This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Humean Constitutivism:
A Desire-Based Account of Rational Agency and the Foundations of Morality

by

Giles H. Howdle

PhD in Philosophy
The University of Edinburgh
2022
Declaration of Authorship

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification.

Giles Howdle 30/11/22
To my parents, Elizabeth and Mark.
Acknowledgements

I am deeply grateful for the opportunity to immerse myself in a topic I find so compelling. A doctoral dissertation is never created in isolation, and I wish to express my heartfelt appreciation to the many people who helped guide and support me along the way.

My profound thanks to my supervisor, Michael Ridge, for the many thoughtful discussions and invaluable feedback on countless drafts over the years. Your insights and generosity with your time made this work immeasurably better. I wish to thank my secondary supervisor, Guy Fletcher, for the sage counsel and your generous and insightful assistance in refining my ideas.

I am incredibly appreciative of my examiners, Matthew Silverstein and Michael Cholbi, for their discerning and constructive critique of this dissertation. Our rich discussion gave me much to think about for future areas of inquiry.

To my friends, especially Luke, George, Fausto, Tom and Pavel - thank you for the camaraderie, encouragement, and welcomed distractions along the way.

Most of all, my heartfelt gratitude goes to my family for their endless love and support. To my wife Vera, thank you for your kindness, patience, and unwavering faith in me. To my parents, Elizabeth and Mark, thank you for always cheering me on and enabling me to follow my passions. I could not have done this without all of you.
Why should I be moral?’ The moral sceptic is one of philosophy’s trickiest characters. Socrates’ answer seems to have left Plato so unsatisfied that he wrote *The Republic* in part to provide a better answer. This thesis provides a new answer to the moral sceptic and thereby vindicates the robust normativity of morality—its significance in determining what we ought to do.

Broadly, three philosophical traditions have given three answers to moral scepticism. Kantians (arguably in the tradition of Plato) claim that morality is a part of rationality—that moral norms are rational norms and hence that any rational agent ought to be moral. Aristotelians claim that we ought to be moral because it is good for us given our human nature, as social and rational animals. Those in the Humean tradition claim we ought to be moral because (or if) we *care* about other people, or acting morally would satisfy some desire.

In this thesis, I offer my own answer to the moral sceptic, combining elements of these three traditions. I take the fundamental account of normativity—of why we ought to do what we ought to do—from the Humean; I claim an agent’s reasons are explained by his desires, his commitments, by features of his conative psychology broadly. I supplement this account with the Kantian argumentative strategy of constitutivism according to which the nature of rational agency is such that rational agents have reasons to be moral. The slightly Aristotelian element of my account is a more substantial analysis of rational agency which highlights the *linguistic* nature of the kind of thought that makes rational agency possible. I argue that we are committed to moral norms as competent language users, and hence as rational agents. In short, my thesis plays a Kantian tune on a Humean instrument.
Abstract

The thesis combines the Humean theory of normative practical reasons with constitutivism. Humeanism claims, roughly, that an agent’s reasons are explained by his desires. Constitutivism claims that the foundations of moral normativity can be found in the constitutive features of agency. This combination, ‘Humean constitutivism’ as I call it, claims there is some desire-like mental state, the possession of which is partly constitutive of rational agency, which explains the normativity of morality.

I first argue that the union of constitutivism and Humeanism makes for an attractive marriage. As with any good relationship, the two bring out the best in one another. Humeanism offers the constitutivist a compelling account of how we get from the ‘is’ of the nature of rational agency to the ‘ought’ of normative conclusions. In doing so, it permits the constitutivist to circumvent the influential ‘shmagency’ objection. The constitutivist strategy provides the Humean with the resources to accommodate moral rationalism and hence answer sceptical challenges to morality’s practical normative authority.

Having established its attractiveness, I then defend the truth of Humean constitutivism. I argue that there are desire-like mental states (constitutive commitments), the possession of which is partly constitutive of agency. This is because rational agency involves distinctively linguistic forms of cognition—dwelling in something like Sellars’ ‘logical space of reasons.’ Engaging in this linguistic form of cognition, in turn, requires us to take on certain second-personal practical commitments to our (hypothetical) interlocutors. Hence, all agents necessarily have these constitutive commitments. Piggy-backing on Darwall’s derivation of contractualist moral norms from the constitutive presuppositions of the second-person standpoint, I claim that the Humean constitutivist can vindicate the intuitive existence, modal status, universality, and special weight of (many) moral reasons.
Table of Contents

Lay Summary..................................................................................................................................................4
Abstract..........................................................................................................................................................5
Table of Contents .........................................................................................................................................6
Humean Constitutivism: An Introduction to the Thesis.................................................................8
1. Preface.........................................................................................................................................................9
2. Three Key Concepts..................................................................................................................................10
3. Chapter Guide .........................................................................................................................................12
Chapter One: A Humean Solution to the Shmagency Objection..................................................14
1. Introduction..............................................................................................................................................15
2. Modest (Humean) Constitutivism ..........................................................................................................15
3. The Basic Shmagency Objection ...........................................................................................................18
4. Chess and Agency, Shmess and Shmagency ....................................................................................22
5. Insufficiency ............................................................................................................................................25
6. Implausibility ..........................................................................................................................................32
7. Irrelevance ..............................................................................................................................................43
8. Irrelevance and Insufficiency Objections to Humeanism ...........................................................48
9. Summary and Conclusion .....................................................................................................................56
Chapter Two: A Constitutivist Defence of Conservative Humeanism ..................................58
1. Introduction..............................................................................................................................................59
2. Humeanism .............................................................................................................................................60
3. The Amoralism Objection ....................................................................................................................63
4. Answering the Amoralism Objection ....................................................................................................68
5. Conclusion ............................................................................................................................................83
Chapter Three: The Weight of Moral Reasons........................................................................84
1. Introduction..............................................................................................................................................85
2. The Intuitive Weight of Moral Reasons ................................................................................................86
3. Humeanism and the Weight of Moral Reasons ..................................................................................107
4. The Nature of Constitutive Commitments ..........................................................................................117
5. Conclusion and Stock-Taking .............................................................................................................121
Stock-Taking ..............................................................................................................................................122
Chapter Four: Second-Person Constitutivism and the Optionality Challenge..............125
1. Introduction .............................................................................................................................................125
2. The Promise of Second-Person Constitutivism ................................................................................126
3. A Problem: The Optionality of the Second-Person Standpoint ..................................................133
4. Darwall’s Solutions..............................................................................................................................136
5. Korsgaard’s Agential Solution ..........................................................................................................154
6. Conclusion and Outline of my Proposed Solution ....................................................................157
Chapter Five: Linguistic Constitutivism ................................................................................159
1. Introduction ............................................................................................................................................... 160
2. The Promise of Linguistic Constitutivism ............................................................................................. 163
3. Language, Morality, and the Second-Person Standpoint ................................................................. 169
4. Assertion and Language ......................................................................................................................... 201
5. Language and Agency .......................................................................................................................... 211
6. Objections: Condemning the Rogue’s Gallery ................................................................................... 228
7. Conclusion ............................................................................................................................................... 234

**Concluding Remarks** .................................................................................................................................. 237
1. Answering the Moral Sceptic .................................................................................................................. 237
2. Postscript: What about Animals? .......................................................................................................... 239

**Bibliography** ............................................................................................................................................... 241
Humean Constitutivism: An Introduction to the Thesis

It might turn out that when we properly think about it, we shall find that we are committed to an ethical life, merely because we are rational agents. [...] If so] there is what I have called an Archimedean point: something to which even the amoralist or the skeptic is committed but which, properly thought through, will show us that he is irrational, or unreasonable, or at any rate mistaken.

– Bernard Williams, Ethics and the Limits of Philosophy (1985, 33)

1. Preface

‘Why should I be moral?’ This thesis is an attempt to answer the moral sceptic by uncovering Williams’ Archimedean point. The answer it gives the sceptic is the one Williams considers; ‘You are committed to an ethical life merely because you are an agent.’

Seen through the lens of existing philosophical theories and argumentative strategies, this thesis combines substantial elements of Humean and Kantian thought about practical normativity to deliver a vindication of the robust normativity of morality. Specifically, it combines the Humean theory of reasons with the Kantian strategy of constitutivism, yielding what I label ‘Humean constitutivism’.

In the first half of the thesis, the first three chapters, I argue that this combination of views offers profound philosophical benefits to both parties. Broadly, Humeanism offers the constitutivist a much-needed account of how constitutive features of agency could have normative entailments (how we get from the ‘is’ of the constitutive features of agency to the ‘ought’ of (moral) norms and reasons) while constitutivism offers the Humean the resources to conserve ordinary moral intuitions and accommodate moral rationalism.

In the second half of the thesis, chapters four and five, I argue that Humean constitutivism is not just appealing, but that it is true. I argue that there are desire-like mental
states (constitutive commitments), the possession of which is partly constitutive of agency. This is so because rational agency involves a distinctively linguistic form of cognition—dwelling in what Sellars labelled ‘the logical space of reasons.’ And this linguistic form of cognition, in turn, requires us to take on certain second-personal practical commitments to our (hypothetical) interlocutors. Hence, all agents necessarily have these constitutive commitments. Piggy-backing on Darwall’s derivation of contractualist moral norms and principles from the constitutive presuppositions and commitments of the second-person standpoint, I claim that the Humean constitutivist can vindicate the intuitive existence, modal status, universality, and special weight of (many) moral reasons.

2. Three Key Concepts

I will generally define key terms and ideas as they arise. But it would be helpful to clarify three key concepts from the outset. These are constitutivism, Humeanism, and moral rationalism.

Constitutivism, in its broadest interpretation, claims the normativity of (some) normative facts follows from the constitutive features of some capacity, practice, or kind. But the kind of constitutivism which is most discussed, and my focus, is Kantian in two respects. First, it claims the normative facts which the constitutivist account vindicates are moral. Second, it claims that constitutive features of rational agency ground these normative facts. Hence, this agential and moral constitutivism (henceforth simply ‘constitutivism’) claims that agents have moral reasons merely in virtue of the fact they are rational agents. The justificatory foundation of moral facts is to be found in the nature of rational agency.

---

1 Smith offers a particularly ecumenical characterisation along these lines (Smith 2017, 371)
The Humean theory of normative practical reasons (henceforth: Humeanism) claims that our reasons are explained by our desires.\textsuperscript{2} I have in mind a particular version of this theory developed by Schroeder in \textit{Slaves of the Passions}.\textsuperscript{3}

**Reason.** For R to be a reason for X to do A is for there to be some p such that X has a desire whose object is p, and the truth of R is part of what explains why X’s doing A promotes p.

(Schroeder 2007, 59)

I explore Humeanism in detail in the second chapter. But the key point is that the account claims that all our normative reasons are ultimately grounded in our desires.

The final term I want to mention up-front is ‘moral rationalism’. This thesis encapsulates orthodox intuitions about the existence, modal status, and rational weight of our moral reasons. I provide more detail in the third chapter where I go into more detail, but for most of the thesis, I get by with the following characterisation, courtesy of Lord and Plunkett:

**Moral Rationalism.** Necessarily, if there is a moral reason for A to phi, then there is a normative reason for A to \textit{phi} and, at least sometimes, normative reasons provided by moral reasons have significant weight. (Lord and Plunkett 2017, 330)

This thesis combines Humeanism and constitutivism and thereby vindicates moral rationalism. How do I combine these ‘rival’ metaethical theories? As I explain in the first chapter, I interpret constitutivism \textit{modestly}, as a strategy for addressing metaethical challenges to the normativity of certain reasons/norms rather than as a full-blown metaethical theory. Interpreted in this way, constitutivism can be (indeed, should be) combined with traditional metaethical views. I argue that the union of constitutivism and Humeanism makes for an attractive marriage. As with any good relationship, the two bring out the best in one

\textsuperscript{2} This was most likely not Hume’s own view.

\textsuperscript{3} Schroeder labels this specific version of Humeanism ‘hypotheticalism’.
another. Humeanism permits the constitutivist to circumvent the shmagency objection, while constitutivism allows the Humean to accommodate moral rationalism. In doing so, my thesis answers sceptical challenges to morality’s normative authority.

3. Chapter Guide

In Chapter One, ‘A Humean Solution to the Shmagency Objection’, I explain how constitutivism of a certain kind can circumvent Enoch’s influential ‘agency-shmagency’ objection. I interpret constitutivism modestly as an argumentative strategy for dealing with sceptical challenges to the robust normativity of a particular domain, rather than ambitiously as a full-blown metaethical view. I then claim it is naturally and productively supplemented by the Humean theory of reasons, a reductive account of normative reasons which provides the bridge that the constitutivist needs between the ‘is’ of the constitutive features of agency and the ‘ought’ of normative facts. Humean constitutivism claims that the possession of certain practical commitments (‘constitutive commitments’) is a constitutive requirement of rational agency.

Chapters Two and Three examine the benefits constitutivism provides to the Humean. Specifically, they show how constitutivism provides the resources for the Humean to stay extensionally conservative about moral reasons—for the Humean to accommodate moral rationalism. In the second chapter, ‘A Constitutivist Defence of Conservative Humeanism’, I argue, primarily contra Schroeder (2007), that a Humean constitutivist strategy can provide Humeans with a plausible, non-revisionary account of both the existence and, more interestingly, the modal status of moral reasons. In the third chapter, ‘The Weight of Moral Reasons’, I continue the task of reconciling Humeanism and moral rationalism, focusing on the weight of moral reasons. I first explore what the intuitive view of moral reasons’ weights is, concluding that a strong version of moral rationalism best reflects common sense. Then,
drawing on the resources of constitutivism, I show how Humeanism can accommodate strong moral rationalism.

In the fourth chapter, ‘Second-Person Constitutivism and the Optionality Objection’, I explore the derivation of substantive moral norms and reasons with content from the constitutive features of agency. I lean heavily on Darwall’s second-person constitutivism, as I frame it. I present his argument for the conclusion that something like the moral norms we find in Scanlon’s contractualism, or ‘morality as equal accountability’, in Darwall’s terms, follow from the necessary presuppositions we make whenever we occupy the ‘second-person standpoint’. But I also raise a problem for Darwall’s second-person constitutivism: the optionality challenge. The normativity of morality as equal accountability cannot be secured, at least not with the right modal robustness, because it seems perfectly optional to engage second-personally with others, and hence it seems perfectly optional to be subject to moral norms. I consider and largely reject possible lines of response in Darwall and Korsgaard’s work, and conclude that we need more secure foundations for the second-person standpoint if we are to vindicate the normativity of the norms of morality as equal accountability.

In the fifth and final substantive chapter, ‘Linguistic Constitutivism’, I cap the thesis off by arguing for the non-obvious conclusion that these much-discussed constitutive commitments exist. The main steps of that argument are as follows. In the first half of the chapter, I defend linguistic constitutivism, piggybacking on Darwall’s second-person constitutivism. Unlike Darwall, however, I claim that being a competent language user requires occupying the second-person standpoint, and hence that all language users are implicitly committed to morality as equal accountability. I also claim that linguistic competence requires the adoption of a commitment to giving second-personal claims some weight in deliberation. In the second half of the chapter, I defend a linguistic account of rational agency, according to which full-blooded rational agency is made possible by forms
of thought which require language. Hence, all full-blooded rational agents are language users and they are subject to the same moral norms and have the same moral reasons as language users. This answers the optionality objection. I conclude that there are constitutive commitments which explain the robust and universal normativity of moral reasons.
Chapter One: A Humean Solution to the Shmagency Objection

1. Introduction

In this chapter, I give the first volley of my case for combining constitutivism with Humeanism about normative practical reasons. I argue that supplementing constitutivism with the Humean theory of normative reasons provides the constitutivist with the resources to respond to Enoch’s (2006, 2011) influential ‘shmagency’ objection. The next two chapters explain the benefits that union with constitutivism brings to the Humean.

This chapter consists of seven sections. Section 2 explains the modest interpretation of constitutivism and how it can be combined with Humeanism about reasons. Section 3 presents the basic version of the shmagency objection to constitutivism and a basic modest constitutivist response. Section 4 explains my general strategy of response to Enoch’s (2011) restatement of the shmagency objection. Unlike most constitutivists, I accept Enoch’s chess-agency analogy; I don’t appeal to any special features of agency such as its inescapability to explain agency’s reason-providingness. In sections 5, 6, and 7, I explain how Humean constitutivism can respond to three objections Enoch makes to the kind of solution I favour: that the solution is insufficient (section 5), implausible (section 6), and irrelevant (section 7). Finally, in section 8, I respond to some shmagency-related objections to my appeal to Humeanism.

