to many of its constitutive norms, are still actions but are (obviously) not good actions. A bad action is however *less* of an action than a good action – this is just what it means for them to be “good” and “bad” respectively. This is the force of the original point.

But we might deny that the relevant sense of ‘good’ exists in relation to the concept action. It will help to consider the analogous case. It seems ‘house’ is not obviously misapplied to structures which (say) do not shelter inhabitants *at all*, e.g. they are built without roofs. We might think houses only *fail* in respect of not sheltering inhabitants *if that is our purpose in building them*. But if contingent purposes provide an external standard by which we judge the products of our actions in cases like this, then why think those purposes are, as constitutivists argue, themselves subject to *constitutive* standards which determine our reasons for having them? In response, I simply concede the charge with respect to houses. It makes sense to appeal to contingent goals in assessing whether the fact houses do not perform certain functions represents a failure – the only kind of failure which seems possible here is a failure to give effect to our (rational) intentions. In which case the concept house is *not* governed by constitutive standards. But the example was only meant to illustrate how such standards *would* differ from other standards we appeal to in our evaluations of houses – and many other things besides. The argument that action *does* have constitutive standards will have to wait until 3.4, but we must emphasize that it rests in no way upon how other concepts operate. We have seen that because we sometimes produce actions for contingent purposes there is a mode of assessment for actions in terms of whether they succeed or fail to manifest such intentions. But this is compatible, with their *also* – necessarily – serving the constitutive aim of action, which could not vary from agent to agent or from action to action.

³⁵ I thank Ulrike Heuer for pressing me on these issues.

Finally, we might deny, in general, that what is constitutive of a kind of thing provides an evaluative measure for things of that kind – let alone a normative standard in the case of action. The objection is that we can only understand constitutive features of a kind of thing as a *descriptive* criterion which something must satisfy in order to fall under a given concept. Perhaps if a house has to be anything in order to be a house, it has to be three-dimensional. But the language of a ‘good’ house is strained if we are referring to this feature of houses alone. If this picture generalizes, then constitutivists would have “smuggled” normativity into their account of action by wrongly supposing that constitutive features of a kind of thing provide the things which fall under the kind with an evaluative standard. But this picture does not generalize. Take the concept *rainforest*. If a rainforest has been deforested such that it can no longer sustain itself, there is something bad about this – besides any moral failure of the deforesters. In failing to sustain itself, the rainforest fails, to some degree, to be the kind of thing it is – which is just a *rainforest*. This does not reflect human *purposes* for rainforests – but rather that we understand the concept rainforest through the good case of a fully self-sustaining ecosystem, and so other cases are judged to be deviant. The problem with the previous example is that something is either three-dimensional or it is not – and how can something be assessed *as a three-dimensional object* if it is not three-dimensional? This is why if three-dimensionality were all that is demanded of a kind of thing constitutively, this would not be an evaluative standard of things of that kind *qua* things of that kind. What the rainforest example shows is that some features of a thing which are necessary for the application of a concept can be found in degrees. Rainforests may be self-sustaining ecosystems even if some rainforests are not very good at sustaining themselves. We find this “space for failure” more generally in functionally organized things in the natural world. If some other concepts work this way – being governed by constitutive standards – this is a further reason not to discount (at the outset) the hypothesis for the concept action.

2.2 The Constitutivist Account of Practical Authority

Some qualifications are needed before outlining the constitutivist explanation of practical authority. Part of what we want to explain is how unconditional reasons – considerations to which nobody can rationally be indifferent – are possible. In 1.3, I suggested *moral* reasons provide the clearest case of this. Thus, in describing the constitutivist explanation of practical authority I will speak freely as if the constitutive aim of action grounds moral reasons. But at this stage, I will not attempt to *prove* anything of the kind. This is because I will not have established that *any* aim is constitutive of action. The latter project will be postponed until

Chapter Three – where I will also discuss its possible moral implications. The conclusion I am defending here is: *if there is a constitutive aim of action*, this would explain practical authority.

There is a further aspect of the constitutivist framework which I take for granted. (This will be important for dealing with objections later.) What I have said so far may suggest that constitutivism is only committed to the following: constitutive standards of agency are *one* source of reasons; ordinary contingent desires and, even, facts which are conceptually independent of desires *may* be another. This is *not* how I conceive the framework: constitutive standards of action are *the* source of reasons. In 1.6 I rejected as incoherent the idea that facts which are conceptually independent of desires could constitute reasons. Thus, the constitutive aim of action is itself conceived as being grounded in prior motives: it is an aim we are rationally committed to in virtue of having *any* set of desires. What then of ordinary contingent desires? I noted in 1.3 that these cannot be the grounds of unconditional reasons. But if this burden is taken up by the constitutive aim of action, might they still be the grounds of reasons that apply to some agent or group of agents only? In fact I deny it makes sense that ordinary contingent desires *as such* could constitute reasons – rather they only do so, when they do, in virtue of their relation to the constitutive aim of action. The argument for this claim must wait until 3.4. Thus, in what remains of this chapter, ‘constitutive standards of action’ will refer to (a) any demand that we *do* something which follows directly from the constitutive aim of action (b) any requirement that some things *not* be done which also follows directly from the aim *and* (c) any reason generated by those ordinary contingent desires which relate to the aim in some (as yet unspecified) way. So understood, we can define ‘constitutivism’ as the view that constitutive standards of agency are *the* source of reasons.

The explanation of practical authority that constitutivism yields is this. We may again frame our discussion in terms of someone who denies the normativity of (as she sees them) “moral” standards. Moral reasons apply, necessarily, to all agents. This means that agents will be motivated by moral considerations (assuming they are conclusive) insofar as they are rational – indeed directly, and not in virtue of some ulterior motive.³⁶ The skeptic is then irrational *both* because she will not be so motivated, and because she denies she has reasons she in fact has. How does constitutivism interact with these claims? Constitutivism agrees with Williams

³⁶ I include in such ulterior motives the fact that moral considerations *are reasons*. This was the crux of my objection to realism.

about the nature of rational powers. The starting points of deliberation are our antecedent desires, and these partly determine its conclusions. Thus, the explanation of how an agent could be irrational for failing to be motivated by moral considerations, when she knows all the non-evaluative facts of the case, is also the same: she must fail to care (enough) about something which she is already committed to caring about (to that degree) given her present desires. The problem with Williams's approach is that agents may never be irrational for failing to be motivated by "moral" considerations, which means those considerations may fail to be genuinely normative. He holds that we possess our existing motives – and are rationally committed to certain other motives in virtue of them – only contingently. There can be no guarantee that "moral" standards will give us reasons – and thus Williams must reject practical authority. Constitutivism avoids this conclusion. If moral standards *just are* constitutive standards of action then in forming intentions at all we are necessarily committed to the normativity of moral considerations. This means that the desires which moral considerations are directed towards must be retrievable from *any* set of desires whatsoever by subjecting them to a process of rational deliberation. It follows that *all* agents will see themselves as having moral reasons and will be motivated by them (when they are conclusive³⁷) insofar as they are rational. (They are quite literally *less* of a rational agent if not.) I argued earlier that if a consideration appeals to constitutive standards of agency this is *sufficient* for its being a reason. This is partially reinforced by the argument that those considerations satisfy a necessary condition for being a reason: we will be motivated by them, if they are conclusive, insofar as we are rational. This is because they are directed towards aims we are, as a condition of our agency, committed to. Thus, "Why be "moral"?" – where this question challenges the normativity of "moral" considerations – is incoherent. Any agent who asks it *is already committed* (and rightly so) to taking moral standards into account – since even the (defective) skeptical question represents an attempt at deliberating practically. This is just to say that all agents, necessarily, are subject to moral reasons. The skeptic's *attempt* at normative reasoning shows she is an agent; her failure to take moral considerations into account shows she is *irrational*.

This last remark is the kernel of a response to the following objection: constitutivism undermines the possibility of "defective" – including perhaps *immoral* – action. If acting morally, say, is constitutive of agency, then how can we *act* immorally? If the impossibility of immoral action followed from constitutivism, this

³⁷ We might wonder how reasons which can be established directly by the constitutive aim of agency could *fail* to be conclusive, since it is *the* source of reasons. I deal with this in 3.5.

would be troubling on at least three counts. First, agents obviously do act immorally. Second, “immoral” people would not be criticisable for failing to be moral. Moral standards, I have argued, only give us reasons because they are constitutive standards of action and we are beings who act. But acting is not, the objection goes, something which an “immoral” person does – so they do not have moral reasons. Criticism of behaviour is illegitimate in the absence of reasons to behave otherwise. Finally, if acting morally were inevitable (if we act at all) then moral standards would not be normative even *for agents*.³⁸ As Silverstein says:

... if I am falling from a great height and cannot arrest my descent, questions about whether I *should* be falling are inapt, to say the least. More generally, when there are no alternatives, normative thoughts seem out of place. (2015: 1131-32)³⁹

Silverstein is correct in what he says. But the kind of inevitability which governs the (constitutive) relation of agents to moral standards is not analogous to the inevitability of a physical object conforming to laws of gravity. It is not the kind of inevitability which would entail that all agents actually conform to moral standards – thus the objection misfires. Constitutivists only claim that *being subject* to moral standards is inevitable for agents because agency itself *commits* them to taking those standards into account. Whether they do or not is a separate question. Hence, the many possibilities for, and degrees of, failure reflected in someone’s trying *badly* to be an agent – ignorance of their reasons, akrasia, and so on – are (so-to-speak) the “alternatives” which are necessary for the constitutive features of agency to be normative.⁴⁰ This may seem strange. How can something be constitutive of an activity when we cannot use it as a criterion for determining whether someone is performing that activity (since what I have said amounts to the claim that someone can act without adhering to constitutive norms of agency)? What then stops us from saying that the actions of non-human animals are also (constitutively) subject to moral standards? The answer is that whilst the idea that people are committed to these standards does not require that they recognize or follow them, it does rely (for any plausibility it has) on people doing something which resembles *an attempt* at performing the activity of which they are constitutive. That is, *acting rationally* – and we cannot intelligibly say that non-human animals try to do this. Though the point will not resonate until 3.3, when I

³⁸ I thank Klara Andersson for first bringing this to my attention.

³⁹ This point is closely related to one made earlier (2.1). When some constitutive feature of a kind of thing either is or is not present (e.g. three-dimensionality) it cannot provide an evaluative measure for things of that kind *qua* things of that kind.

⁴⁰ See Korsgaard (2009: 159-76) for related discussion. Silverstein seems to think that the alternative that is required is the possibility that *not being an agent* is a rational option. Whilst I agree it can be, I do not think this is essential to resolving the current problem.

discuss a particular aim as constitutive of agency, someone who did not ever conform to constitutive standards of agency would be unrecognizable as an agent, given our pre-theoretical grasp of what agency involves.⁴¹

2.3 “Why be an agent?”