2. Modest (Humean) Constitutivism

Constitutivists claim that we can derive the normativity of some reasons or norms from what is constitutive of agency. But what does this derivation give us, philosophically speaking? We might think of constitutivism ambitiously, as providing a theory or analysis of
normative practical reasons. In this reading, constitutivism claims that *what it is* for an agent to have some reason to act just *is* for the reason to be derivable from something constitutive of agency. In providing an analysis of normative reasons, ambitious constitutivism is a rival view to traditional metaethical theories. According to the ambitious interpretation of constitutivism, the following biconditional holds; R ‘is a reason for action *if and only if* it somehow can be derived from something constitutive of agency.’ (Ridge 2017, 2952) However, a chorus of philosophers have harmonised on the thought that constitutivism (or its cousin, constructivism) may have to curtail these metaethical ambitions if it is to be at all plausible (Hussain and Shah 2006, Silverstein 2012, Ridge 2012).

Only recently has there been serious philosophical consideration of what a more metaethically *modest* constitutivism would look like, and if it would be an interesting view in its own right (Silverstein 2016, Ridge 2017, Ferrero 2019). So, stripped of its lofty metaethical ambitions, what *is* constitutivism? Modest constitutivism claims merely that R is a reason for action *if* R can be derived from something constitutive of agency (*if* rather than *if and only if*). Modest constitutivism, unlike its more famous brother, allows for the existence of ‘reasons which *cannot* be derived from what is constitutive of agency.’ (Ridge 2017, 2952)

Modest constitutivism does not attempt to provide a *general theory* of normative reasons. The promise of modest constitutivism is a new way of deriving and vindicating normative reasons. This, of course, is still a significant promise. Vindicating the normativity of various things is one of the oldest and most seemingly intractable philosophical problems. Plato’s *Republic* is, roughly, an attempt to vindicate the normativity of reasons of justice. More broadly, the moral sceptic has been a continual challenge to moral philosophy. ‘Sure, X, Y, and Z are demanded by morality,’ the sceptic says, ‘but what does that show about what I *really* ought to do?’ It is a notoriously difficult question to answer convincing. Modest constitutivism, with its promise of a novel method of deriving normative reasons, is, then,
best understood as a ‘strategy for meeting sceptical challenges to the authority of some norm or end.’ (Ridge 2017, 2921)

Since constitutivism is not, on its modest interpretation, a freestanding theory of normative reasons, it’s possible, even desirable, to combine it with one of the traditional metaethical candidates (Silverstein 2012, 8-9, 2016, 235-237, Ridge 2017). So who will take modest constitutivism to the ball? Alternatives have been suggested, but an obvious candidate, one that I think would make a happy couple with constitutivism, is the ‘subjectivist’ or ‘Humean’ theory of normative practical reasons (henceforth simply ‘Humeanism’).  

Why is this a promising partnership? An initial reason, and I will give more in the rest of this dissertation, is philosophical parsimony. The conceptual pieces of the puzzle fit neatly together. Humeanism claims (roughly) that an agent has a reason to X if and only if X-ing promotes some desire of the agent (Schroeder 2007, 59). So, if there is some desire-like state(s) of mind the possession of which is partly constitutive of agency, then we could derive normative reasons for action from the agent (necessarily) having those states of mind. This is, as I will explain in the next chapter, also a tempting marriage for the Humean, who is traditionally thought to struggle to explain the existence of certain normative reasons, e.g. universal normative reasons to act morally. A novel method for deriving normative reasons is of obvious appeal to the Humean. But the advantage of Humean constitutivism I want to

4 This is not a totally novel idea. Silverstein suggests combining constitutivism with reductive accounts of reasons (2016, 235-237) and Ridge also considers this option (Ridge 2017, 2596-7), though he ultimately argues in favour of an alternative combination: expressivism and constitutivism.
demonstrate is this; it allows the constitutivist to answer (or perhaps dodge) the influential ‘shmagency’ objection.

In sum, then, I propose that we follow some recent precedents\(^5\) and interpret constitutivism merely as ‘a method of deriving normative reasons’ (Ridge 2017, 2921) rather than a full-blown metaethical theory. I’ll argue that the union of constitutivism and Humeanism about practical normative reasons, a.k.a. metaethical subjectivism, would allow the constitutivist to provide an adequate response to the shmagency objection.

3. The Basic Shmagency Objection

The most influential and well-discussed objection to constitutivism is probably Enoch’s (2006, 2011) ‘agency, shmagency’ or ‘shmagency’ objection. In this section, I introduce the basic form of the objection and explain the basic modest constitutivist response. I also note a key limitation of the basic response.

3.1. The Basic Objection

Constitutivists claim that certain normative truths follow from what is constitutive of (generally) agency or action. Hence, they think that an agent’s *being an agent* entails certain normative truths, e.g. facts about what they ought to do. The versions of constitutivism Enoch

---

\(^5\) I am thinking of Silverstein (2012, 2016), Ridge (2017) and Ferrero (2019) who distinguish, in their own ways, the ambitious and modest interpretations of constitutivism, and Tubert (2010), Lindeman (2017), and perhaps Smith (2015, 2017) who (in my view) define constitutivism such that it obviously can be interpreted modestly, to ground norms/reasons in specific, limited domains. Arguably, Rawls could be understood as offering a modest constitutivist account that pertains merely to basic political norms.
considers specify that the constitutive aim of agency is self-constitution (Korsgaard’s view) or self-understanding (Velleman’s view). Since then, further alternatives have been proposed. For example, according to Katsafanas’s Nietzschean constitutivism (2013), the constitutive aim of action is both agential activity and power. The shmagency objection targets all of these views.

Enoch introduces the shmagency objection to constitutivism in the voice of a moral sceptic:

Classify my bodily movements and indeed me as you like. Perhaps I cannot be classified as an agent without aiming to constitute myself. But why should I be an agent? […] I am perfectly happy being a shmagent—a nonagent who is very similar to agents but who lacks the aim (constitutive of agency but not of shmagency) of self-constitution. I am perfectly happy performing shmactions—nonaction events that are very similar to actions but that lack the aim (constitutive of actions but not of shmactions) of self constitution. (Enoch 2006, 179)

Even accepting that there are constitutive features of agency, the shmagency objection suggests that constitutivists will fail to deliver on the promised vindication of normative claims. To succeed in vindicating these normative claims, thinks Enoch, we would need good reason to be agents in the first place (rather than, say, shmagents). ‘Of course,’ he says, ‘if there were some independent reason to be an agent (for instance, rather than a shmagent), or to perform actions, this objection would go away.’ But constitutivists do not provide such reasons and, he thinks, they cannot; ‘The price would be too high’ claims Enoch, ‘for such an independent reason [to be an agent]—one not accounted for by the constitutivist story, but rather presupposed by it—would make it impossible for constitutivism to be the whole, or most foundational, account of normativity’. (Enoch 2011, 209).

To summarise, the basic argument of the constitutivist is:
1. If C is a constitutive feature of agency, then an agent has normative reasons to do whatever follows from C

2. C is a constitutive feature of agency

3. (Conclusion) An agent has normative reasons to do what follows from C

The shmagency objection assumes (2) and denies (1).

Implicitly, the objection adds a further requirement to (1): that the agent ought to be an agent. So we can, I think safely, take Enoch to accept the following: if C is a constitutive feature of agency, then an agent has normative reason to do whatever follows from C only if the agent has a normative reason to be an agent in the first place. Then, and only then, does normativity flow through to whatever follows from C. I’ll call this view ‘trickle-down normativity’.

So why can’t the constitutivist give Enoch what he wants—a reason to be an agent? There is a threat of circularity. The (ambitious) constitutivist thinks that reasons are what follows from what is constitutive of agency. Any reason we appeal to, be an agent, must both follow from the constitutive features of agency and explain why we ought to be agents in the first place. And this seems viciously circular. A non-circular answer to the question, ‘Why be an agent?’ will have to draw on an ‘unconstructed’ source of normativity. This is why answering the shmagency objection would, Enoch thinks, force constitutivists would have to abandon their claim ‘to be the whole, or most foundational, account of normativity,’ (Enoch 2011, 209)

3.2. The Basic Modest Response

Appealing to non-constitutively derived reasons to be an agent in the first place would involve admitting that ‘constitutivism [is not] the whole, or most foundational, account
of normativity.’ But this is acceptable to a modest constitutivist. He only wants the constitutivist story to deliver an account of some reasons’ normativity. Going modest is a perfectly viable option for the constitutivist (see section 2) as he can appeal to non-constitutively-derived reasons. Of course, the modest constitutivist still would need to answer Enoch’s question—why be an agent, rather than a shmagent? But this is a rather different challenge, and, prima facie, a far less troubling one. The modest constitutivist is entitled to draw on whatever (non-constitutive) reasons he wants to answer it.

As dialectically powerful as the basic modest constitutivist response is, I suspect that this way of dodging the shmagency objection leaves a central question unanswered: the sense in which constitutive aims/features are understood to be reason-providing. Katsafanas (2013, 48-9) makes a helpful distinction here. Constitutive features can be thought either to merely ‘transfer’ normativity (‘if you have reason to participate in A, then the constitutive aim of A is reason-providing’) or they can be thought to be ‘originative’ of reasons (‘if you participate

__________________________________

6 Enoch’s explicit target was the ambitious interpretation of constitutivism which existed at the time of his 2006 paper, defended in the work of Velleman (1989, 2000) and Korsgaard (1996, 2002), so it makes sense that he does not consider the modest option.

7 Silverstein (2016) identifies this potential ‘gap’ in the constitutivist strategy between the teleological aim of agency and the normative conclusions which supposedly follow (i.e. what reasons we agents have given the constitutive e.g. aim of agency). Silverstein’s suggested solution—that constitutivism should be ‘supplemented with a reductive metanormative account of normativity or reasons for action’ is the solution I am pursuing here (Silverstein 2016, 215-6). Silverstein outlines a reductive account according to which ‘reasons for action just are considerations that bear on practical reasoning’ (Silverstein 2016, 235). I suspect the Humean theory of reasons would fall under this description.
in an activity A, then the constitutive aim of A is reason-providing’). If all the modest constitutivist is claiming is that the constitutive features of agency transfer the normativity of other non-constitutive reasons to some other reasons, then the view seems a lot less interesting. That would yield only a partial vindication of the reasons we have given the constitutive features of agency. For the remainder of the chapter, I shall argue that the Humean constitutivist can do more than this. He can explain how constitutive features are originative of reasons, and how they are themselves reason-providing, without needing to appeal to other reasons.

4. Chess and Agency, Shmess and Shmagency

Few discussions of constitutivism fail to mention chess, and this one will be no exception. Many philosophers have noted the usefulness of competitive games such as chess for thinking about the constitutivist proposal. In fact, each of Enoch’s three objections to aim constitutivism is based on an analogy between chess and agency (2006, esp.185-6 and 2011). Enoch claims that just as chess has a constitutive aim but chess-related reasons might not be normative, agency might have a constitutive aim but our agency-related reasons might not be normative. The typical response to the shmagency objection is to reject the analogy between agency and chess, to claim that agency is special in some way. A lot of ink has been spilt in that debate, and I will avoid spilling more. Instead, I’ll argue that the chess-agency analogy is, in fact, grist to the (Humean) constitutivist’s mill.

---

4.1. Rejecting the Chess-Agency Analogy

Perhaps agency is relevantly disanalogous with chess in a way that entails that agency’s constitutive aims/norms are always normative and generate normative reasons for action (while those of chess, as Enoch claims, are not and do not). Rosati, for example, claims that the constitutive norms of agency are special in that, unlike those of chess, they are ‘self-vindicating’ (Rosati 2016, 197). Korsgaard and Velleman both claim agency is special because it is, unlike chess, inescapable. Korsgaard defends (what has been labelled by others as) agency’s ‘natural inescapability’, describing agency as ‘our plight’, that we are ‘condemned to choice and action’ (Korsgaard 2009, 1-2). We can’t simply step outside of agency and ask for reasons to take it seriously. Velleman (2009), Ferrero (2009), and Silverstein (2015) dispute that agency is naturally inescapable. After all, one can kill oneself, self-inflict serious brain damage, or simply fall asleep. But Velleman and Ferrero agree with Korsgaard that agency is inescapable in a relevant sense; they describe it as possessing dialectical inescapability. Velleman writes:

To ask “Why should I have the aim of making sense?” Is to reveal that you already have it. If you don’t seek to do what makes sense, then you are not in the business of practical reasoning, and so you cannot demand reasons for acting or aiming (Velleman 2009, 137).\(^9\)

Ferrero makes a similar point:

The inescapability of agency, however, shows that there is no standpoint external to agency that the shmagent could occupy and from which he could launch his challenge. (Ferrero 2009, 311)

\(^9\) On Velleman’s account, the constitutive aim of action is making sense or self-understanding.
Whether these responses based on agency’s special qualities are adequate or not is a matter of significant ongoing disagreement. But, as I hinted earlier, I think the Humean constitutivist can avoid being thrown into philosophical trench warfare. Constitutivists don’t need to reject the analogy between chess and agency. Indeed, chess is a helpful model for understanding how the Humean constitutivist secures normative conclusions from constitutive features. That dialectical point noted, let’s turn to Enoch’s three objections.

4.2. Embracing the Chess-Agency Analogy

The shmagency objection, at its heart is simply this: to draw normative conclusions, it is insufficient that you are an agent or that you are acting—you also need a reason to be an agent, or perform actions. I’ve called this view trickle-down normativity. But there are some activities where it does seem like we can draw normative conclusions merely from the fact that you’re engaged in the activity—competitive games such as chess, for instance. Hopefully, something similar can be said for agency or action; merely being an agent, or

---

10 Enoch responds in detail in Enoch 2011. For further defences of agency’s specialness see Tubert (2010), Ferrero (2018), and Silverstein (2015). Silverstein argues that it is not agency’s supposed (though illusory) inescapability that makes its constitutive norms normative, but the fact that ‘agency stands in a special relation to practical reasoning and thus to reasons and normativity.’ (Silverstein 2015, 1140). The shmagency question—why be an agent?—is only intelligible from ‘within the standpoint of agency’ as only agents are ‘in the market for reasons for action’ (Silverstein 2015, 1136). This is something special about agency—‘Chess is not like this’ (Silverstein 2015, 1140). For further developments of the shmagency objection see Tiffany (2012) and Leffler (2019).
being in the business of action, suffices to entail that one has normative reasons to do certain things. This (with some adjustments) is the kind of solution I favour.

Let’s return to chess for the moment. The fact that someone is playing chess does seem sufficient to give him reasons to e.g. attempt to checkmate his opponents. These might be weak pro tanto normative reasons rather than conclusive normative reasons, but they still seem genuinely normative. After all, the constitutive aim of the game of chess is winning by checkmating one’s opponent. And that’s not all that playing chess entails, because a reason to phi gives instrumental reasons to do whatever else it takes to phi, or whatever raises the chances of phi-ing. So, having a reason to achieve checkmate provides instrumental reasons to develop one’s pieces, force advantageous exchanges, control the centre of the board, etc.

In the remainder of this chapter, I explain how a solution along these lines can address three lines of objection Enoch provides in ‘Shmagency Revisited’ (2011). In the process, I also explain how constitutive features could be genuinely reason-providing or originative, to use Katsafanas’s term. The first of these objections is that an agent’s engaging in an activity is insufficient to derive any normative conclusions. The second is that it is implausible that agency has constitutive aims, ends, or commitments. The third is that even accepting that agency has constitutive commitments, these commitments are normatively irrelevant. Let’s start with the first—that an agent’s participation in an activity is insufficient to derive any normative conclusions.

5. Insufficiency

Enoch’s insufficiency objection comes in the form of a supposed counterexample. Imagine several versions of chess with different rules, or criteria for winning. One of these versions, let’s call it ‘shmess’, is won by achieving checkmate in an even number of moves. In ‘sufficiently early stages’ of chess, claims Enoch, you are also playing shmess because you
are ‘satisfying some relevant descriptive criteria’ for shmess-playing (Enoch 2011, 211). Here’s the problem. If you’re playing shmess, then the constitutivist must say that this is sufficient to provide the shmess player with shmess-related reasons. The fact that you are, accidentally, playing shmess suffices to provide you reasons to ‘shmeckmate’ your opponent (checkmate them in an even number of moves), to keep track of the number of moves in the game, etc. But this is an absurd conclusion. Intuitively, you don’t have any such reasons. So we have to abandon the idea that simply engaging in an activity with a constitutive aim (whether chess or agency) provides sufficient grounds for deriving normative conclusions. You also need a reason to be engaging in that activity, rather than some other version of the activity with a slightly different set of constitutive norms or aims. In other words, constitutive aims and commitments are merely transferative, and not originative, of normative reasons.

5.1. Two Kinds of Constitutive Aim

Constitutivists of different stripes have suggested various ways to get from the existence of certain constitutive features of agency/action to specific normative conclusions. ‘Commitment constitutivism’, as I’ll call it, follows the general strategy of ‘aim’ constitutivists, such as Velleman (2009) and Katsafanas (2013) who claim that agency or action has a constitutive aim. They then seek to derive the normativity of some reasons/norms from this aim. It is aim constitutivism, in particular, that Enoch has in his sights. But to answer Enoch’s insufficiency objection, aim constitutivists first need to distinguish two senses in which a practice or activity has a constitutive aim. Not all activities that have constitutive aims can respond to the insufficiency objection, but chess can.

Both medicine and chess have constitutive aims—health and winning the game respectively. But these constitutive aims are importantly different in kind. Imagine a cynical doctor who participates in the practice of medicine. He really is practising medicine. But he
does so only to receive his paycheque and not to provide health to his patients. In other words, the cynical doctor is practising his profession without adopting the aim of medicine as one of his own. So some activities possess a constitutive aim that does not necessitate that a participant in the activity adopts the aim as her own. If an activity has only this kind of constitutive aim then the normativity of whatever follows from what is constitutive of the activity is going to remain disputable. ‘Sure, I’m a participant in the activity and I acknowledge that the activity has a constitutive aim,’ the sceptic can say, ‘but why should I take the constitutive aim of the activity to be one of my aims? Why should the standards of success of the activity, given the aim, guide my action?’ At the very least, it’s deeply uncertain that we can derive anything normative from activities with only this kind of aim.

But I think there is another sense in which activities can have constitutive aims, and the constitutivist is on much more promising ground if agency has a constitutive aim of this sort. Think of competitive games or sports. Participation in any competitive sport or game—earning the status as a ‘player’—requires participants to adopt the end of winning as one of their own. My intuition is that someone who superficially appears to be playing chess but does not aim to win—does not take checkmating his opponent to be one of his ends—is not really playing chess. He’s doing something other than playing chess.11

The key point is simply this. The kind of constitutive aim we find in certain activities, such as chess, is such that the participant in the activity must, to count as a participant, adopt the constitutive aim themselves.

11This is a controversial view. I will develop this claim in more detail shortly.
5.2. From Constitutive Commitments to Normative Conclusions

What do I mean when I say that a chess player must, as a constitutive matter, ‘adopt an aim?’ A natural way to think of adopting an aim as one’s own is taking on some kind of practical commitment to realising or promoting it: having a (perhaps special) pro-attitude toward the object of the aim/commitment.\(^{12}\) So constitutive commitments are a class of pro-attitudes: desire-like states of mind, i.e. mental states with a ‘mind-to-world’ direction of fit.\(^{13}\) In other words, it’s the kind of attitude from which a Humean would be happy to derive normative reasons in the *originative* sense, not merely in the *transferring* sense.

Once we know an activity has a constitutive commitment, it is straightforward to derive normative conclusions. To count as playing chess or engaging in any competitive game, you must—as a constitutive matter—take on the aim of winning as one of your ends. This isn’t to say you must care about winning a great deal, or prioritise winning over other

\(^{12}\) My reasons for straying from the traditional Humean language of desire to talk of commitment will become clearer in later chapters.

\(^{13}\) Though she doesn’t use the term ‘direction of fit’, Anscombe famously draws this distinction in Intention (2000, 53). She highlights the difference between a man going shopping with a shopping list and a detective following him with a list describing what the man is buying. The shopper’s list has ‘mind-to-world’ direction of fit (or ‘word-to-world’), whereas the detective’s has ‘world-to-mind’ (or ‘world-to-word’) direction of fit. The shopper would make an error by failing to get the right items, as dictated by the shopping list. The detective would make an error by failing to have the list conform to the way the world is. The same kind of distinction can be drawn between different kinds of mental states, beliefs being the paradigmatic ‘world-to-mind’ states, and desires being the paradigmatic ‘mind-to-world’ states. Some think that there are mental states (‘besires’) which have both directions of fit (Swartzer, 2013)
aims you might have, but it does mean that you take winning the game to be one of your ends. Then we can develop a specific account of chess-related *pro tanto* normative reasons *given your commitment*, in just the same way a Humean would derive a reason to go to the party from the agent having taken on a practical commitment to socialise more.

Perhaps it’s not yet clear how the introduction of constitutive commitments addresses Enoch’s insufficiency objection, which claims that the fact that you are participating in an activity is *insufficient to draw normative conclusions*. To see how constitutive commitments help, let’s return to chess. What makes participation in chess *suffice* to provide chess-related reasons is not *the mere fact that you are playing chess*. On this much, Enoch and I agree. Rather, what’s reason-providing here is the fact that (insofar as you’re playing chess) you’ve adopted some kind of *practical commitment to win the game*.

The same logic can be applied to agency. If there are constitutive commitments of agency then we needn’t look for further *reasons* to be an agent, a shmagent, or neither. The agent *taking on a practical commitment* is what’s doing all the reason-providing work. The fact that those commitments *are constitutive of agency* does not play any role in vindicating their normativity. (The fact that they are constitutive of agency only means that they are *universal* which might be appealing to those interested in vindicating the universal normativity of moral norms and reasons).

5.3. Do We Have Shmess-Related Reasons?

Let’s now see how this Humean-friendly *commitment constitutivism* deals with Enoch’s supposed counterexample. In the first few turns of a chess game, argues Enoch, a chess player meets the descriptive criteria for playing *shmess*. If being a participant in an activity with a constitutive aim is *sufficient* to provide agents with reasons, we would have a bunch of shmess-related reasons (such as keeping track of the number of moves) in the early
stages of playing chess. But this seems false; we don’t seem to have these shmess-related reasons in the early stages of a chess game. So we must reject the supposition that being a participant in an activity with a constitutive aim is sufficient to generate normative conclusions.

The constitutive commitment of chess is to win by checkmate. The constitutive commitment of shmess, if it has one, would presumably be to win by shmeckmate (checkmate on an even number of moves). So a shmess player should be totally indifferent, or even averse, to winning on an odd number of turns. If chess and shmess have different constitutive commitments, then whether a player counts as engaging in the activity—as playing chess or shmess—cannot simply be a matter of satisfying some relevant descriptive criteria unless those criteria include the adoption of the constitutive aim of the activity. Enoch faces a dilemma here. He can either accept my view that part of what makes someone count as a shmess player is adopting the aim of winning the game by shmeckmate, or he must claim that meeting the more minimal descriptive criteria is sufficient—that a shmess-player needn’t adopt the aim of winning the game by shmeckmate.14

Let’s first explore the latter option; what happens if we say that the shmess player doesn’t need to adopt a practical commitment? I think this opens room for the kind of response Katsafanas (2013) provides, drawing on a distinction made by Anscombe (2000) and Davidson (1980) between true descriptions and descriptions under which the action is

14I find it implausible that meeting the descriptive criteria that Enoch has in mind is sufficient for you to count as playing a game like chess or shmess. You might perform identical physical movements and simply be ‘messing about’ or ‘testing the steadiness of your hand’. But I needn’t dig my heels in on this intuition—perhaps you are doing all three at once. In any case, the constitutivist can respond to the objection.
intentional. Katsafanas notes that playing chess or shmess may be ‘a true description of a person’s action, but it is not a description under which the action is intentional […]’. Just so […] “causing a shadow to fall on the pawn” may be true descriptions of what I am doing, but they are not descriptions under which my action is intentional.’ This is significant because ‘constitutivists are attempting to show that intentional action has a constitutive aim […] so it is unsurprising that the constitutive aim of [shmess] isn’t reason-providing.’ (Katsafanas 2013, 51)

The point is this. Assuming a shmess player can play shmess without adopting any practical commitments, then shmess is (like medicine) not the kind of activity whose constitutive aim necessitates that participants take the end of the activity as one of their own. So nothing normative will follow from being a shmess player (unlike being an aiming-to-win chess player). After all, it was the constitutive practical commitments, and not the participation in the activity, that are reason-providing. So Enoch’s observation that you intuitively lack shmess-related reasons can be accommodated by the constitutivist.

Now let’s assume Enoch takes the other horn of the dilemma and agrees with me that to count as a shmess player, you need to adopt the goal of winning the game by shmeckmate. The constitutivist will simply respond that a chess player is not playing shmess in the first few turns of a chess game (as he lacks the constitutive commitment) and he, therefore, has no shmess-related reasons. So both horns of the dilemma lead to a constitutivist accommodation of Enoch’s intuition that you have no shmess-related reasons in the early stages of chess.
But this isn’t the end of the shmagency objection. Enoch anticipates a response along these lines (2011, 211-2).\textsuperscript{15} And he acknowledges that it gets the constitutivist somewhat further. \textit{If} you count as playing chess only if you take on a practical commitment to winning by checkmate, or, in Enoch’s terms, only if you ‘care about checkmating your opponent’ then he agrees it is ‘\textit{much} more plausible that merely playing chess […] suffices for your having a reason to checkmate your opponent (regardless of whether or not you have a reason to play chess).’ (Enoch 2011, 212) Ultimately, however, Enoch is unconvinced by the merits of this strategy. Let’s now turn to his two further reasons for doubt.

6. Implausibility

We’ve seen how the existence of constitutive commitments would help the Humean constitutivist vindicate normative conclusions. But, argues Enoch, it is \textit{implausible} that activities have constitutive pro-attitudes like this.\textsuperscript{16}

Enoch once again presents his objection \textit{via} chess. While, as I’ve mentioned, some constitutivists reject this analogy, I think chess is a simpler and easier case for the

\textsuperscript{15} A terminological note: Enoch refers to what I’ve called ‘adopting an aim’ and ‘undertaking a practical commitment’ as ‘caring’ about an aim. This detail isn’t significant, as caring certainly qualifies as the kind of pro-attitude the kind the Humean finds interesting.

\textsuperscript{16} This objection is not strictly an instance/development of the shmagency objection. I take the shmagency objection to concede whatever the constitutivist wants to say about what the constitutive features of agency or action are and what normative conclusions they entail, whereas this objection contests all that. Nonetheless, the solution just offered to Enoch’s insufficiency objection may make the constitutivist even more vulnerable to these kinds of objections, so it’s a good challenge to raise.
commitment constitutivist than agency. If not even chess has constitutive commitments, that provides licence for pessimism regarding the prospect that agency does. Here is Enoch’s case against chess having constitutive commitments:

Suppose I am playing chess (or, well, sort-of-playing chess) with my daughter; I obey the rules quite strictly, but I do not care who wins. Perhaps I even intentionally let her win. On the (chess-analogue of the) suggested constitutivist account, I am not really playing chess. And this seems like a huge stretch. Certainly, in common parlance we would be happy to describe the situation as one in which I am playing chess with my daughter. (Enoch 2011, 214)

At first glance, this looks like the previous ‘chess, shmess’, objection. But the target of this objection is quite different. Previously, the objection was that an agent engaging in an activity is insufficient to entail any normative conclusions. Here, the target is the very existence of constitutive commitments (of chess).

6.1. Enoch’s View: Chess Lacks Constitutive Commitments

How is this example supposed to show that chess lacks a constitutive commitment? The object of chess’s constitutive commitment, if it has one, is presumably winning the game (and perhaps achieving a draw if victory becomes unachievable). So, if chess has a constitutive commitment then people who lack it (or who didn’t ‘care’ about winning the game, in Enoch’s terms) can’t be playing chess. But this ‘seems like a huge stretch.’ (Enoch 2011, 214) We want to say they are playing chess. Therefore, chess must not have a constitutive commitment. I should also mention that Enoch is not alone in his conviction here. Olson similarly writes—‘Note that a soldier might not desire to comply with the general’s order, and he might have no ends that would be served by his compliance. The same goes for chess players and soccer players; they might not desire to play by the rules and they need not even desire to win.’ (Olson 2011, 65).
6.2. Suits’ View: Chess Has Constitutive Commitments

Not everyone agrees with Enoch and Olson. The orthodox view of what it is to play a game like chess, if there is one, is probably that given by Suits in *The Grasshopper* (2014). On Suits’ view, *contra* Enoch, one of the necessary conditions of ‘playing chess’ is the adoption of ‘a prelusory goal—any checkmate arrangement of the pieces’ (Suits 2006, 7). In other words, Suits’ account specifies that a chess player *must care about winning* to count as playing a game of chess (indeed, to be a chess player).\(^{17}\) I’ll shortly provide some of the reasons I find Suits’ account plausible, in the face of objections. But for now, I should acknowledge the existence of a third position.

6.3. Ridge’s Alternative View: Chess Has *Promise-Like* Constitutive Commitments

Ridge agrees with Suits (and with me) that a necessary, that is, constitutive, condition on someone playing a game such as chess is that they ‘undertake commitments to abide by the relevant rules and pursue its lusory aim’ (Ridge 2019, 78). Hence, Ridge also rejects Enoch’s claim that a chess player is, in the early game, playing shmess.\(^{18}\)

This all might sound favourable to my account. However, Ridge interprets *undertaking a commitment* to be an essentially social phenomenon. Specifically, he claims

\[^{17}\] Other accounts of what it is to play a game go even further. Kreider (2011), for instance, claims that taking on such a commitment is *sufficient* to count as playing a game. Admittedly, Kreider’s sufficiency claim seems implausible. As Ridge (2019, 77) argues, surely even the most caring-about-checkmate person in the world needs to at least ‘show up’ to count as having played chess. Suits’ necessity claim is far more plausible.

\[^{18}\] Ridge confirmed this to me in correspondence.
that these commitments are ‘communicative acts of a certain sort’—effectively promises’ that players must make ‘to one another’ (76). So, while undertaking these commitments is constitutive of earning the status of playing the game, like promises, they needn’t be sincere. You can, in the merely social sense, be committed to your opponent to attempt to win by checkmate but simultaneously lack any desire- or intention-like mental state we might call ‘having a commitment’ toward that end or goal, just as you can promise to wash the dishes while lacking any desire or intention to do so. It’s just a false promise or a false commitment. As an illustration, Ridge discusses what is involved in chairing a meeting. He asks, ‘What if I show up, follow the constitutive rules, but do not try at all to pursue the committee’s constitutive goal?’ (Ridge 2019, 82) (That goal may be ‘to set a new policy on plagiarism’.) ‘Plausibly in that case’ claims Ridge, ‘I have chaired the committee, albeit poorly. […] Here we find the same pattern as I posit for playing a game.’ (82)

This is significant, and troublesome for me, because my Humean-friendly account of constitutive commitments assumes that undertaking a constitutive commitment means something like caring about the constitutive goal of the practice (in the sense of having a desire-like mental state toward that goal). But Ridge’s sense of constitutive commitment requires no such mind-to-world mental state. I’ll refer to Ridge’s sense of commitment as social commitments and the Humean-friendly version, following Calhoun’s terminology, attitudinal commitments (Calhoun 2018, 93).

Before I give my argument for the view that constitutive commitments sometimes involve undertaking commitments in the attitudinal sense, I first want to note that even Ridge’s account solves the shmagency objection (insofar as agency is analogous with chess-playing). That’s because his account also posits the existence of constitutive commitments to the goal of the practice. And being committed to that goal, even in the social, communicative sense Ridge is interested in, plausibly entails that you have certain normative pro tanto
reasons, in just the same way that having promised to do X gives you some reason to do X, to take the necessary means to X-ing, etc. Indeed, we can only understand what it means to be committed to something as being ‘on the hook’ of normativity. We should not be at all surprised if normative conclusions follow. So, this is another way for the shmagency objection to be answered—though it is not the one I am most interested in.