I have argued that constitutivism explains why the “Why be “moral?”” question is incoherent by showing how moral considerations can be unconditional reasons. But some argue that even if (because agency has constitutive standards) the question “Why be “moral?”” is incoherent when asked by someone who wants to be (and perhaps has a reason to be) an agent, we *can* understand the question “Why be an agent?”, and so constitutive standards of agency (including, possibly, moral standards) have merely conditional force (Enoch 2006, 2011). That is, the imperatives they establish have the form “If you should be an agent, then you should do *x*”. So, following constitutive standards of agency would only be demanded, necessarily, of agents if the antecedent of this statement is necessarily satisfied. But we can doubt whether it is. In answering these doubts, we cannot appeal to constitutive norms of agency – as constitutivists do to account for *all our reasons* – because those are the norms whose authority is challenged. Thus, constitutivism cannot provide a full explanation of how agents, necessarily, have reason to conform to these norms. But this objection fails because it misunderstands constitutivism. I will explain this presently. We must keep two questions in view. First, is the question “Why be an agent?” intelligible? If not, the challenge is unsuccessful. It will not have been shown that a further rational basis for conforming to constitutive norms of agency is needed, since it will not even make sense to ask for one. But even if it does make sense, we still have to ask whether the justification of agency must be provided by something other than constitutive norms of agency. Only if it does will it have been shown that a further rational basis for conforming to constitutive norms of agency is needed.

2.4 Normative Inescapability

Some critics have thought the question “Why be an agent?” is unintelligible because agency is (in some sense) *inescapable* (Ferrero 2009; Katsafanas 2013:

⁴¹ To use an example from earlier, the difference between applying the concept agent to someone who is trying and failing to form and implement intentions and applying it to the sort of being just discussed, is like the difference between applying the concept rainforest to the rapidly declining Amazon and applying it to a desert.

47-67; Velleman 2009: 115-58).⁴² That is, it would only make sense to reflect upon whether I should be an agent if there is some possibility I might not be one. In response, it might be argued that agency is *not* (in the relevant sense) inescapable. It might be argued further that the mere intelligibility of the question shows that constitutive norms of agency cannot, by themselves, ground unconditional reasons. In what follows I reject every one of these claims. I argue that constitutivists are entitled to claim that agency is inescapable and that this explains why constitutive norms of agency, by themselves, ground unconditional reasons. But this explanation does not depend on the claim that the question “Why be an agent?” is unintelligible. The sense in which agency is inescapable is compatible with the intelligibility of this question – and that it *is* intelligible is plausible in any case.

What does the claim that agency is “inescapable” mean? It might mean we are at all times *active* in the role of agent – forming intentions and acting. But when I am asleep, such activity is impossible – and this does not mean I no longer exist. If this is what the claim that agency is inescapable means, it is false. Worse, this observation may embolden those who claim that a further rational basis for conforming to constitutive norms of agency is needed. I can exist without being subject to those norms: when asleep I am not capable of having aims, including the constitutive aim of agency, and if I am not an agent (however temporarily), then norms of agency are not *my* norms. What is troubling here is not simply that when asleep I have no reasons. Rather, the problem is that if being subject to constitutive norms of agency is what gives me reasons, and I could simply *choose* not to be subject to them by bringing it about that I am asleep (e.g. by taking sleeping pills), then it seems like it would be valid to make that choice on the basis that the demands of agency do not suit my interests on a particular occasion. But it does not seem valid to opt out of (say) moral reasons on these grounds – they apply to us whatever our interests – so it follows that constitutivism cannot account for the possibility of such unconditional reasons.

We might respond that agency is inescapable in a different sense: we cannot *coherently* choose not to be agents. This could be developed into an argument that the question “Why be an agent?” is unintelligible in the following way.

⁴² The following is indebted to Silverstein’s conclusion that “There are no standpoints outside of agency where questions about one’s reasons are apt (or even intelligible). And thus, there are no normative questions about agency... left unanswered by agency’s constitutive norms (2015: 1141).” My decision to make this point, unlike Silverstein, in terms of “inescapability” is merely a presentational difference.

Normative questions structure our deliberations – they demand justification for options that are available to us and in this way guide our choice between them. But if we cannot choose to stop being agents, we cannot really choose to *be* agents either – there is no alternative. Therefore, the question “Why be an agent?” could not structure deliberation in the way normative questions do, since there is no manifold of options for it to guide choice between. It is therefore unintelligible. But *why* can we not coherently choose to stop being agents? The best support for this claim lies in the fact that choice is an intentional activity and thus itself governed by constitutive standards of agency. But as it stands this suggestion only highlights something trivial. The fact I am governed by norms of agency (constitutive or otherwise) whilst I am still an agent merely reflects what it means for some standard to be genuinely normative for some activity. As long as I am playing chess, I can decide not to follow the rules, but this just identifies me as a cheat.⁴³ But nobody thinks I cannot choose to stop playing chess. So the initial claim has to be developed further: in order to show that leaving agency is not a coherent option, we need to show there is something especially problematic about how that intention (whose formation is governed by constitutive standards of agency) relates to the actions (which are also governed by those standards) that would be needed to bring about its fulfilment. Maybe the problem is that whilst I can stop playing chess “just like that” – without performing any moves within chess that are subject to its norms – the same cannot be said for agency. I have to “be an agent for a bit” in order to stop being one. There is something awkward about this – especially if I want to leave agency *because* I reject its rules as, say, silly or malevolent. In order to be outside these constraints, I first have to do something that depends for its success or failure – *qua* thing done – on the very rules I reject. Nonetheless, it overstates things to say this involves the kind of inconsistency or confusion which would mean leaving agency is not an option. On the contrary, cases with this structure are familiar from everyday life, when we would say that genuine choice is exercised. Suppose a prisoner seeks her freedom by appearing in front of the parole board, even though, and *because*, she regards prison rules as unjustified. To be clear, she is not *pretending* to go along with the procedures of the parole board (which are part of the prison rules). She really feels remorse about her crimes and is able to express this to their satisfaction. She just does not think this “panel of suits” ought to have the powers it does. I think this is analogous to someone who intends to leave agency because she rejects its rules. Since the

⁴³ There is a fine line between wildly deviant attempts at participating in some activity and *pretending* to participate in it.

prisoner can choose to seek her freedom in the way described, it follows that agents can choose to exit agency by exercising their agency.⁴⁴

If my responses are successful, the preceding arguments, based upon the claim that agency is inescapable, do not show that the question “Why be an agent?” is unintelligible. I will now argue that it *is* intelligible. Consider Korsgaard’s remark that “Human beings are *condemned* to choice and action (2009: 1).”⁴⁵ Her word choice is revealing – it suggests we can regard our agency as regrettable. It is possible we could be condemned to our agency, by experiencing the suffering it brings, without being able to identify it as the source of our suffering and thus as something regrettable. But this cannot be Korsgaard’s point: *she* has identified agency as something to which we are condemned. That such beliefs are possible is obvious – that they can sometimes seem *rational* is also plausible. It makes sense to lament the burden of responsibility for our actions – having to reflect on their worthiness and experiencing the guilt that comes with (inevitably) failing to do all that we judge worthy. It makes sense to find this condition regrettable, even if we have no intention to leave it for something else – perhaps the life of an animal whose ends are simply given to it, and for whom the question of their value does not arise. The point is that this would *not* make sense if we could not coherently entertain any other possibility than being agents. If we really could not imagine anything else, we would be in a position not much different from someone who could not identify their agency as the source of their suffering. We could experience agency as painful, but we could not form complex emotions such as *regret* about it. That emotion depends upon the idea that things could (conceivably) have been different. This suffices for the intelligibility of the question “Why be an agent?”. If we have multiple possibilities we can regard as better or worse than each other – which is again implied by our regret – we have the resources we need for normative enquiry.

Perhaps I have missed a salient point: we cannot *actually* pursue alternatives to agency. We cannot become, for example, non-human animals. This seems to support the view that the question “Why be an agent?” is unintelligible because of a

⁴⁴ There are still constraints. If I intend to end my life, it would be incoherent for this to be based on the desire that I not be subject to constitutive norms of agency – since if I end my life, ‘I’ does not refer, and there is no one such that constitutive norms of agency do not apply to them. Minimally coherent desires here would have the negation take wide scope: I want *not* that I be subject to constitutive norms of agency. This desire is fulfilled either if constitutive norms of agency cease to apply to me *or* if (as is my intention) I cease to exist and so there is no one to whom constitutive norms of agency could apply. I thank Fiona Leigh and Naomi Alderson for urging me to be clearer about this.

⁴⁵ Emphasis original.

claim I made earlier. That is, in order that I have a reason, it must be possible for it to explain my action. This requires that performance of the action for which it is a reason is physically possible, which becoming a non-human animal is not. Crucially, if there cannot be reasons to pursue alternatives to agency, because doing so is “deeply” physically impossible, then it cannot make sense to ask whether to pursue them. Without a manifold of options to guide choice between – which would require that there are reasons for those options – the question “Why be an agent?” is unintelligible. But the premise which is needed to support this conclusion is false. Whether or not I can become a non-human animal, there are conceivable alternatives to agency: I can think it would be better if I was dead or asleep, and I can actually kill myself or take sleeping pills. The objector might weaken her claim. Restricting ourselves to the case of suicide, even if this course of action is possible, there is also a familiar sense in which many people could not “bring themselves to do it”. Perhaps this is enough to deny they have reasons to pursue this option.⁴⁶ The thought is that we could only presently have reasons to pursue alternatives to agency (or indeed any other course of action), if we can *now* be rationally motivated to pursue them. But this goes beyond any plausible restriction that the possibility (or not) of performing an action places on whether we could have reason to perform it. Whether we have reasons to perform an action does not depend on whether we could now be rationally motivated to perform it, but only on whether (insofar as we are rational) we would be motivated to perform it if (we suppose) these were conclusive reasons to do so. Having a reason to perform some action does not even depend on it *now* being likely that I could successfully perform it (though it must be not be “deeply” physically impossible). Suppose I am playing chess and someone has tied me to my chair. All I can *actually* do is continue playing. If I ask, “Why play chess?” and someone responds, “You just have to!”, their point is not that it lacks meaning to ask whether a life beyond chess playing would be better than this – whether it would be a life which I have reason to live – but rather that I should not bother thinking about this in my predicament. I conclude that the present objection fails, and thus that regret about agency is evidence *for* the intelligibility of the question “Why be an agent?”.