6.4. Who Is Right?

Which of these three views is most plausible? Does chess lack constitutive commitments, as Enoch claims? Do they have them, as Ridge and Suits claim? And, if they do have them, are they communicative acts which generate commitments in the social sense, as Ridge claims, or are they attitudinal commitments: mental states which constitute commitments to some goal, as Suits claims? I suspect Suits has things right.

The disagreement between Suits and me, on the one hand, and Ridge on the other, is, in a sense, narrow. We all agree, contra Enoch, that playing chess necessarily involves taking on constitutive commitments to aim to achieve the goal of chess—checkmate. We disagree on whether a certain character is possible, a character Ridge labels the ‘trifler’. This is ‘someone who is meant to be playing a game, but who does not try to win.’ (Ridge 2019, 86) On this issue, Ridge agrees with Enoch that such a character is possible, in the case of chess, while Suits and I disagree. We don’t think the trifler is really a chess player.

Let’s begin with Enoch’s case for his intuition that chess lacks any constitutive commitment—the case in which he is teaching his daughter chess with no intention of winning. It seems clear to Enoch that he is still playing chess. (Enoch 2011, 213-4).

Perhaps it’s a symptom of reading too much constitutivist literature, but I lack this intuition that a person who isn’t trying to win is still playing chess—that a trifler, in Ridge’s
terms, is still a chess player.\textsuperscript{19} Admittedly, in ordinary conversation, we might call what Enoch is doing with his daughter ‘playing chess’. I think that calling this practice ‘playing chess’ will often be fitting. But whether we would or wouldn’t assent to such a description depends on the context of the conversation.

Imagine an observer criticising Enoch’s seemingly poor chess moves. The observer points out that Enoch had the opportunity and reason to take his daughter’s queen, that he must either be ignorant of the potential move or irrational given he’s playing chess—a competitive game! At that point, Enoch may feel inclined to have a private word with the observer and re-describe his activity as ‘teaching my daughter how to play chess’ rather than ‘playing chess’. Under this new description, the observer would certainly find it easier to make sense of Enoch’s behaviour as rational—the intentional action of an agent who is responsive to reasons. And I think the commitment constitutivist about chess has a plausible account of what makes that difference (perhaps among other things). Enoch the chess teacher does not have the same aims, the same practical commitments, as Enoch the chess player. Most notably, Enoch the teacher lacks the practical commitment to winning. I take this to be some evidence of the existence of constitutive commitments, at least in the case of competitive games such as chess.

Other examples lend further plausibility to the existence of constitutive commitments. Indeed, when the difference between the ‘standard’ version of an activity and the ‘training’ version is more important, we are careful to use different terms to refer to the activities—boxing and sparring, for example. Sparring’s constitutive aim, if it has one, is presumably something like to improve the skill of the sparrers. This looks like a constitutive aim of the

\textsuperscript{19}I am not unique in this. Others who intuit that triflers are not genuine players include (Ferrero 2009, 313) and (Katsafanas 2013, ch.2).
kind possessed by medicine; it does not entail a participant adopting the constitutive aim as one of their own by taking on a practical commitment. Boxing’s constitutive aim, on the other hand, is winning the fight. I think boxing, like chess, has a constitutive aim in the sense that requires participants to take on certain specific practical commitments. We interpret those who are boxing to be aiming, in some sense, to win the fight. We don’t think it’s required for those who are merely sparring to have taken on the same aim. Indeed, if someone does try to ‘win’ at sparring, we think they are confused about the nature of the activity in which they’re engaged. Enoch says that regarding his ‘sort-of’ playing chess, ‘the claim that I don’t even do anything similar to playing chess is too much to swallow.’ (Enoch 2011, 215) But it seems to me that chess and ‘sort-of’ chess are as different from one another as boxing is to sparring. These activities are superficially similar and yet really rather different. You wouldn’t want the person you intended to spar with to be confused about the difference! And, in both cases, part of the explanation of the distinction between the activities (chess/teaching chess and boxing/sparring) is that one activity constitutively involves taking on certain practical commitments while the other doesn’t.

Of course, our ordinary language use is not a sharp metaphysical scalpel ‘carving at the metaphysical joints’. It might be objected that we ought to refrain from using our ordinary linguistic practices, such as the distinction between boxing and sparring, as an evidence base when it comes to thinking about whether there are constitutive commitments. So, perhaps we ought to ignore this evidence. But Enoch's objection to constitutive commitments is based on the same evidence base, namely the claim that we still call chess without the constitutive commitment ‘chess’. So either Enoch’s objection stands on the linguistic evidence, in which case my reply also stands, or Enoch’s objection can be dismissed. Either way, the objection to constitutive commitments is not compelling.
A brief recap. My thesis is that some activities, chess being my example, plausibly have constitutive commitments. Enoch argues that our everyday practices for differentiating activities suggest chess lacks a constitutive commitment. But it’s plausible that there are cases in which a participant in an activity, to count as a participant in the activity, must take on a practical commitment to the activity’s aim. In other words, it’s plausible that some activities have constitutive commitments. So the main thrust of the implausibility challenge is unconvincing.

Let’s now turn to Ridge’s argument against chess having a constitutive commitment in the attitudinal sense. Ridge begins with the same intuition as Enoch, that someone who doesn’t care about winning, a trifler, can still be a chess player. It seems possible, thinks Ridge, for a parent who wants to ‘encourage’ their child or ‘even just pass the time’ to be ‘actively trying to lose [a game such as chess…] Yet even in this case, it is plausible to say that the parent is playing the game.’ In other words, Ridge shares Enoch’s intuition that chess lacks a constitutive commitment in the attitudinal sense that chess players must care about winning.

But, as the Walt Whitman line (almost) goes, ‘Parents are large, they contain multitudes’. A parent might simultaneously undertake losing as a goal and winning as another goal. It may be practically incoherent, perhaps even practically irrational, to want (or intend) to achieve both goals simultaneously, but it seems psychologically possible. So, the fact that a parent might be aiming to lose is no counterexample to Suits’ view that the parent must also, in some sense, be aiming to win insofar as the parent is rightfully considered to be playing chess. So Ridge’s example is just a recapitulation of Enoch’s intuition. I have provided an explanation of why I don’t share it.

Now, Ridge provides some additional argumentation which might get us beyond this intuitive stalemate. Ridge points out that the trifling parent might accidentally win, and the
child point out the checkmate. But ‘you cannot win a game you never played.’ (Ridge 2019, 87) Triflers can win and lose games, and, therefore, it must be possible for them to play. And if they can play, then chess lacks constitutive commitments (in the attitudinal sense). The inference is good. If you win or lose a game, you must have been playing it. So the argument really rests on the intuition that the trifler can win or lose a game. But, as the saying goes, one man’s *modus ponens* is another’s *modus tollens*. Those such as Suits who share my intuition that the trifler is not playing chess will infer that he can’t have *won* a game of chess. He has performed physical movements that would, in a game of chess, constitute checkmate and hence winning. But without the trifler aiming to win by checkmate, he may as well have performed those physical movements to stretch his fingers. On the other hand, those such as Enoch and Ridge will infer from the intuition that triflers can win or lose games. But we may still have a simple disagreement at the level of intuition.

Can we break the tie? Ridge continues intuition-pumping with the following cases:

Suppose I enter and turn up at a chess tournament, sit down at the board and make a series of legal moves but make *no* attempt to win. When I then find myself in checkmate, I complain to the arbiter that this should not count as a proper loss, but as a default, on the grounds that if I did not try to win then no game was played, and you cannot lose unless you played. [...] Such a complaint would quite rightly be heaped with scorn.

Moreover, if the trifler accidentally checkmated her opponent, and the arbiter spotted the mate, then the trifler would have won the game. If the opponent complained that the trifler was not trying to win, but just ‘got lucky,’ then the arbiter would correctly reply ‘Tough luck for you; she still won.’ Again, you cannot win a game you never played. We had therefore better not follow Suits in making trying to win into a necessary condition for playing the game. (Ridge 2019, 87)

I take it that Ridge is seeking to reinforce the intuition that triflers are game players by calling on some intuitive allies. In particular, we would seem justified to have certain moral
(or at least interpersonal in the Strawsonian sense) reactions and these reactions only make sense if the trifler is playing. But I find this less than fully compelling.

The trifler’s complaint in the first instance and the opponent’s in the second seem like the rightful objects of scorn. But I’m not convinced that the explanation for the aptness of scorn straightforwardly entails the metaphysical truth of the trifler having actually played chess. In the position of Ridge’s arbiter, the trifler’s complaint (and that of his opponent in the second variation) would come across, to me, as insincere. The issue is that the arbiter should not believe that the trifler is a genuine trifler, that he showed up to the competitive contest, played legal chess, and in no sense took on any aim of checkmate as one of his own. It’s an unbelievable story, the kind of story that a sore loser with a degree in philosophy might come up with. The arbiter’s reasonable belief that the trifler’s claim is an insincere attempt at saving face for the loss seems to do all the work of licensing the interpersonal reaction of scorn. We don’t seem entitled to draw any metaphysical conclusions about the nature of chess. The fittingness of our reactive attitudes does not imply that the trifler was both playing chess and failing to care about checkmate.

Let’s remove the distracting interpretation that the person making the trifling-therefore-not-playing appeal is insincere in his claims that he had no desire to win and is simply attempting to save face for having lost. Consider this variation of the case. The trifler won, by accident, and that checkmate was noticed by the arbiter. But then the trifler insists he wasn’t really playing, that he was moving pieces at random, and that he does not deserve to have the match counted toward his record as a win. We might still react with scorn, and yet we couldn’t accuse the trifler of lying to save face. Even in this scenario, I still think it is reasonable to believe that the trifler is lying, perhaps to humiliate his opponent (who, according to the trifler’s claims, lost a game of chess to randomness rather than a skilled opponent). We need to do better than this for our reactive attitudes to be truly informative.
Let’s think about how the arbiter should respond if she does, somehow, justifiably believe the trifler’s claim is sincere. She might, like Enoch and Ridge, think the trifler was playing chess and is confused about the metaphysics of game-playing. Or she might, like Suits and myself, think the trifler is an even more confused individual, perhaps even that he is doing something like crying out for help, but not that he is a chess player. A more apt response than scorn to the sincere trifler might be pity, concern, or compassion—providing the phone number of a therapist seems more appropriate than scorn. It is, after all, very strange to go to professional chess tournaments and simulate/play the competitive game of chess (in a competitive context, no less) while sincerely lacking any intention or desire to checkmate one’s opponent. So the reactive attitude of scorn seems appropriate only if we think the trifler is a liar (about trifling). If we think the trifler is sincere, that he really wasn’t trying, then scorn seems inappropriate. And reactions of pity and compassion do not entail that the trifler was playing chess.

Intuitions on these matters seem controversial and hard to isolate from distracting features. Of course, neither Enoch nor I really care if chess has a constitutive commitment; we’re interested in whether agency does. Chess is merely meant to be proof of the principle that some activities or practices can have constitutive commitments (in the attitudinal sense). Interestingly, while Ridge denies that chess (and most multiplayer games, more generally) has a constitutive commitment in this exact sense, he does accept the existence of constitutive commitments in the Humean sense for single-player games such as solitaire:

Interestingly, though, we sometimes use ‘undertake a commitment’ in a sense which need not involve communicative acts. Here I have in mind the sense invoked when we say someone has undertaken a commitment simply by setting themselves a goal. Perhaps we should say that when one plays a single-player game that this consists, in part, in undertaking a commitment in this sense.
A commitment to do what, though? Plausibly, a commitment to achieve some goal which is intuitively the point of the game, but using only means allowed by the rules of the game. (Ridge 2019, 91)

So it might be that Ridge’s account of games is right, and I’ve simply opted for the wrong analogy as my proof of principle that there are such things as attitudinal constitutive commitments. Perhaps agency is not analogous with chess, but with solitaire.

More broadly, what Enoch and constitutivists really care about is whether agency, not chess or solitaire, has constitutive commitments. Perhaps the implausibility objection should be interpreted as claiming that the existence of constitutive commitments is implausible specifically when it comes specifically to agency. For instance, Wiland writes:

Constitutivism is ambitious. It attempts to extract an account of reasons for action from reflection on the bare ideas of action and agency. So we shouldn’t be surprised if doubts remain. Those who claim to extract reasons out of agency might remind us of those who claim to pull rabbits out of hats. You suspect that there must be a trick (Wiland 2012, 141).

The difficult ‘trick’ constitutivists must pull off is to provide a sufficiently minimal account of agency such that it remains plausible whilst leaving enough substance in agency that we can ‘extract’ substantive normative practical reasons. We might worry that the idea that the possession of certain practical commitment is constitutive of agency puts too much substance into the notion of agency for it to remain plausible (Enoch 2011, 213-4).

I acknowledge the challenge, but as an objection to commitment constitutivism, this specific form of the implausibility challenge is inconclusive. The possibility of an objection is

---

20 As Ridge has pointed out to me in correspondence, the solitaire-agency analogy raises its own questions about the nature of commitments with no social component, i.e. whether having such a commitment requires recognition of others and challenges in the line of private-language style arguments. As I don’t pursue this option, I won’t try to settle those issues here.
not an objection. Having more substance to our notion of agency will presumably make deriving normative conclusions easier. Whether it also creates an implausibly overinflated notion of agency can only be determined once we have a specific account to examine. I’ll demonstrate the possibility of such an account by providing one Chapter Five.

7. Irrelevance

The last of Enoch’s objections, and the most complex, is that it is normatively irrelevant if agency has constitutive aims. I argue that constitutivism can avoid the irrelevance objection’s bite if we understand constitutivism modestly, combining it with a Humean theory of normative reasons.

7.1. The Objection

The Humean commitment-based constitutivist identifies practical commitments that agents necessarily have—as a constitutive matter—and then specifies the reasons agents have given that they care about X, to use Enoch’s terminology. For example, given that you care about animal welfare you presumably have some kind of reason, perhaps not conclusive, to refrain from purchasing many animal products. But is an agent having a constitutive commitment relevant at all to what he ought to do?

In the case of chess, Enoch claims that for chess-related reasons to be normative, ‘it is not sufficient that you do in fact play chess. Rather, it is also necessary that you have a reason to play chess’ (2006, 185). I’ve called the generalised version of this claim ‘trickle-down normativity’. In ‘Shmagency Revisited’, Enoch concedes that the issue may not be as clear-cut assuming something along the lines of what I’ve called commitment constitutivism. If you count as playing chess only if you ‘care about checkmating your opponent’ then it is ‘much more plausible that merely playing chess […] suffices for your having a reason to
checkmate your opponent (regardless of whether or not you have a reason to play chess).’ (Enoch 2011, 212) Nevertheless, he maintains that, contrary to the initial plausibility of the commitment constitutivist solution, it is still normatively irrelevant whether or not agents care, or take on practical commitments. Here is the objection:

The thought here is very simple: noting that I do X is never a good answer to the question of whether I should X. This is true for actions, and it is just as true for carings. Perhaps I do care about something; but how does noticing this fact count as an answer to the normative question of whether I should care about it, or indeed as a reason for caring about it? (Enoch 2011, 215)

Admittedly, the normativity of our attitudes and of our reasons-given-our-attitudes are deeply linked. Someone might come up and suggest that you shouldn’t care about animal welfare, and, therefore, that the ‘reasons’ to refrain from purchasing animal products that you have given that you care are not normative. Perhaps the normativity of the reasons for action you have given that you care about X hinges on the normativity of your caring about X. To vindicate the normativity of your reasons, it’s no good to simply state, ‘I care about animal welfare’. You also need your caring to be normative. Applied to commitment constitutivism, the worry generalises. The commitment constitutivist can never succeed in vindicating the normativity of whatever reasons agents have given their practical commitments without first vindicating the normativity of those practical commitments. The fact that some practical commitments are partly constitutive of agency is irrelevant. That only shows agents have these commitments, not that they should.

Notice that Enoch’s irrelevance objection targets the normativity of the constitutivist’s reasons for action indirectly. Their normativity is in doubt, he thinks, because we need reason to care about something for our caring to be normatively significant. He is assuming trickle-down normativity and disputing not that we do care but that we have reason to care. Why can’t the constitutivist provide any such reason (to care)? Because doing so would be an instance of the naturalistic fallacy:
The point is not merely an ‘is-ought gap’ kind of point. [...] What we are up against here is an especially problematic instance of such a move [from is to ought] – it is the move immediately from someone caring about something to it being the case that she should care about it, or at least that she has reason to care. I take it that even those of us with the strongest stomach for naturalistic fallacies should not be happy with such a move. (Enoch 2011, 215)

Hence, the mere fact that we do care is ‘utterly irrelevant’ (215, my emphasis). Of course, the constitutivist is inclined to dig in—of course, the fact that you are committed to something has at least some normative significance in determining what you ought to do!