If the original objection that a further rational basis for conforming to constitutive norms of agency is needed fails, this is not because the question “Why be an agent?” is unintelligible. But there is another reason why the objection fails. A

⁴⁶ I ignore the complication that those who *could* bring themselves to pursue this or any other alternative to agency might have reasons to exit agency – so this still would not show that the question “Why be an agent?” is *always* unintelligible and so that constitutive norms of agency are unconditionally authoritative.

helpful way to proceed is by noticing that a criterion for the intelligibility of a question is that it could be answered, even if we do not know its answer. “Why be an agent?” is a *normative* question – it asks for considerations for and against pursuing certain courses of action (in this case, those that involve remaining an agent). Thus, normative questions are answered by giving *reasons*. But what *are* practical reasons according to constitutivists? They are directives to promote the constitutive aim of agency. So, whether you should be an agent (or do anything at all) is determined by appealing to constitutive standards of agency. To capture this idea, let us say that agency *is* inescapable in a sense not discussed so far. Call it ‘*normative inescapability*’: agency’s constitutive standards are *the* source of reasons.⁴⁷ If being an agent promotes the constitutive aim of agency, we should be agents. If it does not, we should not. Agency is normatively inescapable because we can only engage in intentional activity if we are committed to the constitutive aim of agency. If we are committed to this aim, we must be trying to conform to constitutive norms of agency – which *just is* what we have reason to do. Thus, we can only demand reasons – an intentional activity – if we are already committed to the constitutive aim of agency.

We can now answer the charge of circularity against constitutivism. The claim was that the intelligibility of the question “Why be an agent?” shows we have reason to do what conforms to constitutive norms of agency only if we have reason to be agents. Since constitutivists think all reasons are grounded in constitutive norms of agency, all that can be appealed to in showing we always have reason to conform to them is themselves. It is true that if agency is something which can be justified or not (which it must be if the question “Why be an agent?” is intelligible) it would have to justify itself. That is because we cannot settle whether it (or anything else) is justified without appealing to constitutive standards of agency. But why think this shows that constitutivism lacks the resources to vindicate the idea that we necessarily have reason to conform to constitutive norms of agency? Here is another suggestion: this shows that whether we necessarily have these reasons *was never at stake* in the question of whether we should be agents. Clearly, if it were, my defence of the view that regret about agency can be rational suggests the objection to constitutivism is sound – since I acknowledged there can be circumstances in which we should not be agents.⁴⁸ But the explanation why we always have reason to conform to constitutive norms of agency is not that “You should be an agent, and so you should have the constitutive aim of agency, and do

⁴⁷ De Maagt uses this terminology for the claim that agents necessarily have a reason to be agents (2019: 5). This should be severely distinguished from the meaning I assign it.

⁴⁸ See also 3.5.

what promotes that aim”; it is rather that some aim is constitutive of agency and hence establishes which aims we have reason to promote – including itself. It follows that anyone who is, as a matter of fact, an agent is *necessarily* committed to the constitutive aim of agency, and so should do what promotes that aim – even if it sometimes follows from this that they should be not be agents.

2.5 Justification Without Agents’ Reasons?

So, the question “Why be an agent?” does not establish that a further rational basis for conforming to constitutive norms of agency is needed, because *the answer* to it must itself be derived from constitutive norms of agency. The only rational considerations available to us are those that appeal to constitutive norms of agency – so whether we have reason to be agents is beside the point of whether constitutive norms of agency give us reasons. But an objection remains. If the only normative questions agents can sensibly entertain are asked and answered by reference to constitutive norms of agency, then this is how *we* must understand the question “Why be agent?”. But why assume this is how *anyone* asking the question must understand it? I have assumed the question is posed from the perspective of agency, which entails commitment to its constitutive aim in all intentional activity. It is only because of this that when we demand reasons we must be interpreted as demanding considerations that appeal to constitutive norms of agency. To state the objection it will help to adopt the terminology of ‘*agents’ reasons*’ for those considerations which appeal to constitutive standards of agency and which we have argued are genuinely normative for agents. The thought is, if we do *not* assume that the question “Why be an agent?” is posed from the perspective of agency then it need not be concerned with agents’ reasons at all. We suppose this is the case only because there is nothing else *agents* can appeal to by way of rational considerations, and so no other way of framing normative questions except in terms of them. But the person who now asks the question is not bound by this restriction – at least we have been given no reason to suppose this – because she occupies a perspective outside agency. Constitutive norms of *agency* cannot be, in the first instance, where she looks for rational considerations, because constitutive standards only apply to things which belong to the kind of which they are constitutive (and she does not belong to the kind *agent* because she is not committed its constitutive aim). What this thought experiment is meant to show is that there could be a challenge to the authority of constitutive norms of agency which has to be answered *without* reference to the norms themselves. Since constitutivism cannot give such an answer, this is meant to show that it cannot vindicate the idea that constitutive standards of agency are unconditionally

normative. That is because, whatever might be true from within the perspective of agency, its constitutive norms may or may not be justified from perspectives outside it. Hence, whether those norms give us reasons *does* ultimately depend upon whether we should be agents, since it seems agents' reasons may not, ultimately, provide rational grounds on which to act.

Silverstein argues that constitutivists are entitled to reject this interpretation of the question as unintelligible, so the explanation of practical authority given in 2.2 is not undermined:

If agency's constitutive condition is what explains why agents have the various interesting features they do... at best he [who poses the question in this form] would resemble Harry Frankfurt's wanton: he would have "no identity apart from his desires" and would be "no different from an animal".⁴⁹ He would, in other words, be incapable of intentional or autonomous action. (2015: 1137)

Silverstein's point is that assuming *we* could pose normative questions without being committed to the constitutive aim of agency begs the question against constitutivists who claim precisely that this commitment *just is* the essential mark of our intentional activities. So if the objection depends upon the idea that we can understand the challenge to the authority of constitutive standards of agency, then it cannot claim to accept constitutivism for the purposes of argument. This is correct – but the objection could be made without the offending assumption. It could grant to constitutivists that their account of intentional activity is the right one *for us*, and that constitutive norms of agency appear to us to yield unconditional reasons, because we can only deliberate from our own perspective. But it could still claim these norms do not *actually* ground such reasons because there is a more minimal account of intentional activity that is conceivable, and for which there are no considerations that appear unconditionally normative from within that perspective. A simple belief-desire account of intentional activity, and a purely instrumental account of practical reason would suffice here – we admit only the features which are necessary so that we could conceivably recognize activity as intentional.⁵⁰ That such accounts have been argued about is itself evidence of their possibility.⁵¹ The objection is: suppose we had been that way; what we now take to be unconditionally normative we would not take to be so; but then how can

⁴⁹ Silverstein is quoting Frankfurt (1998: 18).

⁵⁰ See Tiffany (2012) for something like this strategy.

⁵¹ The claim is not that Humeans and Kantians are arguing about different things: the former, the minimum features necessary for intentional activity, and the latter the right account of intentional activity *for us*. It is just that even if Kantians can show the Humean model is *wrong* for us, this alone would not show that intentional activity along these lines is *inconceivable*.

what we now hold to be the case have a claim to be *the* normative truth? We might respond by insisting that *only* the richer account of intentional activity – one which allows that some considerations appear unconditionally normative to us – is conceivable.⁵² In which case, it was wrong to think the question “Why be an agent?” could be posed from perspectives other than this one; so there is no challenge to the authority of constitutive standards of agency that must be answered without reference to them. I will not discuss attempts to develop this strategy (e.g. Korsgaard 1997). Rather, I deny we need it in order to defend the constitutivist account of practical authority.

The objector acknowledges that *she* cannot occupy perspectives other than agency in order to challenge the authority of its constitutive standards. This follows from accepting that the constitutivist’s account of intentional activity is the right one *for us*. The objection is that this fact – that the role of agent is the only role we can occupy *as intentional beings* – is normatively insignificant. This might seem plausible. I argued earlier that questions about the worth of activities we cannot now (realistically) avoid may be valid, and moreover that appealing to such limitations by way of response could at most serve to highlight the *pointlessness* of, say, the question “Why be an agent?”. Thus, if the only resources constitutivists have to demonstrate the normative significance of *being an agent* is to appeal to some sense in which we cannot avoid being agents, we have already shown that the argument will fail.⁵³ The situation here is actually worse. The appeal to some sense in which we cannot avoid being agents would not simply fail to do more than show the question “Why be an agent?” is pointless in the present context: the response “You just have to!” given outside the perspective of agency is actually false. Of course, I argued that appealing to this response does not exhaust the resources for explaining why constitutive norms of agency must appear normative *to agents* – but these resources are redundant when the issue is how the question “Why be an agent?” can be answered outside agency.

⁵² See Lavin (2017) for discussion of this strategy.

⁵³ See also Enoch (2011). We might respond that it is “deeply” physically impossible for us to occupy the role of the non-agent we have been describing. But this *alone* would not show that the question “Why be an agent?” posed from outside agency is no threat to the authority of constitutive norms of agency for us (though the point is importantly related to my argument in the next paragraph). That we could not have reason to do what is “deeply” physically impossible – exist as an intentional being but *not* as an agent – would not show that constitutive norms of agency are authoritative for us after all. It would just turn the question of how we should respond to their lack of authority into a question about whether we should engage in any kind of intentional activity at all.

But the crucial point, again, is that *the objector*, cannot pose the question from this perspective – she only *imagines* what it would be like for it to be posed in this way. But that has no significance for what *she* should do. The question “Why be an agent?” in her mouth always concerns constitutive norms of agency. It may be objected that this only shows she is prevented from knowing the true basis of her reasons. But *this* is incoherent. I have argued already that conclusive reasons will motivate agents insofar as they are rational. So, if there are considerations outside agency which we cannot understand, and which could never motivate us, then this means they have no significance for what agents should do – they have no bearing on agents’ reasons. The correct implication of the constitutivist framework is that even if there could be a rational being who is not committed to the constitutive aim of agency, this being would still have to acknowledge that *insofar as one is an agent*, what *they* should do is exhausted by agents’ reasons. That follows even if constitutive norms of agency do not (directly) give this intentional being reasons, since she lacks a commitment to its constitutive aim and perhaps has no reason for adopting it.⁵⁴ So there is no perspective from which the question “Why be an agent?” could be posed such that it undermines the claim that *agents* always have reason to conform to constitutive norms of agency. The situation is not altered if the objector could *herself* occupy the perspective of a rational being that lacked any commitment to the constitutive aim of agency – say, if there was a pill she could take in order to become such a being. Still, this perspective is not one which generates insights into what we should do *as agents*.⁵⁵ But maybe the point is that if this pill existed, and we rejected constitutive norms of agency, we could simply take the pill in order to become beings for whom conforming to those norms was no longer justified. But this possibility would not threaten my argument. We could take the pill – but whether we *should* will still be determined by constitutive norms of agency, since *taking the pill* is an exercise of agency. It remains the case that constitutive standards of agency always give agents reasons. Suppose these tell us to refrain from taking the pill on the basis that we reject, say, moral demands. I can choose to do so – but it will have been a *bad* choice (that is, less of a choice).

⁵⁴ I speculate that her reasons *for* or *against* becoming an agent (or doing anything at all) derive from the constitutive aim of whatever concept *she* falls under in virtue of which she is a rational being.

⁵⁵ This may be Velleman’s point when he says, “If there are indeed alternative agential constitutions, the Kantian strategy doesn’t take sides between them; it merely insists that questions must be asked and answered within the framework of some constitution (2009: 143).”

Practical Authority and Moral Authority

3.1 Taking stock

I began by describing the phenomena I call ‘practical authority’ which a theory of the nature of reasons should explain. I argued that neither certain desire-based views nor realism is successful in this. If we follow realists in claiming reasons are not conceptually related to desires then we cannot explain why conclusive reasons will motivate us insofar as we are rational. But certain theories which accept what realists deny (e.g. Williams 1981b, 1995) cannot accommodate unconditional reasons – and some reasons (e.g. moral reasons) *do* possess this character. Next, I presented an alternative constitutivist strategy. Constitutivism agrees with views of the latter kind that reasons are conceptually related to antecedent motivations. But they are distinguished by rejecting the claim that the content of our desires is always ultimately contingent: we are committed to certain evaluative attitudes insofar as we are agents. This provides the resources to accommodate unconditional reasons.

I then considered an objection to this framework which abstracts from the details of particular constitutivist views. I rejected Enoch’s (2006, 2011) thought that the authority of considerations bearing on constitutive features of agency depends upon whether we want to be agents – which it *seemed* we do, if at all, only contingently. We can account for the authority of such considerations without appealing to that aim. In essence, I have argued that if agency has a constitutive aim, this would account both for why conclusive reasons will motivate us insofar as we are rational, and how unconditional reasons are possible.

To complete the argument that constitutivism explains practical authority, we must show there *is* a constitutive aim of agency (3.4). We must also address the *content* of this aim (3.3). I mentioned moral reasons earlier as an example of unconditional reasons. If we want to establish that constitutivism can accommodate not just unconditional reasons, but the unconditional reasons *we think we have*, we must show how the constitutive aim of agency might plausibly give us *moral* reasons.

3.2 Some Preliminaries

I will approach the task of showing agency has a constitutive aim via a further question that confronts desire-based views: since desires do not straightforwardly render their intentional objects good, *and yet only considerations which engage our desires can be reasons*, what determines that we should treat certain desires as reasons? Following Korsgaard (1996), I will argue that agents are committed to valuing practical reason itself, and do not need further reasons for doing so. This provides the criterion by which other desires can be selected as reasons. In 3.5 I will explore an apparent tension between this conclusion and my earlier claim that we can intelligibly *disvalue* our agency yet this does *not* undermine the authority of constitutive standards of agency.

My argument that agency has a certain constitutive aim *presupposes* aspects of the constitutivist framework. Realists will not be satisfied by my argument here since I assume what they deny: that reasons are conceptually related to antecedent motivations. I will pose a question about the internal coherence of desire-based accounts, and then suggest how to resolve these difficulties. I will ignore the possibility that the best response to the appearance of incoherence in desire-based views is to reject outright the claim that reasons are necessarily related to antecedent motives. Whether this begs the question depends upon whether the argument against realism in 1.6 is successful.

I have so far described constitutivism as a view about the nature of reasons. It is also, unavoidably, a thesis about the nature of value. I am mentioning this because I will talk about the constitutivist account of value in what follows, and I take my arguments to support *both* the account of value and the account of reasons (indeed, if they work at all, I think they must). There is an intimate relation between reasons and values. I will assume Raz's (2011) view that practical reasons just are the values there are in performing certain actions. Now, we might suppose that facts about value are fixed independently of desires (call this 'realism' about value), but that we only have reasons to promote, or act in otherwise appropriate ways towards, those values which are connected in the right way with our existing desires. Thus, we would agree with Raz (whilst disagreeing with his realist account of reasons) that some ends may be valuable whilst there is no reason to promote them. But this combination of claims is untenable.

Recall from 2.2 that I understand the constitutive aim of action to be some desire which we could retrieve from any set of antecedent motives whatsoever by a

process of rational deliberation. It is *the* source of practical reasons, so they are exhausted by the following: (a) any demand that we *do* something which follows directly from the constitutive aim of action (b) any requirement that some things *not* be done which also follows directly from the aim *and* (c) any reason generated by those ordinary contingent desires which relate to the aim in some (as yet unspecified) way. Now, suppose for the sake of argument that some consideration qualifies as a reason according to these criteria but that it does *not* concern the value there is in performing some action. What are we to say about this? Realists about value might simply accept there could be practical reasons where these do not at all concern the value there is in performing an action – and thus it is at least conceptually possible we could have reason to perform actions that are not valuable. This does not strike me as an attractive position.⁵⁶ Alternatively, we could deny the situation is conceptually possible. But how could that be? It could be so, assuming the truth of constitutivism, only if the standards that determine our reasons – constitutive standards of action – also determine what is valuable (i.e. establish at least sufficient conditions for the value of something⁵⁷). But these standards are necessarily related to antecedent motives. So taking this position means rejecting realism about value and embracing the following thesis:

... value is an *attitude-dependent* property... the attitude of valuing [is] the more fundamental explanatory notion and... value itself [is] a “construction” of that attitude. Things are valuable ultimately because we value them, not the other way round. (Street 2012: 41)⁵⁸

Call this the ‘Conferral Model’. Is this a kind of relativism about value? (In asking this, we are also asking whether constitutivism entails that reasons are relative.) The answer is both “yes” and “no”. Yes, because you and I may value different things – and the view claims there are conditions under which the object of your valuing will be valuable for you, and the object of my valuing will be valuable for

⁵⁶ It would also mean departing from Raz’s position – which I am assuming here – that practical reasons just are the values there are in performing certain actions.

⁵⁷ It is logically possible there could be other independently sufficient conditions, perhaps realist ones, for the value of something – values which would not provide us with any practical reasons. But it is unclear why we would think this kind of value and the kind of value which is necessarily connected to antecedent motives are species of the same thing since there is no feature they must both have in order to belong to it (the standards set by the constitutive aim of agency have to be sufficient for the value of something – and realist kinds of value, by definition, do not have to meet these standards). It seems more plausible that constitutive standards of agency set *necessary* as well as sufficient conditions for the value of something. I will assume this is the case. Finally, this means that whilst we accept Raz’s view that practical reasons just are the values there are in performing certain actions, we depart from his claim that that some ends may be valuable whilst there is no reason to promote them (and, obviously, from his view that values do not depend conceptually on desires).

⁵⁸ Emphasis original.

me, but neither of us will have a reason to value what the other values.⁵⁹ This will happen when desiring some end is compatible with constitutive standards of agency, but not demanded by them. *No*, because there are some ends the pursuit of which is incompatible with constitutive standards of agency. Hence, there are some ends which constitutive standards of agency either demand we value or prohibit us from valuing. My claim has been that constitutive standards of agency are *the* source of reasons and values, so there are some things which are valuable for all agents. These values are in an important sense not relative – even if they are “conferred” by each of us in virtue of being, necessarily, committed to the constitutive aim of agency. Some may still reject the relativistic aspects of the proposal. I cannot defend them here. But it is important not to project onto the Conferral Model the idea that literally anything we desire (once we rule out mistakes in instrumental reasoning, false non-evaluative beliefs, and so on) is thereby genuinely valuable for us. Indeed, part of my ambition in this chapter is to show that this idea is incoherent (3.4.2).

3.3 The Moral Character of the Constitutive Aim of Agency

I will argue that it is a condition of our agency that we are committed to valuing ‘the capacity for practical reason’ (‘CPR’ henceforth) – our ability to reflect upon whether some consideration justifies a particular action, that is, whether there are reasons to perform it.⁶⁰ This commitment grounds certain reasons directly, and provides the criterion by which desires confer the values that ground most of our practical reasons. This, being the constitutive condition of *rational* agency, determines (at least partly) the nature of practical *irrationality*: we are irrational if we fail to value CPR by failing to follow (through action or inaction) the requirements our commitment to its value establishes. I have argued that it is because there are constitutive standards of agency that we can account for unconditional reasons. The rational requirement that we act on reasons which the constitutive aim grounds directly, and not act on certain considerations if doing so is inconsistent with this aim, applies necessarily to all agents. So, if there are (unconditional) moral reasons this is because a commitment to valuing CPR grounds certain moral imperatives directly and imposes moral constraints on which desires confer value. We can examine this thesis independently of the argument

⁵⁹ At least we will not have a reason to value it for ourselves – we may have a reason to value us both getting what we value, and so a reason to help one another pursue what we each value (see 3.3).

⁶⁰ ‘Committed’, in the sense that our attitudes imply we would actually value it if we were fully rational.

that a commitment to valuing CPR is constitutive of agency.

What would failures to value CPR consist in? Would this capture the core examples of immoral action? I can only give some cause for optimism. First some general remarks on valuing. Valuing is not identical with having a positive “feeling” towards something. The latter is not even necessary for the former: we may value acts of charity as a matter of duty, even if giving up our cash or time never fills us with joy. The important point is that we must be disposed to act in distinctive ways in order to count as valuing something: generally speaking, I do not value doing yoga if I have rarely been inclined to do it. On the other hand, a mere disposition to perform certain actions is not sufficient for valuing them – someone who compulsively washes their hands does not therefore value doing so.⁶¹ For someone to count as valuing an end, we must be able to attribute to them the belief that achieving it would be in some way good.⁶² The two ideas are related insofar as if someone values an end, the disposition we would expect her to have is to pursue it when she judges herself to have most reason to do so.

How then are we to understand valuing CPR? First, valuing some capacity, in general, involves being disposed to exercise it in appropriate conditions – those cases in which we judge ourselves to have most reason to exercise it. CPR differs from other capacities in that exercising it is necessary for intentionally pursuing any other perceived good, since it is our ability to grasp that some end is valuable in the first place. So, the appropriate conditions for its exercise always obtain – there is reason to exercise CPR as part of any action we have reason to perform. Second, we must judge that our exercise of CPR is good in some way. The preceding suggests we might value its exercise instrumentally, as a necessary means to pursuing other goods. But this is to misconceive the relation, if the Conferral Model is true, between the ends which are valuable for us and our capacity to conceive of their value. If value is ultimately dependent upon our evaluative attitudes, then our ability to form such attitudes (CPR), if it is judged valuable at all, could only appropriately be valued for its own sake – since it is not

⁶¹ I thank Ulrike Heuer for this example.

⁶² I do not take a stand on whether we *want* to perform compulsive actions. Perhaps some such attitude with a “world-to-mind” direction of fit is needed to explain them. But if so, the notions of valuing and wanting come apart. To clarify, in using the term ‘desire’ throughout I have been assuming we judge the objects of our desires to be good in some way, so the notion of desire excludes such cases of wanting, if they exist. ‘Desiring’ is interchangeable with ‘valuing’ for my purposes.

instrumental to other goods, but rather our possession of it is a precondition of their value for us.⁶³ Both points will be important later (3.4.3).