7.2. Against an Internalist Solution

How can the constitutivist respond? A tempting (but ultimately unpromising) move would be to appeal to internalism.

Constitutivists are typically internalists and an important motivation for constitutivism is existence internalism about reasons. Existence internalism claims that a necessary condition on R being a reason for an agent A to phi is that A has some pro-attitude that has the right connection to phi-ing. So perhaps this strong connection between our practical commitments and our reasons can answer the irrelevance objection. Perhaps our actual commitments are relevant to our reasons in the way existence internalists describe.

But an appeal to internalism won’t do. Internalism merely places a necessary condition on something. It does not seek to provide a theory of normative reasons. So it is

21 Enoch also rejects the appeal to internalism but on slightly different grounds to me (Enoch 2011, 217). He observes that internalists don’t claim that the possession of a pro-attitude is sufficient to provide reasons, and that, therefore, they can’t answer the objection. He is right that internalism merely makes a necessity claim about the relationship between agents’
incapable of providing the kind of down-to-the-roots vindication of the reasons we have given our constitutive commitments that the constitutivist needs to answer the irrelevance objection. Of course, dialectically, appealing to a full metaethical theory isn’t an option for an ambitious constitutivist; he can’t appeal to a rival metaethical theory to explain the normativity of an agent’s constitutive reasons. But a modest constitutivist, as I’ve already mentioned, can.

7.3. A Humean Solution

Take Mark Schroeder’s reductive Humean theory of normative reasons:

**Reason.** For R to be a reason for X to do A is for there to be some p such that X has a desire whose object is p, and the truth of R is part of what explains why X’s doing A promotes p.

(Schroeder 2007, 59)

On Schroeder’s Humean account, the agent having a relevant attitude (desire) for p is part of her having a reason. For example, Ronnie likes dancing. He has a desire to dance. There’s going to be dancing at the party. And, since going to the party will promote the object of Ronnie’s desire—dancing—that means that the fact that there will be dancing at the party is a reason for Ronnie to go to the party. Assuming this analysis of reason is correct, the claim that agency has constitutive desire-like attitudes such as commitments is clearly not ‘irrelevant’ to the question of what normative reasons an agent has. An agent having a pro-attitude, such as being committed to something, or caring about something, partly constitutes her having a reason for action. Practical commitments fall under the umbrella of ‘desire-like attitudes and their reasons, and not a sufficiency claim. But, as I explain later in this chapter, constitutivists don’t need the sufficiency claim, exactly. They only need constitutive commitments to be reason-providing—originative of reasons (perhaps in combination with other facts)—not sufficient to provide reasons all by themselves, in isolation of other facts.
states’ that the Humean thinks are reason-providing. So constitutive commitments, if they exist, would clearly be originative of normative reasons, on a Humean account.

Humeanism, then, denies the implicit assumption behind the irrelevance objection: trickle-down normativity. Humeanism grounds normative truths in facts about what agents are, or would ideally be, practically committed to (or desire more generally). The highest court of appeal, to the Humean, is the agent’s (ideal) desire-like states. Once we know what an agent (or his ideal counterfactual self or advisor) wants (or would want under ideal circumstances), then there is no deeper justification of what he ought to do. And we don’t need a further reason for our desires for them to be reason-providing.

So, Humeanism explains why our commitments and other desire-like states are reason-providing. Of course, the Humean solution to the shmagency objection requires that we understand constitutivism modestly, in the way I’ve outlined in the previous section—as a plug-in option for full-blown metaethical views to use, a method of solving certain metaethical challenges.

22 The Humean account is, of course, not as simple as this brief characterisation suggests. For instance, on most Humean accounts, not old desire-like state is necessarily reason providing. But, in later chapters, I shall argue that constitutive commitments are plausibly reason-providing on a Humean account, and furthermore, that practical commitments being constitutive of agency would make their reasons particularly weighty.

23 With the few exceptions I discussed earlier, why is this option generally overlooked in the shmagency literature? I suspect it is because most of the well-known defences of constitutivism adopt the more metaethically ambitious interpretation of constitutivism. At least, the ambitious interpretation was the only one knocking about when Enoch presented his objection. So, in a sense, my favoured solution could be considered sidestepping rather than refuting the
8. Irrelevance and Insufficiency Objections to Humeanism

Here is a final bundle of worries worth considering. I have claimed the constitutivist can stand on the shoulders of Humean giants to escape the irrelevance objection. But maybe Humeanism itself faces similar objections.

How could we resolve an apparent deadlock between the Humean picture of how normativity is explained and the sceptic about the reason-providingness of our desire-like mental states? I don’t pretend to solve that question satisfactorily here, but we can at least get clearer on the nature of the disagreement and how it follows from deeper (perhaps implicit) metaethical assumptions.

8.1. Is The Humean Analysis Plausible?

Humeanism explains why an agent’s being committed to something can provide him with normative reasons. But this raises a new worry. Perhaps the ease by which the Humean moves from commitments to normative conclusions should give us pause. Enoch sees such a move, at least in the hands of a constitutivist, as a ‘move immediately from someone caring about something to it being the case that she should care about it [...] an especially problematic instance’ of the naturalistic fallacy (Enoch 2011, 215). His own position is that being committed to some end—without some independent reason to be committed—is ‘utterly irrelevant.’ (2011, 215)

---

shmagency objection; I am not confident that a metaethically ambitious constitutivist, Enoch’s original target, could adequately answer the shmagency objection.

24 Enoch’s irrelevance objection is built upon the principle that participation in an activity (or a practical commitment to some end) is only reason-providing if the agent already has a reason to
Being precise, the Humean does not make this inferential move from what one does care about to what one should care about. Rather, he moves from what one does care about to what one ought to do. Nevertheless, perhaps the fact an agent is committed to something is, insufficient\textsuperscript{25} to determine or, contra Humeanism, ‘utterly irrelevant’ to what reasons he has.

8.2. Blunting the Objection

A preliminary point is that the Humean slide from the ‘is’ of our pro-attitudes and practical commitments to the ‘ought’ of our normative reasons for action is not quite as

\textsuperscript{25}Why does Enoch assume the constitutivist must endorse such a strong claim as sufficiency? I think it follows from his commitment trickle-down normativity. If trickle-down normativity is true, then the constitutivist must claim that you have a reason to care about whatever aim is constitutive of agency. And all the constitutivist actually does is claim is that you do care. So to fill that gap, Enoch thinks, the constitutivist must think he can move immediately from the agent caring to the agent having a reason to care. But the Humean denies trickle-down normativity, so he isn’t forced to tell this story.
slippery and implausible as might be feared. I’ll briefly mention a few ways in which Humeanism can pump the breaks.

The general point is that there is room for Humean to occupy a middle ground here—something between the extreme views that (1) having any old commitment is sufficient to provide specific weighty normative reasons and (2) that our actual commitments are utterly irrelevant to our reasons. Recall that the constitutivist only needs agents’ commitments to be originative of reasons, i.e. for agents’ commitments to be reason-providing. But they might be reason-providing in combination with many other facts. The existence of our commitments alone might not be doing all the work.

8.2.1. Desires and Commitments (Alone) Are Insufficient To Provide Reasons

How strong a connection between one’s reasons and one’s commitments does the Humean posit? The claim that mere possession of a commitment is sufficient to derive normative conclusions is clearly too strong, even for the most ardent Humean. The fact I am committed to becoming twenty metres tall, or being born a year before my actual birthdate, or even intergalactic justice, is not reason-providing in part because there’s nothing to be done

26 A modest constitutivist wants is to vindicate the normativity of something that’s proved hard to vindicate by other philosophical methods—the normativity of morality, for example. Suppose the constitutivist can show that some important piece of the puzzle of normative vindication, though not the whole thing, is constitutive of agency. If something constitutive of agency is the missing piece of an otherwise complete vindication of the normativity of the domain, then that piece is going to be relevant enough to our normative reasons for the constitutivist to score some valuable points in the vindication game. The existence of a commitment needn’t be the entire story.
to promote the end to which I am committed. At the very least, ought implies can; for me to have normative reasons to pursue my ends, it must be the case that the world is cooperative enough such that my desires have the possibility of being satisfied.27 If the constitutivist did require the possession of commitments to be sufficient for deriving substantive normative conclusions, it would indeed be an implausible position.

8.2.2. Some Desires and Commitments Are Not Reason-Providing

Not only does the world have to cooperate for there to be normative reasons, but many versions of Humeanism have the resources to say some practical commitments are not reason-providing. It may be that you would abandon some of your commitments if only you reasoned or imagined sufficiently. (Or, you may have a desire only due to ignorance of relevant facts, like Bernard Williams’s accidental petrol drinker who believes he is eyeing up an enjoyable glass of gin and tonic (Williams 1981, 45)). A Humean needn’t accept any old desire or commitment like these as reason-providing.

8.2.3. Only Intrinsic Desires Are Originative of Reasons, Not Instrumental Desires

The Humean also does not (standardly) claim that all our pro-attitudes and commitments are originative of reasons. Specifically, it is standard to think that intrinsic desires, desires ‘for states of affairs that are wanted for themselves’ (Schroeder 2020, 2.2) are the grounds of our normative reasons. This is in contrast with instrumental desires—desires for those things wanted merely as a means to an intrinsically desired end. These instrumental

27 A voice of dissent here is Brown (1977, 218) who claims that an ought-judgment is ‘pointless’ rather than ‘meaningless’ when the agent cannot perform the required action.
desires do not originate normative reasons but they do transfer the normativity of our intrinsic-desire-based reasons to our instrumental ends and desires. Let’s take an example.

Say you intrinsically desire pleasure and you work out the necessary means to obtaining pleasure is, in this case, eating a slice of cake. You may now have an instrumental desire to eat the cake, as a means of realising your intrinsic desire for pleasure. But your instrumental desire to eat cake isn’t what grounds the normativity of your reason to eat the cake—your intrinsic desire for pleasure does that. This is another way in which it is false, on the Humean picture, that any commitment (or other pro-attitude) is sufficient to originate normative reasons.

8.2.4. Weak Reasons and Unspoken Reasons

If, as the Humean thinks, a chess player’s being committed to winning is reason-providing, why can we still doubt the normativity of a chess player’s chess-related reasons? Imagine, for example, you have been blackmailed to go lose a chess match. If you ever want to see little Timmy again, you presumably have conclusive reason to lose. It might seem that you now have no reason to win the game, nor to realise the instrumental necessary means to winning the game. So the fact you are committed to the constitutive aim of chess seems at least insufficient to provide normative reasons, and perhaps totally irrelevant.

I think our intuitions about what reasons we have in the blackmailed chess game case may be distorted by certain implicit assumptions. When we talk about someone having a reason to *phi*, or having reason to *phi*, we generally mean they have good reason to *phi* or even conclusive reason to *phi*. This tendency has, as Schroeder has argued, more to do with the pragmatics of language (and perhaps conscious thought) than the ontology of reasons (Schroeder 2007, ch.5). Ordinarily, there’s no need to mention or even think about weak
normative reasons. They don’t tend to have much influence on what we ought to do, all things considered. Given these dynamics, we should give less credence to the intuition we have no reason to play chess well in the blackmail scenario and more credence to the proposition that we have weak, pro tanto reasons to play chess well, given our (necessary) commitment to win the game. These weak reasons are easily outweighed by our blackmail-related reasons to lose the game of chess so easily that we typically ignore their existence.

Of course, there is also the distinct possibility that upon being blackmailed you completely abandon your commitment to win the chess game. Our conative economy often streamlines itself into more coherent patterns in this kind of way, and perhaps it is even rational to do so. If you really did abandon your commitment to winning upon being blackmailed, then I don’t think you have any reason to win, not even a pro tanto reason. After all, you’re not committed to winning the game.

In summary, there are two psychological possibilities in the blackmail case. Either you retain your chess-related constitutive commitment, in which case you have a weak pro

28 One complication is that we might think that only ‘all-things-considered’ or ‘conclusive’ reasons are normative reasons for action (and not pro tanto reasons). The fact that I have some reason to rob a bank (to get more money) does not entail I have an all-things-considered reason to rob a bank. I presumably have reasons stemming from my interests in avoiding immorality and jail to refrain from robbing the bank. And these might outweigh my reasons for. There are at least two ways to think about the normativity of my reasons here. The first is to think of all my pro tanto reasons as genuinely normative. The normativity of the all-things-considered reason is presumably inherited from the balance of pro tanto reasons. A second way to think about it is that none of my pro tanto reasons are really normative. That something provides me some reason to do something says nothing about whether I really should do that thing—only all-things-considered reasons can have this normative property. I assume the former account.
normative reason to e.g. develop your pieces that’s outweighed by your conclusive normative reasons to lose the game. Or, you abandon your chess-related commitment in which case you indeed have no reason whatsoever to play chess or to play chess well (after all, you’re not playing chess). Neither case threatens the Humean constitutivist account. The Humean does not claim that any old desire or commitment will provide conclusive reasons to act.

8.3. Irrelevance 2.0

Humeanism clearly has some resources to control the slide from is to ought, from our commitments to our reasons. But such allowances (for the possibility of commitments that are not reason-providing) lay the ground for an updated version of the irrelevance objection, this time aimed at Humeanism rather than constitutivism. Without a reason to think my commitment is a reason-providing one, I don’t have any reason to think my commitment-given reasons are normative. So the fact that I am committed is normatively irrelevant. The question is still whether I should be committed.

My first point: This objection inappropriately shifts the burden of proof. Humeanism claims that an agent’s reasons are more-or-less a (perhaps complicated) function of her attitudes and the state of the world. In this picture, our commitments are innocent until proven guilty. That is, we ought to think our commitments are reason-providing until we have a specific reason to think they’re not.

Admittedly, burden-shifting is not the most convincing argumentative move in the face of sceptical objections. My second, and more significant, point: the updated irrelevance objection is question-begging against Humeanism, and possibly against metaethical naturalism more generally. The implicit assumption of the irrelevance objection is that the only vindication of normative facts (i.e. what reasons we have) that there could be would
involve reference to further normative facts (i.e. why we should care/have that commitment). So the reductive naturalistic vindication of our reasons’ normativity, of the sort the Humean provides, will never satisfy the critic. He can always ask ‘But why care?’ and demand an answer with normative content. Of course, this is a respectable challenge, one that non-naturalists who are convinced of the sui generis and autonomous nature of the normative are likely to want to make. Just as Moore can always ask, of any naturalist analysis of goodness, ‘But is that good?’ Similarly, for any kind of naturalistic fact about agents’ commitments that is meant to originate, provide, or partly constitute reasons, the irrelevance objection can ask, ‘But do we have a reason for that commitment?’

What does this exchange reveal? Merely that the Humean theory is reductive and naturalistic and therefore faces standard open-question argument objections. But there is nothing particularly problematic about the Humean constitutivist strategy that makes it any more vulnerable than other naturalistic theories of normativity. So the irrelevance objection to Humeanism is merely an ‘is-ought gap’ kind of point, and it has the limited force of all such objections.

29 In other words, the irrelevance objection to Humeanism is just a new dress on the traditional non-naturalist criticism of naturalist metaethical theories, such as Moore’s open question argument. The charge is that naturalist theories are committing the naturalistic fallacy in attempting to move from an ‘is’ to an ‘ought’. Many philosophers find the open question argument a convincing objection to naturalist metaethical theories. But the argument seems to be taken less seriously in naturalist circles. I don’t seek to adjudicate that debate here.
9. Summary and Conclusion

A brief summary of how commitment Humean constitutivism deals with Enoch’s shmagency objections. Enoch’s argument is that agency is like chess. Just as the constitutive aims of chess are normatively impotent without some reason to play chess in the first place, Enoch argues, so too are the constitutive aims of agency normatively impotent without some reason to be an agent in the first place.

I first argue that interpreting constitutivism modestly provides an easy answer to the basic form of the shmagency objection. But this discussion still leaves the reason-providingness of constitutive features underspecified.

I then move on to discuss Enoch’s development of his shmagency arguments in ‘Shmagency Revisited’. As I explain in section 4, instead of rejecting the analogy between agency and chess, I’ve argued that constitutivism (at least, modest Humean commitment-based constitutivism) can accept the analogy and still answer the shmagency objection.