Consider, then, an akratic – who is irrational. She acts against her considered judgements about what she ought to do, as laid down by her CPR. In light of the preceding, we can describe this as a failure to value (at least on this occasion) her CPR – since valuing it is, in part, to judge that we always have reason to exercise it, which she has not done here (at least not well). If so, we may draw the parallel conclusion that undermining, or failing to support in reasonable ways, the plans of others in accordance with what *they* judge themselves to have reason to do, also involves a failure to value CPR. It is natural to describe this as a moral failure.⁶⁴ Some might disagree. Perhaps what morally constrains our actions are objects, persons, and so on which we and others value *not* the fact that humans instantiate certain rational powers. This is a false dilemma. We can agree that these are the immediate objects of moral concern. This is compatible with insisting that we must care about them because they *are valued* – because they are so related to CPR.

A fair point to make about the argument I am going to present is that it does not establish the conclusion I am at this point endorsing – that we are committed to valuing *the* CPR insofar as we are agents. Rather, it shows that our agency commits us to valuing *our own* CPR. This is problematic because even if the latter commitment did imply we are subject to certain rational requirements, this would not give us a system covering *moral* requirements. Again, my suggestion is that we fail *morally* (at least it is one important way to fail morally) when we undermine or fail to support in reasonable ways the rational plans of *others*, that is, fail to value CPR *in them*. But this idea flounders if we do not fail rationally at all (let alone morally) by failing to value CPR in others – if we are not, as a constitutive condition of our agency, committed to its value. I will not be pursuing this line of discussion, but rather (for reasons of space) assuming without argument the following inference:

⁶³ Valuing CPR would in this respect be akin to how we value certain other of our capacities for their own sake – which involves taking pride in them (an idea which is, in the context of how we view our own rationality, familiar from Kant (e.g. *Gr* 4: 410-11)). The salient difference is that we may be committed to valuing CPR in others too (see below), so the relevant attitude is really a *collective* pride and respect for our rational form of life (thus it is not practically unhealthy, as other forms of pride can be). I leave aside questions about our relations to non-rational creatures. I thank Ulrike Heuer for pressing me on this.

⁶⁴ Our commitment to the value of CPR may generate (unconditional) requirements that are not obviously morally significant. It is plausible that an agent who values spending her days counting blades of grass is also irrational. This behaviour is inconsistent with valuing CPR since she is, necessarily, not pursuing any activities that might cultivate it. Pursuing some such activities is demanded as part of a rational response to the value of CPR.

- P) My agency is a sufficient condition of the value of my CPR.
C) I am necessarily committed to the value of any agent's CPR.⁶⁵

The intuitive idea is that if we judge the presence of some feature *x* to be enough for the presence of some other feature *y* in some concrete case, then, other things equal, we should judge *y* to be present whenever *x* is. Of course, someone might deny that other things are equal when I take my own agency to be sufficient for the value of my CPR, and when I consider whether your agency is sufficient for the value of your CPR – after all, I am me, and you are you. But this is at least not how moral agents understand their relations to each other: the simple fact of us being different persons is irrelevant to our value from the moral point of view. In fact, the original proponent of the argument I offer here – Korsgaard – has not always accepted this route to the claim we must value CPR in others (1996: 132-45).⁶⁶ I will not examine this disagreement here. Relatedly, I will not attempt to reconstruct Korsgaard's argument since my purposes are not exegetical. But the argument set out here is greatly indebted to it.⁶⁷

3.4 The Argument

3.4.1 Rational Agency

One of my main ideas in 1.2 was that our applications of the concept of a reason are bound up intimately with the intelligibility of praising and criticising actions. Indeed, praising and criticizing actions *just is* to subject them to a distinctive form of *normative* assessment – an assessment based on the reasons for performing them. Hence, such assessments are only intelligible when an action was, or could have been, done for reasons. Recall too the connection we established between whether some consideration can be a normative reason and the possibility of it explaining action. This allows us to distinguish cases where merely *evaluative* (that is, non-normative) assessments concerning our position in the world are appropriate.⁶⁸ For example, I may think it would be good (for her and for me) if my

⁶⁵ See de Maagt (2018: 891). To be clear, 'the value' here refers to values as understood by the Conferral Model (see 3.2).

⁶⁶ The failure of Korsgaard's alternative strategy is much discussed (e.g. Hills 2010: 96-101).

⁶⁷ See especially 1996: 120-25. I also draw upon work that *does* take itself to be reconstructing Korsgaard (Stern 2015; Street 2012), and some objections I address my proposal to be originally directed towards her position (Brady 2002; Fitzpatrick 2005; Ridge 2005; Stern 2015; Street 2012).

⁶⁸ It should be distinguished too from the kind of assessment of non-rational, but functional, life forms I described earlier, e.g. there is something amiss when certain plants fail to flower in the spring. The salient difference between this case and the one I am about to describe is that it is implied about the subject of the assessment, i.e. the plants, that they are entities

friend were to recover from her illness by tomorrow so we could get lunch. But assuming she takes steps to look after herself between now and then, it is unintelligible for me to criticize her if she does not recover. The thought that it would be good if this happened is incapable of functioning as a reason for her to actually secure that outcome – this is why it is unintelligible to criticize her merely for the failure of that outcome to materialize – because it is in important respects beyond her control.⁶⁹

But in virtue of what are we ever *in control* of whether something happens such that we might be criticisable with respect to it happening or not? Why do non-human animals never act for reasons? The preceding reveals that the possibility of the value of some end explaining our action is essential to the appropriateness of criticism. So, the answer to our questions is that we have the capacity to reflect upon whether certain considerations *justify* particular actions. It is the fact they justify, when they do, that guides us in the rational performance of actions. Certain other species can engage in reasoning of a kind – they adopt ends as responses to features of the world they encounter and then figure out how to realize them. In one sense we might say they “act for reasons” by pursuing courses of action that effectively realize these ends – which they even see as “good”, in the sense of being attracted to pursue them. But this is not the sense of acting for reasons that applies to us. *This* sense implies that the agent takes her ends to be justified – which is only possible if she can reflect on whether the considerations which she takes to support her ends really do support them, i.e. if she has CPR. The question of whether their ends are justified does not arise for non-human animals – their ends are just their ends. Hence, their actions cannot be said to be justified or not, in the sense of supported by reasons, and it is inappropriate to praise or criticize them as if they were. Rational agency, because it is normative, presupposes and is entailed by exercising CPR.

3.4.2 The Problem

The objection to realism in 1.6 was that if we deny reasons are conceptually connected to desires, then we cannot explain why reasons will motivate rational agents – which an account of the nature reasons must do to explain how any

which are incapable of ever having reasons – not merely that they did not have certain reasons on a given occasion. See below.

⁶⁹ The thought may still provide reasons for her to avoid negligence with her health. So it may be intelligible, and even warranted, to criticize her for failing to take certain prudent measures.

standards are normative for action. So, these standards must be dependent upon our desires. This conclusion *should* strike as overly hasty. If the charge against realism is that the facts it appeals to as the grounds of our reasons are not normative, then exactly the same charge can be levelled against desire-based accounts. That is, it is unclear how the addition of desires into the picture could succeed in accounting for the *normative* dimension of action, since desires are not in themselves normatively significant. That we desire some end (once we rule out mistakes in instrumental reasoning, false non-evaluative beliefs, and so on) is not straightforwardly a reason to pursue it. This is the familiar point that we can desire ends, sometimes very strongly, which lack value entirely. This bears on desire-based views in the following way. Desiring some end is a necessary condition of having a reason to bring it about. But now it seems the contribution our inclination to pursue some end makes to our having a reason to do so cannot be a *sufficient* contribution. That is because the value of the end as it appears to us may be non-actual.⁷⁰ We need a reason to treat our desires as reasons – that is, we need a reason to hold they are desires to promote or act in otherwise appropriate ways towards some genuine good. The problem for desire-based views is that they claim values and reasons are *always* dependent upon evaluative attitudes. But it now seems we have to use the notion of reasons – the notion evaluative attitudes are supposed to explain – in order to explain when desires are linked to genuine goods, and hence can be treated as reasons. A vicious form of circularity threatens. I will argue that the objection is avoided because we must treat certain desires as reasons insofar as we are agents – and we do not need further reasons for doing so (3.4.3).

Before this, the nature of the problem needs to be clarified further. In order to act, we must take some considerations to be reasons. I have indicated that any exercise of agency is governed *constitutively* by constraints which, it seems, can only be established through theoretical arguments. Am I suggesting that failing to establish this result here – or, if that is successful, of failing to act in accordance with the constraints in question – means it is impossible to act for reasons? This seems absurd. After all, agents go about their lives without worrying whether it is beyond doubt that they should take the considerations they do to be reasons

⁷⁰ I take this point to be completely general. Someone might object that because *whenever* we desire something obtaining it will be pleasurable, then obtaining it will result in *some* good because pleasure is good – even if this good is ultimately outweighed by its bad consequences. So desiring does, by itself, make its intentional object good (in one respect). I do not think this is right. The pleasure that will result from an action because it is desired may not be a good-making feature of it (if the action is sadistic, say). We must be attentive to other features of the pleasurable action to determine whether this is so.

(Brady 2002: 409). If there is a solution to the “problem” that we must take some considerations to be reasons in order to act it seems to be: take some fact, any fact, to be a reason. And we do that without trouble anyway. But this is to misunderstand the problem. The problem *is* that we act for reasons without any trouble – or rather of explaining this. The argument I will offer is “transcendental” (Stern 2015; Korsgaard 1996: 123-24). It assumes we act for reasons to see whether this can be given a coherent explanation. There are familiar worries about this form of argument (e.g. Stroud 1968). Perhaps the fact we must think and behave in certain ways in order that normative thought is possible does not show that it latches onto an aspect of reality. The thought which it makes possible may itself be illusory. But as Stern (2015: 61) notices, such objections have limited force against a view which already concedes that what normative thought is about (reasons) is dependent upon subjective experience: our rational aims and desires.

So, to return to the original point, if my aim is not to offer a solution to a problem which is trivially overcome – of getting us to act for reasons – is it rather to offer a characterization of how agents deliberate when they successfully act for reasons? This too is misguided. Since the characterization would be the product of theoretical arguments, it is implausible that it could feature at the level of ordinary conscious processing. Rather, I will describe the *commitments* which our reasoning has to embody in order that it leads to action. Simply because we may not realize we are committed to something does not undermine the idea, which is crucial to this essay, that such commitments will motivate us insofar as we are rational. It will motivate us insofar as we are rational precisely because its endorsement is *rationaly* implicated in any set of ends that an individual endorses.⁷¹ That is, building upon Williams’s (1981b, 1995) conception of rational powers described in Chapter One, they are directed towards attitudes which we can acquire by subjecting our antecedent desires to a process of rational deliberation. It does not follow that agents will necessarily care about satisfying these desires – that is, they may not care about being rational. Moreover, there may be nothing we can say to such a person to persuade her that she should care than to point out (unhelpfully) that she is irrational if she does not. But an account of reasons was never meant to rule this situation out.⁷²

⁷¹ Thus, the proposal is not hostage to empirical fortune (as Ulrike Heuer suggested to me). It does not follow from the view that if someone does not actually possess the attitude in question she is not an agent. Rather, it follows from the fact she is (minimally) rational – that she takes some considerations to be reasons, something we might in principle establish empirically – that she is committed to having it. I take it that rational requirements are not something which can be established empirically.

⁷² So, our criticisms of realism should not, and do not, hinge on the claim that it has nothing to say to someone who does not care about being rational. See Fitzpatrick (2005: 675).

3.4.3 The Solution

In sum, the challenge for desire-based accounts is that, since the mere fact we desire something does not make it good, there must be a reason to treat certain desires as reasons, and by implication their intentional objects as good. If this reason is itself a desire, as desire-based theories hold, then given the preceding thesis about desires, we need another reason to treat this desire as a reason, and so on indefinitely. If there is a solution to this problem, it must take the following form: there are some desires which we are committed to treating as reasons, which we do not need further reasons for so treating, and that give us reasons to treat other desires as reasons. The intentional objects of the former kind of desire would be unconditionally good. The condition of the goodness of any other end would be that these desires ground reasons to treat desires for the other ends as reasons to pursue them. I will argue that a certain desire can play this role – a desire directed towards practical reason itself.

But why think only a desire with a certain content can do this: are there not an abundance of evaluative attitudes that do not stand in need of further rational support, and which yield reasons to take other considerations to be reasons? For example, Stern asks (2015: 67-8), if he is a loving son, what further reason does he need to act in ways that befit this identity? His identity gives him reasons to act in these ways and the identity itself can be brought into doubt only by giving him reasons to reject it. But why should he accept any such consideration *as a reason* unless he has already given up the identity? That is, his identity gives him reason to deny they are reasons. But the issue is not whether the considerations we treat as reasons in ordinary deliberation are normatively suspect unless we can vindicate them completely.⁷³ The problem is about how practical reason is possible at all if we accept the metaethical premise that reasons are conceptually related to antecedent desires. Returning to Stern's example: it is *qua* desire (judged from the philosopher's perspective) that his desire to act as befits a loving son is normatively suspect, and not *qua* perception of a reason to act in this way. The ease with which this perception makes a claim on him is what we are trying to explain – not what we wish to put into doubt. It is only by abstracting from the concerns that occupy us in everyday life that we can bring the problem into focus. It is also for this reason that no ordinary desire can resolve our difficulties, but only a fairly abstract one which is suitable to the context of the problem.

⁷³ Stern recognizes this point. I have benefited from his take on Korsgaard's argument.

I must now show that the normative force of certain desires would be explained if we were committed to valuing CPR. I have to show that it (i) gives us a reason to treat other desires as reasons (subject to the constraint that doing so is compatible with valuing CPR) and, (ii) that we do not need further reasons to value it.⁷⁴ If we value CPR this involves thinking we have reason to use the capacity for its own sake, as I argued in 3.3. This gives us reason to take other things to be valuable, so we can see them as giving rise to reasons – thus satisfying (i). This is because, as noted in 3.4.1, a rational agent *just is* someone who exercises CPR – which means treating some considerations as reasons. As Street puts it, rational agents “*need reasons to act*” (2012: 47).⁷⁵

It remains to be shown that we do not need further reasons to value CPR – thus satisfying (ii). This is needed to complete the argument that we are committed to valuing CPR insofar as we are agents. We must remind ourselves that even if it is true that the value of CPR gives us reason to treat other desires as reasons, if we need further justification for valuing CPR our initial problem remains. If a desire directed towards practical reason itself was like any other desire, we would need reasons to hold that it was a desire for some genuine good in order to treat it as a reason. But this further justification would according to desire-based views just be rooted in further desires, for which the same issue can be raised. Desires do not by themselves make their intentional objects good, and thus are insufficient to ground reasons *qua* desire. But we do *not* need further reasons for valuing CPR – our commitment to valuing it provides its own justification. The answer we can give in response to the question “Why value CPR?” is that we have the same reason to value it as we have to value things in general – we must value some things in order that we can see ourselves as having practical reasons.⁷⁶ But we have *this* reason because we are already committed to valuing CPR.

In short, the argument is this. We are committed to valuing CPR insofar as we are rational *if a desire-based account of reasons is true* because only such a commitment could explain how genuine normativity is possible within that framework. Since I have argued in favour of desire-based accounts, it follows that

⁷⁴ The next part of the argument draws upon Street’s (2012: 47) reconstruction of Korsgaard.

⁷⁵ More on this in 3.4.5.

⁷⁶ Someone might re-raise the objection from 3.4.2 why this does not simply give us a reason to value anything – rather than, specifically, CPR. So the structure is clear, I am suggesting we do have this general reason to value things. But if we accept a desire-based theory, that we have this reason (and, by implication, any reason at all), only makes sense if we assume commitment to some value which implies we have this reason and is itself justified by it. The value of CPR fits this specification.

we are committed to valuing CPR insofar as we are rational. This commitment is constitutive of agency and explains the phenomenon of practical authority in the way described in Chapter Two. We can now see that the explanation of how valuing CPR satisfies (ii) and the explanation of how we act for reasons are mutually supporting. We are entitled to the assumption that we are *already* committed to valuing CPR, and hence we do not need further reasons to value it (satisfying (ii)), because *acting for reasons is possible* (what needs to be explained) and a commitment to valuing CPR is a condition upon acting for reasons (the conclusion of the argument). So too, a commitment to valuing CPR is a condition upon acting for reasons (in part) because it satisfies (ii). This *is* circular. But not viciously so, since we only wanted to show there is a coherent explanation of what seemed obvious anyway – that we act for reasons.

3.4.4 The Unconditioned Condition of Value

I will now discuss some implications and objections in the rest of the chapter. The conclusion that we are committed, as a condition of our agency, to *valuing* CPR is surprising because it assigns a role to CPR which is not demanded by the minimal conception of rational agency outlined in 3.4.1. It may be a trivial fact that when we act – when we exercise CPR – we *have* CPR. But it does not follow directly that CPR is ever (let alone always) valuable or that we have reason to exercise it.⁷⁷ There is, after all, no general principle connecting mere possession of a capacity to the value of that capacity or the value of its exercise on a particular occasion. I have the capacity to swim: that this is a *good* skill to possess may be a reason to nurture it in various ways, but this does not preclude circumstances in which there is *no* reason to exercise it (e.g. if the tide is very strong and swimming in the sea is the only way I could exercise it right now) (Raz 2011: 158-59). It is also true that most human beings have the capacity to kill each other. Arguably, this is a capacity we *never* have reason to exercise – and the reasons we have to *prevent* its exercise exist because it is *unfortunate* that we have this ability. We cannot usually say what reasons are generated by possessing capacities without first specifying whether it is good or bad (on balance) for us to have them, and whether it is good or bad that we exercise them on a given occasion.

⁷⁷ And to exercise it *well*. Someone might worry that even if I can show we always have reason to exercise CPR it does not follow that we always have reason to exercise it well, and hence to conform to constitutive standards of agency. But actually our reason to exercise CPR does imply we should do it well because of the manner in which we must value it – *for its own sake* – which gives rise to this reason.

I have insisted we are committed to valuing CPR insofar as we are agents. Thus, CPR is not only valuable, but unconditionally so – valuable in every conceivable circumstance – since I am arguing that the possibility of all value depends upon our commitment to *its* value.⁷⁸ This is a welcome conclusion for my argument. I hold that moral requirements are grounded in constitutive standards of agency. These are unconditional and hence presuppose values which are unconditional: practical reasons just are the values there are in performing certain actions; so if a reason applies in every conceivable circumstance, it must appeal to something which would be valuable in every conceivable circumstance. If when we criticize an agent's immoral action we point to the good an alternative action furthers besides the moral good itself, our criticism is not a *moral* one. Morality requires we do the right thing for its own sake, and not for these additional benefits. So, the goodness that acting morally instantiates cannot be explained as derivative from further valuable goals which are *contingently* furthered whenever we act morally. The moral good is unconditionally good, and hence moral reasons are unconditional. This is just the character of the requirements imposed by constitutive standards of agency.

Ridge worries that arguments for the conclusion that CPR is unconditionally good (like mine in 3.4.3) depend on the idea that “for anything to be the unconditioned condition of the good of anything else, that thing itself must be unconditionally good (2005: 59)”. If there is a condition on the possibility of all value, then it must be valuable and *its* value cannot depend on the value of anything else. So, supposing my argument did apply the inference, it would have this form: because we must believe some ends are valuable in order that we act for reasons, and the value of ends is conditional upon exercising CPR⁷⁹, we are committed to valuing CPR in virtue of acting for reasons. But Ridge doubts whether the principle is sound: “Why should one suppose that simply because X is the condition for the goodness of everything else that X itself must be good *at all*, much less unconditionally good? (Ibid: 65)”.⁸⁰ Indeed, I myself have provided apparently disconfirming evidence to the principle; namely, my claim that it can be rational for agents to regret their agency, precisely to *disvalue* it. How could this be rational if agency is unconditionally good? I will address this in 3.5.

⁷⁸ So more properly, it is valuable in any conceivable circumstance in which anything at all has value – that is, in which there are agents. This fits with the Conferral Model which holds that values just are dependent upon the (rational) evaluative attitudes of agents.

⁷⁹ This is just the claim, from the Conferral Model, that values are dependent on the (rational) evaluative attitudes of agents, and so on our ability to form such attitudes (i.e. on CPR).

⁸⁰ Emphasis original.

First, why should we accept the principle? The crucial point, again, is our prior acceptance of the claim that if a consideration is a reason, or some end is valuable, this depends, necessarily, on our antecedent desires. I have argued there is a problem about how this dependence allows us to explain the value of any action. I have explained how this problem is solved once we see that acting at all commits us to the unconditional value of CPR. The commitment, if we have it, explains how the fact that something is an object of (rational) evaluative attitudes is what its value consists in. This supports the thesis that we actually have such a commitment. So it is not, as Ridge suggests, that the argument *depends* on the principle above as independently plausible. Rather, it *follows* from this line of reasoning that the unconditioned condition of the good of anything else is unconditionally good. If I am to make anything of Ridge's objection it is this. There seems to be a gap between what I have established, that we are committed to the unconditional value of CPR, and the conclusion that it *is* valuable unconditionally. But this is again to ignore the significance of presupposing that the Conferral Model is true. Values *just are* what we value in accordance with rational standards. I have argued that our commitment to valuing CPR is what provides this standard. It is the constitutive condition of *rational* agency. It follows that it is rational to value CPR – which establishes its actual (unconditional) value.