Enoch’s first objection to aim constitutivism (section 5) is that participation in an activity with a constitutive aim is insufficient to entail anything normative, nor that you should ‘set your sights on that aim’ (Enoch 2011, 210-1). But given commitment constitutivism—the view that participation in agency requires agents to take on certain practical commitments—it seems like participating in an activity can have normative implications in the same way that having practical commitments can have normative implications.

The next objection (section 6) is that it is implausible that there are constitutive commitments. This is unconvincing both as a general claim and a specific one about agency. I rejected the general claim because the distinction between playing chess and teaching chess (and boxing and sparring) is plausibly explained by constitutive commitments. I rejected the specific claim because it remains to be established.
The last of Enoch’s objections I’ve discussed (section 7) is that, even if they exist, constitutive commitments are normatively irrelevant. To derive any normative conclusions it’s not enough that you care about something; you also need reasons to care. My response to this objection is appealing to the Humean theory of practical reasons, according to which one’s commitments are (generally) reason-providing.

Finally, (section 8) I considered whether similar objections to those Enoch raises against constitutivism could be raised against Humeanism. I argued that the Humean is not committed to an implausibly bold thesis about the connection between our reasons and our commitments (8.2) and that the charge that our commitments and pro-attitudes are irrelevant to our normative reasons is question-begging against Humeanism (8.3).

Where does this all leave us? The intended target of Enoch’s shmagency objection is ambitious forms of constitutivism and the arguments provided here do not give any indication that Enoch’s arguments are unsuccessful against such views. However, understanding constitutivism modestly permits the shmagency objection to be skirted. Combining constitutivism with Humeanism solves the problems raised by the shmagency objection. Humean constitutivism explains how constitutive features of agency (our practical commitments) can be genuinely reason-providing.30

In the next chapter, I turn things around and argue that the addition of constitutivism helps the Humean deal with well-known objections to the view. In particular, it permits the Humean to explain what it is often thought to be incapable of explaining: how there could be universal normative reasons to act morally that are nonetheless grounded in features of

30 Of course, I have not yet argued that we have these constitutive commitments, just that they would have normative implications if they existed. My argument that they exist is in Chapter Five.
agents’ psychologies. The proceeding chapter will explain how these reasons could, given constitutivism, be particularly *weighty*. 
Chapter Two: A Constitutivist Defence of Conservative Humeanism

“Would you tell me, please, which way I ought to go from here?”

“That depends a good deal on where you want to get to,” said the Cat.

“I don’t much care where –” said Alice.

“Then it doesn’t matter which way you go”, said the Cat.

– Lewis Carroll, Alice in Wonderland

1. Introduction

Constitutivists seek to derive a justification for the normativity of some reasons, norms, or claims from the constitutive features of rational agency. In the previous chapter, I argued that supplementing constitutivism with a Humean account of normative practical reasons provides the constitutivist with significant dialectical benefits, allowing it to circumvent the shmagency objection. In this chapter, I argue that a view like the commitment constitutivism outlined in the previous chapter offers equally significant benefit to the Humean. Specifically, commitment constitutivism helps the Humean answer the *amoralism objection*, the objection that Humeanism is incapable of handling our reasons to be moral. This is the most influential and difficult instance of the extensional problems that Humeanism is widely thought to face – that Humeanism generates too *many* reasons in some cases and too *few* in others. This chapter focuses on establishing that a constitutivist approach allows the Humean to accommodate the intuitive modal status of moral reasons.

This chapter explains Humeanism about normative practical reasons (section 2) and the amoralism objection (section 3), then offers a detailed argument that the objection can be adequately answered given certain constitutivist assumptions (section 4). I argue that if a
certain kind of constitutivism is true, then Humeanism can (a) get the extension of moral reason claims correct and (b) get the strong modal status of moral reasons correct. In the following chapter, I additionally argue that Humeans can also get the weight of moral reasons right and give an account of the nature of constitutive commitments. The conclusion of this chapter is entirely hypothetical: Humeanism can resolve various important problems if a constitutivist view such as commitment constitutivism is true.

2. Humeanism

To avoid potential confusion, here is a quick note on terminology. A failure of philosophical coordination means that the view I’m referring to simply as ‘Humeanism’ has an unpleasantly large number of synonyms. In the dominant, Hume-inspired category, synonyms include neo-Humeanism and the Humean Theory of Reasons (HTR) (Schroeder 2007). Other labels include subjectivism, or reasons subjectivism (Sobel 2016), the desire-based theory of reasons (Heathwood 2011), and existence reasons internalism (Darwall 1983). Humeanism is a thesis about normative reasons for action, such as the reason you have to drink a glass of water if you’re thirsty. It can be broadly characterised claiming that ‘all reasons must be explained by some psychological state of the agent for whom they are reasons, such as desire’ (Schroeder 2014, 60). In other words, Humeanism claims that the existence of an agent’s normative practical reasons, or something status as a reason, depends on, indeed is explained by, the existence and relevance of her desires (or other desire-like mental states). Tightened up to a definition, Humeanism is the following reductive analysis of reasons:
**Humeanism.** For R to be a reason for X to do A is for there to be some p such that X has a desire whose object is p, and the truth of R is part of what explains why X’s doing A promotes p. (Schroeder 2007, 59) 31

To motivate the coming defence of conservative Humeanism, I’ll briefly mention a few considerations which recommend the view.

One obvious attraction of Humeanism is that it captures ordinary intuitions about what normative reasons agents have across many cases (Schroeder 2007). The fact that Ronnie wants to dance explains his having a reason to go to the party. But Bradley, who hates dancing, doesn’t share Ronnie’s reason for going to the party. If our desires explain our reasons in such cases, it is reasonable to assume that they explain our reasons generally.

Internalist considerations about the close relation between reasons and motivation also favour Humeanism. The theory seems capable of explaining why an agent’s reasons seem to fit her in a special way (Railton 1986). What we want, wish for, and what our life projects are, all seem relevant to determining what we have reason to do and Humeanism can explain why this is the case. The differences in our reasons are explained by differences in our psychologies – differences in what we desire, what we care about, etc.

The traditional argument for Humeanism is premised on the combination of reasons internalism and the Humean theory of motivation, both of which have been seen as

_____________________

31 Earlier philosophical discussion often assumed that the agent’s desire to phi is her reason for phi-ing – that your desire to eat the cake is your reason for eating the cake. This tendency is partially explained by some philosophers insufficiently distinguishing motivating and normative reasons. See, for example, Davidson (1980). But this identity-based characterisation of the relation between desires and reasons is not essential to the view.
independently attractive positions. Definitions of reasons internalism tend to be controversial, but we can understand it broadly as the view that $R$ is a normative reason for $A$ to $\phi$ only if $R$ is capable of motivating $A$ to $\phi$ (in condition C). The second premise, the Humean theory of motivation, claims that desires are necessary (and beliefs insufficient) for motivation. This almost gets us to Humeanism; for $R$ to be a reason for $A$, $A$ needs to have some desire (in condition C), since a desire is required for motivation.

Humeanism also promises a reductive and naturalistic account of normative reasons. As Humeanism bases this explanation in naturalistic facts and properties, facts about agents’ psychologies, it seems well placed to deliver this naturalistic account. Additionally, combined with a ‘reasons-first’ account of practical normativity, according to which we can derive an account of all normative properties from an account of normative reasons, Humeanism would provide the basis of a naturalistic account of all practical normativity.

Humeanism clearly has much to recommend it. However, the theory is widely dismissed because of the amoralism objection. Let’s turn to that objection.

32 This line of argument originates in Bernard Williams’ defence of reasons internalism ‘Internal and External Reasons’ (1981). It is developed in (Smith 1994). For criticism and discussion of internalism and its connections with Humeanism, see also (Nagel 1970), (Darwall 1983), (Korsgaard 1986), and (Markovits 2014).

33 I say ‘almost’ because, strictly speaking, Humeans make a slightly stronger claim than this (Sobel 2016, 16-17). It’s not just that an agent’s reasons and his desires co-vary; Humeans claim that something being a reason is necessarily partly explained by the existence of a relevant desire.

34 This is discussed in (Darwall 1992). Some philosophers deny the attractiveness or even coherence of this naturalistic ambition, e.g. Jean Hampton (1998).
3. The Amoralism Objection

3.1. Introduction

Historically, the most influential objection to Humeanism is what Sobel labels the ‘amoralism objection’ (2016, 17). The amoralism objection comes in a variety of forms, but it claims, roughly, that Humeanism is incapable of providing a plausible account of moral reasons.

The Humean response to this objection is the source of a major dialectical division. *Revisionary* Humeans embrace that the view has revisionary implications for our moral practices, including revisionary implications for the moral reasons we ought to ascribe and our practices of blame. *Conservative* Humeans, on the other hand, argue that Humeanism, properly understood, has none of the revisionary implications that the amoralism objection alleges. It is this *conservative* Humean strategy that I argue constitutivism can help with.

35 Other terms include the Moral Criticism, and Warren Quinn’s objection based on the ‘shamelessness’ of the Humean conception of practical rationality (Quinn 1994, 216).

36 I take Bernard Williams, Philippa Foot in her early work, and Gilbert Harman to be revisionary Humeans. For more on this reading and its dialectical implications, see Christopher Cowie’s ‘Revisionist Responses to the Amoralism Objection: A Reply to Julia Markovits’ (2016).

37 I take Mark Schroeder, Julia Markovits, and Michael Smith to be representative conservative Humeans, although Markovits and Smith could also be classified as Kantian internalists. The boundaries between these views have been drawn in a number of ways but these subtleties aren’t relevant here.
This section proceeds as follows. In 3.2., I explain the amoralism objection. In 3.3., I define what I take to be the intuitive thesis about moral reasons conservative Humeans should seek to accommodate to answer the amoralism objection: moral rationalism.

3.2. The Amoralism Objection

Humeanism claims that for any reason R for an agent A to phi, R’s being a reason is explained by A’s desires (understood broadly as any pro-attitude or conative mental state, mental state with a mind-to-world direction of fit). As both proponents and critics of Humeanism (and the closely related reasons internalism) quickly noticed following Williams’ ‘Internal and External Reasons’ (1981), the view appears to have revisionary extensional implications.38

Since Humeanism claims our reasons are explained by our desires, if agents lack the relevant desires then Humeanism implies they also lack the relevant reasons. So, if an agent happened to lack any desire that would be served by acting morally, Humeanism implies she would lack any normative reason to do what morality demands. To take an example from Williams (1981), imagine an abusive husband has no motivational mental states that could lead him to want to treat his wife better. Humeanism seems to imply he has no reason to treat his wife better. Williams, as a revisionary Humean, embraces this conclusion.

38 These issues have received extensive discussion. For the best existing defence of Humeanism against these charges see Schroeder’s Slaves of the Passions (2007). For a response to Schroeder, see Eden Lin’s ‘Prudence, Morality and the Humean Theory of Reasons’ (2015). For a fairly recent general overview (focused on the closely related view of reasons internalism), see (Lord and Plunkett 2017).
But, intuitively, the husband does have a reason to stop abusing his wife. This intuition about the normative reasons of the abusive husband can be bolstered by thinking about the implications of some adjacent moral intuitions. For example, I also find it intuitive that the husband is blameworthy for failing to treat his wife better. And his blameworthiness seems to imply that the husband ought to treat his wife better. After all, if it were false that ought to treat her better, what would we be blaming him for? The fact that he ought to treat her better, in turn, seems to imply that the husband has sufficient reason to treat his wife better. When we say that someone ought to do something because it’s morally important, most of us think it would be irrational of them not to. Objections have been made to this line of inference, but I don’t think it is exceedingly presumptive to say that, if Humeanism does have these implications, it is in tension with the ‘common sense’.

Less commonly, the amoralism objection also seizes on intuitions that Humeanism conflicts with our moral intuitions by generating too many reasons – reasons that, intuitively, we don’t have. On the Humean view, if an agent has some desires then she ordinarily has some related reasons for action, assuming some potential action would promote the satisfaction of one of her desires. But without any substantive constraints on what reasons agents could have (constraints which Humeans reject), Humeanism seems to entail that bad

39 Some revisionary subjectivists such as Bernard Williams (1981) might take this to imply that we simply have fewer genuine moral reasons than we ordinarily think. The disparity between our profligate moral reason ascriptions and the more austere truth can be explained by a combination of moral ‘bluff’—that is, false moral reason-ascriptions—and the ‘proleptic’ use of moral reason-ascriptions and blame—that is, the very acting as though an agent has a reason and is blameworthy provides them with a reason. But, as I’ve previously mentioned, a conservative subjectivist cannot shrink the set of moral reasons in this way.
or defective desires also provide reasons—‘garbage-in, garbage-out’. The amoralism objection focuses on cases where Humeanism counter-intuitively assigns us normative reasons to act immorally.\textsuperscript{40} Eden Lin (2015, 226) describes an agent who, sitting in a room with a trigger, possesses only the desire to see an explosion which they can achieve by pressing the trigger (doing so would happen to kill a group of school children, an outcome to which the agent is indifferent). Humeanism seems to imply Lin’s monomaniacal explosion desirer has normative reasons he intuitively lacks: reasons to kill the children.

What argumentative strategies are open to the Humean in response to the amoralism objection? I have already mentioned that one, more popular in the seventies than now, is to bite the bullet – to claim that Humeanism is revisionary yet true.\textsuperscript{41} Unsurprisingly, however,

\textsuperscript{40} There are also versions of the ‘too many reasons’ objection that don’t focus on moral issues. For example, there is discussion of ideally coherent eccentrics (Street 2009) who have a few bizarre desires, such as ‘to count blades of grass’ (Rawls 1971, 422). Humeanism seems to imply that ideally coherent eccentrics have normative reasons to pursue the objects of their bizarre desires, whereas common sense says they have no such reasons. Another category of example meant to bring out the ‘too many reasons’ objection involves imprudent desires. Allan Gibbard (Gibbard 1990, 166) describes an ideally coherent anorexic. Humeanism seems to counter-intuitively imply that she has no reason to eat given her desire to be thin and her lack of any desire to remain alive. But my focus here is on the ‘too many reasons’ objection in the form of the amoralism objection.

\textsuperscript{41} Bernard Williams argues that if there’s no relevant feature in his psychology, an abusive husband has no reason to treat his wife better (Williams 1981). He’s brutal, sexist, nasty, etc. We might still have good reason to take action—to lock him up or send him to therapy. But he’s not irrational. For revisionary responses see Foot (1972), Harman (1975), Williams (1981) (1985), and Street (2012).
most philosophers think that if Humeanism has these revisionary moral implications then it is false. Anti-Humeans conclude the view is false, while conservative Humeans contend that the view does not have these counter-intuitive implications (see esp. Schroeder 2007). Conservative Humeans claim that the view, properly understood, can explain why we have the moral reasons we intuitively have, and why we lack the immoral reasons we intuitively lack. My response to the amoralism objection pursues the conservative approach.

3.3. Humeanism and Moral Rationalism

Let’s think for the moment about the version of the amoralism objection that claims Humeanism explains too few moral reasons. It’s the more obvious and influential of the extensional problems the amoralism objection raises. To get some purchase on the target the conservative Humean needs to hit to have accommodated common sense intuitions, let’s think more about the connection between what morality demands of us and what normative reasons to act we think we have. If someone has done something terrible, something we think they are blameworthy for, we think they ought to have not done that thing. And if we think somebody ought not to have done something, we think she had a normative reason to refrain

\[\text{42} \] I take Schroeder (2007) to be defending conservative subjectivism in precisely the sense I mean. I take Velleman (1989) (1996) (2009) and Markovits (2014) to be defending views very close to conservative subjectivism. Markovits calls her view ‘Kantian internalism’. The conception of practical rationality Markovits uses is indeed Kantian, but her view still explains agents’ reasons through their valuings (and what those valuings rationally commit agents to). Velleman’s ‘kinda Kantian’ view is that there is something that all action constitutively aims at. Since this aim is sub-personal, i.e. not the agent’s, Velleman’s view is not the standard Humean account I’m thinking of here.
from doing it, a normative reason that she ignored. So, intuitively, whenever morality
demands that we do some action, we tend to think we have fairly serious normative reasons
to do so. These kinds of intuitions recommend the thesis commonly known as moral
rationalism. Errol Lord and David Plunkett have argued the following formulation of moral
rationalism is likely the most intuitive and plausible:

**Moral Rationalism.** Necessarily, if there is a moral reason for A to \( \phi \), then there is a
normative reason for A to \( \phi \) and, at least sometimes, normative reasons provided by moral
reasons have significant weight. (Lord and Plunkett 2017, 330)

For the remainder of this chapter, I’ll focus on how a Humean can accommodate
moral rationalism (and thereby answer the amoralism objection).