3.4.5 Inside Agency

Street (2012) takes issue with the role of the question “Why value CPR?” in the argument.⁸¹ To recap, I have argued that a commitment to valuing CPR is the constitutive condition of agency. This is the case only if, in addition to giving us reasons to treat certain desires as reasons, we do not need further reasons for valuing CPR. So, the question “Why value CPR?” is a test for whether this condition is satisfied: do we need to go beyond what the value itself commits us to in order to justify our acceptance of it? But Street objects that because our commitment to the value in question is hypostasized as the constitutive condition of agency, the test does not make sense; so we could never establish whether the conditions for the truth of our hypothesis are satisfied:

In effect, one is asking: Does it matter (*as judged apart from the standards that determine what matters*) whether I take anything to matter? The proper reply to such questions is not to say yes, but rather to reject the question and say: Either you take something or other to matter or you don't; either you take something or other to be a reason or you don't. If you do, then something matters for you; then

⁸¹ See de Maagt (2019) for another response to Street.

you have reasons. If you don't, then nothing matters for you; then you have no reasons. (Ibid: 49-50)⁸²

The claim is that because the conclusion is that our commitment to valuing CPR establishes the normativity of any consideration whatever, there can be no question within this scheme of whether that commitment is itself justified. It is just the precondition for the sense of any question concerning *our* reasons. There is some truth in what Street says. But it does not threaten the argument. She is right that (on the present view) questions about reasons, when posed by some creature who is committed in no way to valuing CPR, cannot have any significance for what we should do, and is unintelligible to us. This is an idea familiar from 2.5. Street is wrong, however, that the question on which the argument depends is *this* kind of question. That is, we assume for the purposes of argument that we act for reasons and attempt to show how this is possible. In effect, we are asking: suppose a commitment to valuing CPR were constitutive of agency, then what justification would an agent (by hypothesis, someone who is committed to valuing CPR) have for valuing CPR? Since (*inter alia*) her justification could be provided by her commitment to valuing CPR itself, we conclude that this *is* constitutive of agency because it renders intelligible our original assumption that we act for reasons. The crucial point is that precisely because of this assumption, the question “Why value CPR?” is only ever entertained as a question posed from *inside agency*.⁸³

Street anticipates something like this response when she says “the Kantian constructivist might claim that these claims *are* asserted from somewhere – namely the standpoint of agency itself (Ibid: 50)”.⁸⁴ I have suggested the argument is framed in terms of an agential perspective which is defined by its commitment to a particular value. But Street doubts whether such a commitment could provide reasons, either for itself or for treating other considerations as reasons, which it would need to in order to satisfy the conditions for being constitutive of agency. I claimed that because valuing CPR means, at least, thinking we have reason to exercise it for its own sake, and exercising it means treating considerations as reasons, its value would provide reasons to treat considerations as reasons. This is also the justification it provides for itself. The idea here may be summed up by the locution that we “need reasons to act”. We can focus on this locution to better understand Street’s objection to the idea that constitutive conditions of agency can give us reasons:

⁸² My emphasis.

⁸³ There are limits to the intelligibility of even this interpretation of question. See 3.5.

⁸⁴ Emphasis original.

On one interpretation, such claims state conceptual, “constitutive” truths, such as “To be a parent, one needs to have children.” On this interpretation, these claims merely state what is involved in being an agent... On a second interpretation, however, such claims state substantive normative truths, according to which an agent has a reason to take something or other to be a reason... (Ibid: 51)

So, Street claims the locution that we “need reasons to act” involves either (i) the purely descriptive claim that agents, necessarily, treat some considerations as reasons, or (ii) the explicitly normative claim that agents have a reason to treat some considerations as reasons. Her objection is that whilst the second is needed to support the conclusion of the argument, only the first is true given the constitutivist framework. We need not reargue the case for (i): the claim that agents recognize reasons is just another way of saying their actions are subject to normative assessment. Street’s claim that (ii) is false is essentially an extension of her original objection: the idea we could have reasons to be agents is confused because being an agent is rather a presupposition of having reasons and provides the criteria for determining what they are. But this concern is not empty even once we agree that the question is asked from *inside agency*, so there is a criterion for determining our reasons. The point might be that “reasons to be agents”, or to treat considerations as reasons, still cannot be amongst them, since if we do not conform to such “reasons” we do not have reasons to begin with.

The possibility which Street ignores is the one developed in 2.4: constitutive standards of agency would provide *agents* with reasons to conform to them even if they did not *exactly* provide reasons for them (or non-agents) to adopt or continue to subject themselves to those standards.⁸⁵ The authority which constitutive standards of agency have for agents can be explained simply by the claim that if there are reasons for them to do *anything* then this has to be established by appealing to constitutive norms of agency. Does Street offer anything that undermines this? She provides the following as a counterexample to the idea that constitutive conditions of being an X provide X’s with reasons to satisfy those conditions: “a “parent” who does not have children does not have a reason to have children; rather, he or she is not a parent at all (Ibid: 51)”. This too can be handled by my account of why constitutive conditions of *agency* provide agents with reasons. The explanation why the constitutive condition of parentage (to be distinguished from “parenting”) does not provide even an evaluative measure is that it is not, in a certain sense, a *standard*. That is, it does not admit of degrees of failure in conforming to it. Agency is unlike this. We are all committed to valuing

⁸⁵ The idea that they could provide non-agents with *any* reasons is senseless unless they have an independent reason to conform to them.

CPR, but this does not mean we live up to our commitment. We may exercise our agency *poorly* – such evaluations are rooted in how well we carry out our constitutive function as agents. That we have *reasons* to improve as agents is a consequence of the further arguments I have given in favour of building an account of the nature of reasons out of this evaluative standard.

3.5 The Value of Agency and Regret about Agency

I will conclude by addressing an apparent tension in my argument. This is brought into view by an earlier strategy which I have just re-invoked. This is the claim that reasons to be agents are not needed to explain why, if there are constitutive standards of agency, those standards would give agents reasons. But now I have argued that we *do* have such a reason because our commitment to valuing CPR is constitutive of agency. So, first, why did I need to make the earlier point if I was, anyway, going to claim that constitutive norms of agency are self-validating? The answer is that without that argument, the case I have been making that we have reason to treat considerations as reasons would be vulnerable to the sort of objection raised by Street above. The charge that there could not be reasons to be agents gets much of its sting from the idea that the force of those “reasons” must derive wholly from the fact that constitutive standards of agency are self-validating (if in fact they are). But the self-validation of standards does not seem sufficient to generate reasons to conform to them. If there are incommensurable self-validating standards, how do we determine which standards to follow? This concern is heightened by the fact that the standard in question, being an agent, is so abstract and thus has the appearance of being rationally optional. I will return to this shortly. The significant point is that the sting of this argument is removed once we see that the authority of constitutive norms of agency does not derive from their being self-validating, but from their being constitutive standards of *agency* when such standards are *the* source of reasons. This explains why I could not progress directly to the argument of this chapter.

But this leaves a more serious worry untouched. How did I demonstrate that this alternative interpretation of the authority of constitutive norms of agency is correct? In part, by arguing that it is intelligible to *regret our agency*. This was meant to rule out the suggestion that having reason to be agents could be the sole explanation why constitutive norms of agency are rationally binding. This is because the intelligibility of such regret indicates we do not always have *most* reason to be agents. So if that reason were all that bound us rationally to constitutive norms of agency, then those norms could not provide us with unconditional reasons, e.g.

moral reasons. Suppose I do not have most reason to be an agent, and further that constitutive norms of agency are what generate moral reasons. In this case I do not have a reason to act morally – which leads to the absurdity that moral reasons are conditional. If I do not have most reason to be an agent, and I do as I have most reason to do, then I will not be an agent. If it is actually *being* – or at least having most reason to be – an agent in virtue of which constitutive standards of agency give me reasons⁸⁶, then the moral standards which are included in these will not be binding upon me. Thus, the intelligibility of regret about agency suggests that a different account of the authority of constitutive norms of agency is needed.

I have now argued that we are committed to valuing CPR insofar as we are agents. But then how can it ever be rational to *disvalue* our agency? If these conclusions are incompatible this would create one of two problems for my argument. If we reject the conclusion that it is intelligible to regret our agency, then we cannot appeal to this idea to show that the explanation of the authority of constitutive norms of agency must be other than that we have reason to be agents. But, if there is a compelling case for the intelligibility of regret about agency, this shows there is a defect in the argument that a commitment to valuing CPR is constitutive of agency. Some other value *might* play this role. But we do not yet know what it is, and so whether it holds out the prospect of grounding moral reasons (amongst other unconditional reasons which we recognize). It may then be advisable to just reject the idea that agency has constitutive standards as ill-conceived, and so no progress has been made towards an explanation of practical authority.

Fortunately, we need not fear any of these possibilities because the conclusion that regret about agency is intelligible and the conclusion that a commitment to valuing CPR is constitutive of agency are not in fact incompatible. The reason why is simple. If the latter conclusion is correct, we always have a reason to be agents. But it does not follow that this reason is *indefeasible*.⁸⁷ Thus, it is consistent with the intelligibility of regret about agency – that is, that we could have most reason to *not* be agents. At first glance, this may not seem obvious. I have argued that because our commitment to this value is constitutive of agency, if we cannot value

⁸⁶ This is similar to a point about instrumental reasons. If I have some reason to study in Bloomsbury now, but most reason to shop in Stratford, do I have any reason to get on a tube to Bloomsbury (allowing that these are exhaustive of my primary reasons)? Arguably not. I only have reasons to take the means to an end if I have most reason to pursue that end – or at least if I intend to pursue an end which is supported by reasons that are not greatly outweighed by the reasons for some easily available alternative.

⁸⁷ Notice this does not undermine the authority of constitutive norms of agency so long as our reason to be agents is not the sole explanation why they are binding on us.

some end consistently with valuing CPR, then it cannot be valuable and thus we cannot have reasons to bring it about. So how could we ever have a reason – let alone one which *outweighs* our reason to be agents – that in order to follow we would have to act against reasons which are given directly by this value? That acting on some consideration has this consequence surely excludes it from being a reason. This is true, but subject to a minor qualification which shows we could have most reason to not be agents. Sometimes the world, and the choices it forces upon us, are inhospitable to our agency. We can only do our best to hold ourselves together in the circumstances. So, whilst some consideration cannot be a reason unless it is, *in principle*, possible to act on it and on the reasons which are given directly by the value of CPR, it may sometimes be impossible to act on both.⁸⁸ In which case, it is for first-order normative theory to determine what we have most reason to do. Nothing we have said forces the claim that reasons which are given directly by the constitutive condition of agency are always overriding.⁸⁹ On the contrary, my comments about the intelligibility of regret about agency suggest they are not – at least with respect to considerations bearing only on *our own* rational capacities.