4. Answering the Amoralism Objection

4.1. Section Outline

In this section, I discuss conservative solutions to the amoralism objection. These are
solutions that seek to show how Humeanism can accommodate moral rationalism. The
second half of this section defends the viability of a strategy called ‘Common Desires’. I
engage with and respond to an argument to the contrary by Schroeder which claims Common
Desires wouldn’t explain the counterfactually robust modal status of moral reasons. I claim
that a constitutivist version of Common Desires, which I call ‘Constitutive Common Desires’,
can address Schroeder’s objections.

4.2. A Summary of Proposed Solutions

Theoretically, how could conservative Humeans accommodate moral rationalism?
What options are there? Given the content of the theory, Humeans must say that whenever an
agent has a moral reason to phi she also has some reason-explaining desire. And, given methodological conservativism, the conservative Humean must deny that we have fewer moral reasons than ordinary intuition suggests. So the only way the conservative Humean can accommodate moral rationalism is to claim that we all have desires which explain our intuitive reasons to act morally. Let’s call those desires that explain reasons to act morally ‘pro-moral’ desires. To put it neatly, for conservative Humeanism to be true, it must be the case that we all have pro-moral desires.

One approach to defending the claim that we all have pro-moral desires is to alter the conative franchise—the set of desires which are reason-providing. In the simplest version of Humeanism, the agent’s actual desires at the time of analysis explain her reasons. It’s hard to see how moral rationalism can be accommodated in this version of the view. If an agent happens not to have any pro-moral desires at the time of analysis, she won’t have any reason to act morally. But we might, for example, democratise Humeanism by expanding the set of reason-explaining desires to the desires of others—all those affected by the set of possible actions, for example (Manne 2016). Or, we might include an agent’s future desires in the reason-providing set (Heathwood 2011) (Lin 2020).

The most popular version of the ‘alter the conative franchise’ approach is to make the set of reason-providing desires an agent’s reason-providing desires counterfactual desires. If a man who falsely believes that a glass of petrol is a glass of gin desires to drink the glass’s contents, his desire doesn’t seem to provide him any reason to drink the contents of the glass (Williams 1981). Perhaps only the desires we would have if we were fully informed or

43 The obvious worry is that these alternations undermine the original motivations for having a desire-based view in the first place–internalist concerns about the relevance of an agent’s actual psychology to her reasons, for example.
relevantly informed are reason-providing. What if an agent has a sudden strong desire to jump in front of a train, despite having an otherwise normal psychology and happy life? That desire also doesn’t seem reason-providing. A moment’s reflection would surely make him realise that he has much to live for, that he would be much better served by not jumping in front of the train, and likely even eliminate the desire altogether. Perhaps only those desires we would have if we deliberated sufficiently, were fully coherent, or underwent ‘cognitive psychotherapy’ (Brandt 1979) are reason-providing. Michael Smith (1994) provides a particularly sophisticated version of this strategy according to which an agent’s reasons are explained by the desires of her ideal counterpart who is playing an advisory role to her non-ideal counterpart.

Idealised Humeanism about normative practical reasons is attractive not only because it promises more plausible results, in terms of the reasons that it implies exist (and don’t). It is also attractive because it captures something intuitive states of mind that intuitively count when we think about what we ought to do. It is ordinary for people to take great care attempting to exclude ‘defective’ desires from their decision-making. For example, a person who is giving up smoking will not think the desire for a cigarette provides a good reason to have one and might attempt to suppress or eliminate the desire. Such cases suggest that our actual desires are not necessarily the only ones that could be reason-providing.

All versions of the conative-franchise-altering approach face the same two serious challenges: (1) to show the proposed set of desires gets the right results and (2) to explain why that set of counterfactual desires is the reason-providing set, from desiderata that are internal or friendly to the Humean account of reasons so as to avoid the charge that the idealisation is ad hoc.

Schroeder (2007) has developed his own influential approach to defending the thesis that we all have pro-moral desires. Instead of tinkering with the conative franchise, Schroeder
defends the thesis that moral reasons are ‘massively overdetermined’. Let’s follow Schroeder in calling this claim ‘Overdetermination’. How might Overdetermination be true? It might be that reasons are incredibly ‘cheap’ – easily generated given one’s theory of reasons. This can be achieved by positing a minimal connection between agents’ desires and their reasons, according to which a desire explains something’s being a reason to \( \text{phi} \) just in case \( \text{phi} \)-ing promotes (raises the probability) of the desire’s satisfaction to any extent whatsoever above the baseline (set by not \( \text{phi} \)-ing). Schroeder’s memorable example is that the fact that a car contains the recommended daily intake of iron means you have a reason to eat it (Schroeder 2007, 95–6). If reasons are so cheap, Schroeder argues there would likely be some kinds of actions which promote just about any desire an agent might have. For example, you’re more likely to satisfy your desires if you form true beliefs. So, if you have any desires, then you have a reason to form true beliefs. Schroeder is optimistic that at least some moral reasons are overdetermined in the same way, and hence that you have reasons to be moral if you have just about any desires whatsoever.

Schroeder now says Overdetermination is unnecessarily strong and thinks a weaker position suggested by Shafer-Landau (Shafer-Landau 2012, 438), ‘(P)’, is more plausible and sufficient for his argument (Schroeder 2012, 469). (P) is the view that there only needs to be some desire that explains moral reasons.

The contrast between Overdetermination and (P) is this. Imagine an agent with four desire-like states of mind: D1-4. Overdetermination claims that acting morally promotes D1, D2, D3, and D4. (P) claims that acting morally promotes either D1, D2, D3 or D4. In other words, according to Overdetermination, acting morally promotes the satisfaction of (pretty much) every desire an agent has. Therefore, all agents have normative reasons to act morally. Whereas, according to (P), acting morally promotes the satisfaction of some desire an agent
has (i.e. at least one, but not all). But this less ambitious claim is sufficient to entail that all agents have normative reasons to act morally.

Fundamentally, though, the structure of the solution (P) offers remains the same as Overdetermination. Instead of focusing on establishing that agents have the right set of desires to get reasons they intuitively have, the Overdetermination strategy focuses instead on getting the reasons agents intuitively have via the ‘promotion relation’, the relation between desires and reasons. The weaker that relation, the easier it is for the Humean to generate reasons, including reasons to be moral.

4.3. Common Desires

A final way of defending conservative Humeanism is, in a certain sense, the simplest. Like Schroeder’s strategy, it does not seek to alter the conative franchise. This strategy is based on the claim that all agents share the same particular pro-moral desires. Let’s call this view Common Desires.\(^\text{44}\) Common Desires is logically compatible with both Overdetermination and (P). It might be that moral action promotes the satisfaction of just about every desire (Overdetermination) \(\text{and that there are particular desires that all agents share,}\) the satisfaction of which is also promoted by moral action (Common Desires). Or, it might be that every agent has some desire the satisfaction of which is promoted by moral action (P) and this desire (or these desires) is (or are) the same particular desire(s) in all agents (Common Desires).

More than anything, the difference between Common Desires and Schroeder’s positions lies in their respective argumentative strategies for accommodating moral rationalism. On both of Schroeder’s approaches, Overdetermination and (P), what agents

---

\(^{44}\) I take this term from Eden Lin (2015).
happen to desire in particular isn’t of central importance. Given the weak promotion relation, something or other will be promoted by moral action. This entails we have normative reasons to act morally, as moral rationalism claims. In contrast, Common Desires focuses on establishing what desires are promoted by moral action and, crucially, establishing that all agents have these particular pro-moral desires. For this reason, Common Desires is a stronger claim than (P), and perhaps Overdetermination, and it therefore has a greater explanatory burden. Namely, Common Desires must explain how it is that all agents have the same particular pro-moral desires. But Common Desires also has a benefit; it does not require as weak (and arguably implausible) a promotion relation as Overdetermination or (P) which may over-generate reasons in implausible ways. My solution to reconciling moral rationalism and Humeanism is a version of this approach.

4.4. Common Desires and the Modal Objection

Here’s a first pass at stating Common Desires more precisely:

**Common Desires:** There is some finite set of desires that every agent has, such that for any agent S and any action A that S is morally required to perform, S’s doing A would promote the satisfaction of one of the members of that set. (Modified from Lin 2015, 224)

Common Desires might be taken as a kind of empirical, anthropological claim about the agents we know – psychologically normal adult humans. We could tell some evolutionary story that explains why humans share certain desires or why all cultures (that have endured this far) cultivate certain desires, and why these shared desires explain moral reasons.

This would already be enough to get the extension of actual agents’ reasons right. We would be able to explain why all of us have normative reasons to do what morality demands. But we would still not have accommodated moral rationalism. Moral rationalism claims that necessarily, if there is a moral reason for A to \textit{phi} then there is a normative reason for A to
phi. But the fact that all agents have normative reasons to phi in every case they have moral reasons to phi does not establish that this connection is necessary. As Schroeder puts it, ‘The real trick for the Humean Theory is not to account for the agent-neutrality of reason ascriptions, or the universality of reasons. It is to account for the modal statuses of reasons’ (Schroeder 2007, 107).

This modal version of the amoralism objection is a much harder problem for the Common Desires strategy to solve. In the Humean view, if someone woke up tomorrow lacking pro-moral desires then she would also lack normative reasons to do what morality demands. Or, if one thinks that pro-moral desires happen to be universal amongst humans, imagine an alien who met the criteria for rational agency but lacked any pro-moral desires. Humeanism says this alien would have no normative reason to do what morality requires. Hence, even assuming Common Desires, Humeanism doesn’t seem capable of accommodating moral rationalism because it doesn’t capture the modal statuses of moral reasons.

4.5. Constitutive Common Desires and the Modal Objection

A constitutivist version of Common Desires deals better with the modal status of moral reasons. According to Velleman (1989) (1996) (2009), there are certain aims which are partly constitutive of action. That is, one can’t count as performing actions unless one has (or one’s actions have) these aims. It is partly having these aims that makes behaviour count as action. The commitment constitutivism I outlined in the previous chapter is a similar view, though I think it fits together more neatly with the Humean account of normative reasons. According to commitment constitutivism, there are certain practical commitments the undertaking of which is partly constitutive of rational agency. The adoption of these practical
commitments makes agency possible. The commitment constitutivism this thesis is
developing is a version of a more general view, which can be defined as follows.

**Constitutive Common Desires.** There is some finite set of constitutive desires that every
possible agent has, such that for any possible agent S and any action A that S is morally
required to perform, S’s doing A would promote the satisfaction of one of the members of that
set.

I’ll outline and defend my preferred version of Constitutive Common Desires
(commitment constitutivism) in Chapter Five. For the moment, suppose that Constitutive
Common Desires is true; some desires are constitutive of agency and these desires are both
reason-providing and pro-moral. Call the normative reasons to act morally which are
explained by agents’ constitutive commitments ‘constitutive moral reasons’. It would follow
that constitutive moral reasons are necessarily shared by all possible agents. Such reasons
would be universal in a fairly modally robust sense. But the crucial question is: would
Constitutive Common Desires get the modal status of our moral reasons right?

Schroeder argues not, and that this is a ‘structural flaw’ of this strategy (Schroeder
2007, 107). His objection depends on a distinction between weak modal status and strong
modal status (105-6). Schroeder claims that moral reasons have both weak and strong modal
status. Constitutive Common Desires fails as a conservative Humean response to the
amoralism objection because it can’t accommodate strong modal status.

What is strong modal status? An agent A’s reason to *phi* has strong modal status if
and only if A ‘would have [a reason to *phi*] even he did not have that [reason-explaining]

45 It might be wondered how Schroeder’s weak and strong modal statuses compare to the more
traditional Kantian contrast between hypothetical and categorical normativity. I’ll return to this
shortly.
desire’. In other words, a reason has strong modal status if its existence does not depend on the existence of the desire which explains it.\textsuperscript{46} In other words, a reason has strong modal status if ‘there is no single desire on which it depends’ (109). Schroeder claims that constitutive moral reasons wouldn’t have strong modal status:

\begin{quote}

It would not be true, on this view, that even if you didn’t have the desire that all agents necessarily have, there would still be a reason for you to $\phi$. (109)

\end{quote}

This is correct. Constitutive Common Desires can’t accommodate any reasons with strong modal status; if someone lacked the constitutive desire she would, indeed, lack the corresponding reason(s).

I see three problems with Schroeder’s objection. The first: his objection has no dialectical force because it begs the question against Constitutive Common Desires. The second: strong modal status is only possible on a tendentious account of the truth-conditions for counterfactuals. The third: there is no compelling motivation for the view that moral reasons actually have strong modal status; ‘weak’ modal status is enough to accommodate our ordinary intuitions about moral reasons and moral rationalism. And Constitutive

\begin{flushright}

\textsuperscript{46} Strong modal status might seem strictly incompatible with the Humean theory of reasons. As Schroeder says, ‘nearly all philosophical discussions have taken this for granted’ (108). But Schroeder has shown that they’re theoretically compatible. If we can guarantee that there is always some other desire that also explains the agent’s reason to $\phi$ (if e.g. Overdetermination is true), then that reason would have strong modal status. The ability of Overdetermination to accommodate the strong modal status of moral reasons (and Common Desires’ inability) is a key motivation for Schroeder’s defence of the former as the most promising strategy for defending conservative Humeanism.

\end{flushright}
Common Desires *can* accommodate weak modal status. Hence, Constitutive Common Desires can get the modal status of moral reasons right and provide a conservative defence of Humeanism in response to the amoralism objection.

Let me first explain how Schroeder’s objection to Constitutive Common Desires is question-begging. Schroeder’s argument assumes the truth of the constitutivist view and attempts to show it doesn’t generate strong modal status for moral reasons. But this argument implicitly assumes the falsity of Constitutive Common Desires.

To find whether a reason to *phi* has strong modal status, we perform a counterfactual test. We remove the desire which explains the reason and then we ask ourselves whether the agent still has a reason to *phi*. If he does, we say the reason has strong modal status. If he doesn’t, we say the reason lacks strong modal status. But, on the constitutivist view, since our having these pro-moral desires is partly *constitutive of our agency*, to stipulate that a being lacks such a desire is to stipulate that the individual *isn’t an agent*.47 But this implies that in Schroeder’s constitutive-desire-free counterfactual, you simply aren’t an agent.

This is a problem. My intuition is that non-agents have neither moral nor normative reasons for action. Would you have a normative reason to help a drowning child even if you were a rock? Rocks aren’t the kinds of things to which we want to ascribe reasons, and the same goes for non-agents generally.48 It makes no sense to say that a being A still has a

47 Schroeder acknowledges the possibility of this kind of response in a footnote (2007, 107, fn. 8), but he seems to think it is inconclusive as a response. I disagree.

48 We do sometimes say that certain non-agents have normative reasons to do things. For example, a company might have good reason to downsize during a recession. While some might view companies as agents, I take such claims to be metaphorical. We are treating the company as sort-of-agent-like for the sake of minimising cognitive load, for instance. Or
reason to *phi* even if A is not an agent. The test we use to determine whether reasons have strong modal status can’t be used to see if Humeanism has counter-intuitive implications, because passing the test entails that Constitutive Common Desires is false.

My diagnosis is that Schroeder is unintentionally failing to take the constitutivist proposal seriously. Since he thinks strong modal status is true, he must think that a being who lacked the constitutive desires of agency could (and would) have reasons to be moral. But, since only agents have reasons, Schroeder must be implicitly committed to the view that the constitutive-desire-lacking being is still an agent—the kind of being to whom we want to ascribe reasons. But to hold that assumption is to beg the question against the constitutivist.