Of course, it also follows from my view that even when regret about agency is rational, agents must recognize that they have *some* reason to be agents. Does this undermine the intelligibility of regret about agency? Clearly not. The question “Why be an agent?” (like a certain version of the question “Why be moral?”) which an agent might pose need not imply that she fails to recognize the normativity of considerations bearing on constitutive standards of agency. (If it did, we can explain why *that* question is unintelligible. This person, like the moral skeptic, is asking something like: why care about the thing which I must already care about in order to ask whether I should care about it or anything else?) In order to give voice to the kind of regret about agency which we can understand, an agent need only doubt whether she has most reason, on this occasion, to be an agent.⁹⁰

⁸⁸ See Korsgaard (2009; 2018: 259-62) for related discussion.

⁸⁹ I do not commit myself to any view about the priority of moral constraints on action – which are given directly by the value of agency. However, if someone were to make the case that they are not always overriding, this point would go some way to mitigating the charge that my view fails to recognize the practical importance of considerations which do not fall under the narrowly moral, impartial point of view, e.g. personal relationships and attachments (Williams 1981a). Even if their practical importance is conditioned by the value of CPR, it need not follow that we should act on reasons given directly by the latter when the two conflict.

⁹⁰ Similarly, she could intelligibly ask whether she has most reason to be moral, on this occasion, so long as it is not part of our moral concepts that moral reasons are always overriding (see previous note).

Regret about agency can never be to completely disvalue it; so long as it remains intelligible, it cannot be to give up on agency completely. This reflects generally on the fact that our predicament as agents is imperfect: we cannot do everything we have reason to do. Perhaps this fact can itself be a source of regret about agency. We are in the tragic position of both being conscious of the ways in which our agency is constrained by the societies, smaller social units, and ultimately the bodies we inhabit, and at the same time having the capacity to transcend these limits in thought. We can experience this capacity as a burden in a world in which we are in control of so little. I do not know whether this is a reason to not be an agent.⁹¹ But if it is, it conflicts with the reason we all have to be agents. Maybe this is why regret about agency is often experienced as a persistent unease as we just go on living our lives – and if it emerges as a powerful threat to our existence, this is often experienced as a struggle, and not as an unopposed shift in the balance of reasons.

⁹¹ There are unanswered questions about how we could *in principle* act for this reason and follow all the reasons that are given directly by constitutive standards of agency.

Conclusion

I have just argued that we are committed, as the constitutive condition of our agency, to valuing CPR. Let us recall the significance of this. Chapter One made two claims about what it is to be subject to reasons. First, there is a necessary connection between the existence of normative reasons and the possibility of their explaining action in a certain way. Part of what follows is that conclusive reasons will motivate us insofar as we are rational. Second, there are unconditional reasons – moral reasons being an example. I argued that realism cannot explain the first point, and that certain desire-based views cannot explain the second. Thus, I said that neither can explain practical authority. Chapter Two argued that *if* there is a constitutive aim of agency, this explains practical authority. If all reasons are in some way grounded in this aim then we can agree with the aforementioned desire-based theories that reasons are necessarily related to antecedent motives – thereby explaining why conclusive reasons will motivate us insofar as we are rational. But we now claim that a certain aim is constitutive of agency, and so we reject the idea that the content of our motives is always ultimately contingent. This allows for the existence of unconditional reasons – unlike the desire-based views already discussed. Chapter Three saw this idea through by showing there *is* a constitutive aim of agency. We have, then, an alternative account of the nature of practical reasons which addresses the faults of some major rivals.

I will finish by noting various threads for further development. First, the final argument that there is a constitutive aim of agency relies entirely on the success of the case against realism in 1.6. Whilst I am convinced by the case there presented, I wish I could have made space to consider how realists might respond, given the importance of this disagreement to my wider project. Second, my side of this dispute would be helped if the alternative to realism was more fully developed. Two areas in particular need further attention: (i) how the relativistic aspects of my proposal are consistent with intuitions about value; (ii) how relativism is avoided entirely in epistemic matters. Lastly, I was only able to gesture at what first-order normative conclusions may arise from my thesis about the constitutive aim of agency, e.g. do we have moral reasons? This depends partly on whether we are really committed to valuing *our own* CPR or to valuing these faculties wherever we find them. This is, sadly, for another time.⁹²

⁹² I thank Ulrike Heuer, who supervised this work. I also thank Lucy O'Brien and John Hyman for feedback on my Wollheim submission, and Victor Weber for comments on an early draft of Chapter One.

Bibliography

Brady, M.S., 2002, "Skepticism, Normativity, and Practical Identity", *The Journal of Value Enquiry* 36(4): 403-12

Dasgupta, S., 2017, "Normative Non-Naturalism and the Problem of Authority", *Proceedings of the Aristotelian Society* 117(3): 297-319

Dreier, J., 2015, "Can Reasons Fundamentalism Answer the Normative Question?", in Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson, and Fredrik Björklund (eds.), *Motivational Internalism*, Oxford: Oxford University Press: 167-81

Enoch, D., 2006, "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action", *The Philosophical Review* 115(2): 169-98

-----, 2011, "Shmagency Revisited", in Michael S. Brady (ed.) *New Waves in Metaethics*, Basingstoke: Palgrave Macmillan UK: 208-233

Ferrero, L., 2009, "Constitutivism and the Inescapability of Agency", in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics* 4, Oxford: Oxford University Press: 303-33

Fitzpatrick, W.J., 2005, "The Practical Turn in Ethical Theory: Korsgaard's Constructivism, Realism, and the Nature of Normativity", *Ethics* 115(4): 651-91

Frankfurt, H.G., 1988, "Freedom of the will and the concept of a person", reprinted in his *The Importance of What We Care About: Philosophical Essays*, Cambridge: Cambridge University Press: 11-25. Originally published in 1971, *The Journal of Philosophy* 68(1): 5-20

Heuer, U., 2011, "Guided by Reasons: Raz on the Normative Explanatory Nexus", pp. 353-65, in George Pavlakos, Niko Kolodny, Ulrike Heuer & Douglas Lavin, "Three Comments on Joseph Raz's Conception of Normativity", *Jurisprudence* 2(2): 329-78

Hills, A., 2010, *The Beloved Self: Morality and the Challenge from Egoism*, Oxford: Oxford University Press

- Katsafanas, P., 2013, *Agency and the Foundations of Ethics: Nietzschean Constitutivism*, Oxford: Oxford University Press
- Korsgaard, C.M., 1996, *The Sources of Normativity*, Cambridge: Cambridge University Press
- ., 1997, "The Normativity of Instrumental Reason", in Garrett Cullity & Berys Gaut (eds.), *Ethics and Practical Reason*, Oxford: Oxford University Press: 215-54
- . 1999, "Self-Constitution in the Ethics of Plato and Kant", *The Journal of Ethics* 3(1): 1-29
- ., 2008, "Taking the Law into Our Own Hands: Kant on the Right to Revolution", reprinted in her *The Constitution of Agency*, Oxford: Oxford University Press: 233-62. Originally published in Andrews Reath, Barbara Herman, and Christine M. Korsgaard (eds.), 1997, *Reclaiming the History of Ethics: Essays for John Rawls*, Cambridge: Cambridge University Press: 297-328
- . 2009, *Self-Constitution: Agency, Identity, and Integrity*, Oxford: Oxford University Press
- Kant, Immanuel, 1998, *Groundwork of the Metaphysics of Morals*, Mary Gregor (trans.), Revised edition 2012, Mary Gregor and Jens Timmermann (eds.), Cambridge: Cambridge University Press [Abbreviated to *Gr.* References are given by the page number of the relevant volume of *Kants gesammelte Schriften*, which appear in the margins of most translations]
- Lavin, D., 2017, "Forms of Rational Agency", *Royal Institute of Philosophy Supplement* 80: 171-93
- de Maagt, S., 2018, "Korsgaard's Other Argument for Interpersonal Morality: the Argument from the Sufficiency of Agency", *Ethical Theory and Moral Practice* 21(4): 887-902
- ., 2019, "Why Humean constructivists should become Kantian constructivists", *Philosophical Explorations*: 1-14
- Nagel, T., 1970, *The Possibility of Altruism*, Oxford: Clarendon Press
- Raz, J., 2011, *From Normativity to Responsibility*, Oxford: Oxford University Press

- Ridge, M., 2005, "Why Must We Treat Humanity With Respect? Evaluating the Regress Argument", *European Journal of Analytic Philosophy* 1(1): 57-74
- Scanlon, T.M., 2014, *Being Realistic About Reasons*, Oxford: Oxford University Press
- Shah, N., 2003, "How Truth Governs Belief", *The Philosophical Review* 112(4): 447-82
- Silverstein, M., 2015, "The shmagency question", *Philosophical Studies* 172(5): 1127-42
- Smith, M., 1994, *The Moral Problem*, Oxford: Blackwell
- Stern, R., 2015, "The Value of Humanity: Reflections on Korsgaard's Transcendental Argument", reprinted in his *Kantian Ethics: Value, Agency, and Obligation*, Oxford: Oxford University Press. Originally published in Joel Smith and Peter Sullivan (eds.), 2011, *Transcendental Philosophy and Naturalism*, Oxford, Oxford University Press: 74-95
- Street, S., 2012, "Coming to Terms with Contingency: Humean Constructivism about Practical Reason", in J. Lenman and Y. Shemmer (eds.), *Constructivism in Practical Philosophy*, Oxford: Oxford University Press: 40-59
- Stroud, B., 1968, "Transcendental Arguments", *The Journal of Philosophy* 65(9): 241-56
- Tiffany, E., 2012, "Why Be an Agent?", *Australasian Journal of Philosophy* 90(2): 223-33
- Velleman, J.D., 2009, *How We Get Along*, Cambridge: Cambridge University Press
- Williams, B., 1981a, "Persons, character and morality", reprinted in his *Moral Luck*, Cambridge: Cambridge University Press: 1-19. Originally published in A.O. Rorty (ed.), 1976, *The Identities of Persons*, Berkley: University of California Press: 197-216
- , 1981b, "Internal and External Reasons", reprinted in his *Moral Luck*, Cambridge: Cambridge University Press: 101-13. Originally published in Ross

Harrison (ed.), 1980, *Rational Action*, Cambridge: Cambridge University Press: 17-29

----, 1995, "Internal Reasons and the Obscurity of Blame", reprinted in his *Making Sense of Humanity*, Cambridge: Cambridge University Press: 35-45. Originally published in 1989, *Logos* 10 (Santa Clara University, CA): 1-11