A second, closely related problem is the mysteriousness of the truth value of the counterfactual at the heart of Schroeder’s objection. Recall that a reason to *phi*’s having strong modal status (or not) is determined by the truth of a counterfactual—that A would still have that reason to *phi* if she lacked the desire which, in actuality, explains her reason to *phi*. According to David Lewis’s (1973) influential account, counterfactuals are assertions whose antecedents are false, and whose consequents describe how the world would be if the antecedent had been obtained. According to Lewis, counterfactuals quantify over a set of nearby possible worlds. So, to determine the truth of a counterfactual, we ask if the consequent is true in the most nearby possible worlds in which the antecedent is true. If it is, then the counterfactual is true. But what happens if we can’t find *any* nearby possible worlds in which the antecedent is true? That is, what if the set of possible worlds over which our counterfactuals quantify contains no members with a true antecedent? Then the antecedent is

maybe some people really do think of them as agents, in which case they presumably think they have the constitutive features of agents.

49 My thanks to Michael Ridge for raising this point.
necessarily false, and the counterfactual is trivially true. The counterfactual is trivially true in the sense that all nearby worlds in which the antecedent is true (this being a null set!), are such that the consequent is true.

Running Schroeder’s objection through Lewis’s account of counterfactuals, the counterfactual underlying strong modal status is trivially true. Strong modal status claims that, necessarily, if an agent lacks the desire which actually explains her reason to \( \phi \), then she will still have a reason to \( \phi \). So the counterfactual’s antecedent is something like ‘the agent lacks the reason-to-\( \phi \)-explaining desire’, and its consequent is ‘the agent has a reason to \( \phi \)’. But, if the constitutivist is right, as Schroeder accepts for the sake of his objection, then there are no nearby possible worlds in which the antecedent, ‘the agent lacks the reason-to-\( \phi \) explaining desire’, obtains. Hence, the counterfactual underlying strong modal status is trivially true and the objection (which relies on the non-trivial truth of strong modal status) fails.

Of course, this response relies on the truth of Lewis’s account of counterfactuals. But Schroeder or other critics of the constitutivist approach may be relying on a non-Lewisian theory of counterfactuals which avoids these problems.\(^{50}\) I accept this is a possible move, but it is at least a limitation that Schroeder’s objection relies on particular (non-orthodox) views about counterfactuals.

\(^{50}\) Some philosophers claim that counterpossibles (counterfactual conditionals with impossible antecedents) are always or sometimes true. Williamson (2007) (2010), for example, defends the view that all counterpossibles are true. There are also so-called nonvacuist views that accept counterpossibles but claim they are straightforwardly false (Bjerring 2014), which would be equally problematic for objections like Schroeder’s.
Perhaps we can understand Schroeder’s modal objection more simply. He thinks moral reasons have strong modal status, but this can’t be true given Constitutive Common Desires for the reasons just mentioned. But, insofar as we think that moral reasons have strong modal status, we ought to reject the constitutivist Humean approach and embrace an alternative approach that is compatible with strong modal status, like Schroeder’s Overdetermination. Or, at the very least, if there is any plausibility to strong modal status, then that is enough to give, as he says in a footnote, ‘an advantage’ to his approach over the constitutivist approach (2007, 107, fn. 8).

This less ambitious argument rests on the attractiveness of strong modal status. But it is not even clear that strong modal status is attractive. The modal status of moral reasons may be fairly strong. It may be that not just all actual agents have normative reasons whenever they have moral reasons; it may be that all possible agents have normative reasons whenever they have moral reasons. This is the kind of modally robust position moral rationalism seems to entail. But the important point is that the modal status of reasons is not plausibly so strong that it extends beyond the only class of beings metaphysically capable of having moral reasons — rational agents. Assuming Constitutive Common Desires, strong modal status is impossibly strong.

We appear to be at an argumentative impasse. Schroeder’s objection assumes (or gives some level of credence to) strong modal status and therefore rejects (or gives some level of doubt toward) Constitutive Common Desires. I am assuming Constitutive Common

51 It may also be that some weaker modal claim is the right one. Instead of all possible agents we might think the right group is all possible agents with our kind of cognition or all possible psychologically normal adult humans. The latter would, presumably, result in some kind of Aristotelean position.
Desires and therefore rejecting strong modal status. If there is a deadlock here, we can at least conclude the objection is inconclusive. But I think we can do better than that. In the fifth chapter, I provide an independent argument for Constitutive Common Desires. If this argument succeeds, then Humeans ought to reject the incompatible strong modal status of moral reasons. For now, though, I’ll simply bypass this deadlock, as it turns out that answering the amoralism objection does not require our moral reasons to have strong modal status.

4.6. What is the Modal Status of Moral Reasons?

What is the intuitive position on the modal status of our moral reasons? Can Constitutive Common Desires accommodate this? I think it can. Schroeder claims that our moral reasons have ‘weak’ modal status (as well as strong modal status).\(^{52}\) I agree that our moral reasons, intuitively, have weak modal status. I also argue that if Humeanism can accommodate weak modal status (which most philosophical discussion has assumed it can’t), then it would accommodate our ordinary moral intuitions we distilled as moral rationalism.

A reason has \textit{weak} modal status if its existence is not predicated on the contingent fact that the agent just so happens to have the right, reason-explaining desires. Here is a reason that \textit{doesn’t} have weak modal status: the reason I have to go to the kitchen (given I want a slice of cake and there is cake in the kitchen). This reason doesn’t have weak modal status because it would no longer exist if I just so happened not to desire cake.

\(^{52}\) ‘Weak’ is a meaningful term in the way Schroeder uses it, to contrast with ‘strong’ modal status. But ‘weak’ modal status is misleading if read in an absolute sense, for reasons I will shortly explain.
It might seem that all possible reasons would lack weak modal status, on the Humean view. If one’s reasons are explained by one’s desires, then surely one’s reasons depend on what desires one happens to have. However, constitutive moral reasons, the reasons that follow from our constitutive pro-moral desires, would have weak modal status. Their existence as reasons wouldn’t be hostage to our contingent desires. Schroeder agrees with all this. Assuming the constitutivist is right, he says:

No matter what agents there were, holding fixed that it is a reason, it would be a reason for all of them. Such reasons would be explained in the same way for everyone by the desire that everyone shares. Since there couldn’t be an agent who didn’t have this desire, there couldn’t be an agent who didn’t have this reason. (Schroeder 2007, 107)

In an absolute sense, ‘weak’ modal status is actually fairly robust. If an agent has a reason with weak modal status, then she has that reason no matter what she contingently happens to desire. Most of the philosophical literature has always assumed Humeanism cannot deliver weak modal status. But if that assumption is wrong, weak modal status would seem sufficient to accommodate ordinary moral intuitions and moral rationalism, and thereby answer the amoralism objection. And that assumption is wrong if Constitutive Common Desires is true.

Furthermore, moral rationalism does not seem to demand anything stronger than weak modal status for our moral reasons. Nothing about the view entails that moral reasons have strong modal status. Recall:

**Moral Rationalism.** Necessarily, if there is a moral reason for A to phi, then there is a normative reason for A to phi and, at least sometimes, normative reasons provided by moral reasons have significant weight.

Moral rationalism contains a necessity claim, which is why Humeans must get the modal status of our reasons right, but it does not contain the kind of counterfactual claim strong modal status requires.
To sum up, how does Constitutive Common Desires accommodate moral rationalism and account for our moral reasons having weak modal status? According to Humeanism, an agent has a reason to \( \phi \) if and only if she has some desire that would be promoted by \( \phi \)-ing. According to Constitutive Common Desires, some desires are constitutive of agency and explain normative reasons to act morally. I’ve called these ‘constitutive pro-moral desires’ and ‘constitutive commitments’. It follows that whenever an agent has a moral reason to \( \phi \), she necessarily has (as a constitutive matter) pro-moral desires and these desires provide her with normative reasons to \( \phi \). The link between having moral reasons and having normative reasons is secured, and its necessity is also explained. As the only way we could lack these reasons to act morally would be to lack these pro-moral desires, and as these pro-moral desires are constitutive of agency, the only way we could lack these reasons is if we didn’t have reasons altogether (in virtue of not being agents). This means that our normative reasons to act morally have weak modal status. Their existence doesn’t depend on our just having, as a contingent matter, the right desires. Constitutive Common Desires accommodates moral rationalism and gets the modal status of our normative reasons to be moral right.

4.7. The Categoricity of Morality

A closely related challenge is if Constitutive Common Desires can enable Humeanism to capture morality’s ‘categoricity’.

Kant’s distinction between categorical and hypothetical imperatives is between the conditions under which an imperative can be considered normative. Hypothetical imperatives are only normative conditional on the adoption of some end, whereas the categorical imperative is not conditional the adoption of any end. ‘Do not steal’ is a categorical imperative, whereas ‘Do not steal if you want to stay out of jail’ is a hypothetical imperative.
The worry is that even given this constitutivist version of Common Desires, we are still giving moral reasons this second, hypothetical kind of normativity; act morally if you have pro-moral desires. But the constitutivist claim means things aren’t this simple. If the possession of pro-moral desires really is partly constitutive of rational agency, then we’re also saying ‘act morally if you are a rational agent’. Even Kant thought the categorical imperative is only valid for rational agents. Adding the condition ‘if you are a rational agent’ to a categorical imperative does not thereby make it a hypothetical imperative. So to the extent morality can meaningfully be understood to be categorical, the constitutivist version of Common Desires is accommodating.

5. Conclusion

What has this chapter established? On structural grounds, Common Desires has been unduly dismissed as the basis of a viable defence of conservative Humeanism in the face of the amoralism objection and, more specifically, in the face of a modal version of the amoralism objection. I’ve argued that the constitutivist version of Common Desires is capable of generating normative moral reasons for all agents and, much more interestingly, getting the right modal status for those reasons. This chapter has merely shown that if Constitutive Common Desires is true, we are in a promising position to reconcile Humeanism and moral rationalism and thereby answer the amoralism objection. Of course, this conclusion is entirely hypothetical. Constitutive Common Desires needs to be true for it to form the basis of a conservative defence of Humeanism. In the next chapter, I turn to the final part of moral rationalism that requires Humean accommodation: its claims about the special weightiness of moral reasons.
Chapter Three: The Weight of Moral Reasons

It is more important that our theory fit the facts than that it be simple.

—W.D. Ross, *The Good and the Right* (1930, 19)

1. Introduction

In the previous chapter, I argued that, if Constitutive Common Desires is true, if there are constitutive practical commitments the possession of which is partly constitutive of rational agency, then (1) Humeanism can answer the most influential form of the too few reasons objection—it can demonstrate moral reasons where they intuitively exist\(^3\) and (2) Humeanism can get the modal status of moral reasons right—it can capture the intuition that all possible rational agents have some normative reason to \(phi\) whenever they have a moral reason to \(phi\).

Allow me to introduce another challenge. We tend to think moral reasons out weigh other reasons—reasons of aesthetics, prudence, politics, etc. Recall the thesis that encapsulates the intuitions the conservative Humean seeks to accommodate:

**Moral Rationalism.** Necessarily, if there is a moral reason for A to \(phi\), then there is a normative reason for A to \(phi\) and, at least sometimes, normative reasons provided by moral reasons have significant weight. (Lord and Plunkett 2017, 330)

Even if Constitutive Common Desires can reconcile Humeanism with the first half of moral rationalism (‘Necessarily…for A to \(phi\)’), it may not reconcile it with the second

---

\(^3\) Admittedly, the ‘too few reasons’ is not just about moral reasons—my focus. Nonetheless, solving the moral version of the problem would be a significant result and does not exclude the possibility of a similar solution for e.g. prudential reasons.
half—that moral reasons have significant weight. The final obstacle to reconciling moral rationalism and Humeanism I wish to address is that Humeanism must explain why moral reasons are generally weighty. I argue here that, in addition to the extension and modality of moral reasons, a constitutivist version of Humeanism can also deliver a plausible account of the special weight of moral reasons.

I first outline what I take to be a plausible account of the weight of moral reasons the conservative Humean needs to accommodate (section 2). I then explain and discuss Schroeder’s solution, which is to reject a Humean thesis about reasons’ weight called proportionalism (section 3). I argue that there are compelling reasons for Humeans to avoid abandoning proportionalism if it is possible to do so. I then show how a constitutivist version of Humeanism can get the weight of moral reasons right even assuming proportionalism (section 4). Finally, I respond to an objection to my account according to which my solution outlined in section 4 rests on an ambiguity between desires and intentions (section 5). My solution involves a more detailed account of the conative mental state, the possession of which is partly constitutive of agency, on the Humean constitutivist account.

2. The Intuitive Weight of Moral Reasons

In the previous chapter, I took up the task of reconciling the Humean Theory of Reasons with our ordinary intuitions about moral reasons. Specifically, I focused on establishing that Humeanism can be reconciled with

Moral Rationalism. Necessarily, if there is a moral reason for A to \(phi\), then there is a normative reason for A to \(phi\) and, at least sometimes, normative reasons provided by moral reasons have significant weight. (Lord and Plunkett 2017, 330)

So far, I have only argued that Humeanism can accommodate the first half of this thesis—that necessarily, if there is a moral reason for A to \(phi\), then there is a normative
reason for A to \(\phi\). I now want to turn to the second half—that, at least sometimes, normative reasons provided by moral reasons have significant weight.

The Humean offers an account of normative practical reasons. But our normative reasons have weight. Our reasons can be stronger or weaker, more or less important. We do not simply have intuitions about what normative reasons agents have; we also have intuitions about how much weight those reasons have. The conservative Humean must be able to accommodate these intuitions.

First, a brief note about what it means for normative reasons to be weighted is required. Normative reasons must all have some weight; a reason to \(\phi\) with no weight whatsoever would be no reason to \(\phi\) at all. Reasons can also be outweighed by other reasons. So all normative reasons have some weight, and it seems conceptually possible for all normative reasons to be outweighed by other reasons or coalitions of reasons. It might be objected that there is a kind of reason that is plausibly incapable of being outweighed: the decisive, conclusive, or overwhelming reason. Decisive reasons are indeed never outweighed, otherwise they would not be decisive. Hence, decisive reasons always determine the thing to do—the thing that one has an all-things-considered reason to do. If you (somehow) know that your reason to \(\phi\) is decisive, you can rationally skip the arduous task of weighing your reasons and simply start \(\phi\)-ing. But decisive reasons are not decisive because they have some special intrinsic property of decisiveness. Rather, the decisiveness of a reason must be a relational property. A reason \(R\) has the property of being decisive if and only if there is no other contrary reasons that collectively outweigh \(R\). Given that decisiveness is a relational property, it follows that even the reasons that happen to be decisive could be outweighed, if the relevant facts that made it decisive were different. If the relevant facts were different, then another reason would be decisive. The next question is whether, as a matter of actual fact,
moral reasons (or some subset of moral reasons) are systematically decisive such ‘that an
action being morally required is decisive reason to perform it.’ (Dorsey 2012, 5)

Before seeking to show that Humeanism can accommodate intuition about reasons’
weights, I want to flesh out the intuitive position. It is no good to argue that the Humean can
accommodate intuition if I end up reconciling Humeanism with an unintuitive position. So we
first need to make some progress on the following questions. When do normative reasons
have significant weight, and what are the exceptions, if there are any? My focus will be on
the most challenging kind of reason for the Humean—moral reasons. Just how weighty are
they compared with our other reasons? Are they always of greater weight? Are they always
overriding/conclusive reasons or are moral reasons merely pro tanto?

Moral rationalists believe that moral reasons are something like heavyweights of the
realm of practical rationality: weight-advantaged compared to other kinds of reasons—
prudential or aesthetic reasons, for example. This certainly reflects something intuitive. It
doesn’t matter how much you want the baby’s candy, nor how much you will enjoy eating it,
nor how easy it would be; you shouldn’t take candy from a baby. The moral reason we have
to refrain from taking candy from a baby plausibly outweights the reasons that favour taking
it.

I’ll briefly introduce the two theses defined by Brink (1997) that arguably define
moral rationalism—Authority and Supremacy. The Authority thesis claims that there is
always a pro tanto normative reason to behave morally. In other words, our moral reasons
have some normative weight—moral reasons are always normative reasons. The Supremacy
thesis is that moral requirements are also rational requirements. Whenever our moral reasons
demand we phi, we have most all-things-considered reason to phi. Supremacy holds, roughly,
that our moral reasons are not just normative, and not even just particularly weighty, but that
they are collectively overriding in determining what we ought to do, all things considered.
I’ll consider three responses, in broad strokes. What I’ll call *anti-rationalism* rejects both Authority and Supremacy. What I’ll call *strong moral rationalism* accepts both theses. There are also philosophers who take a middle ground. Dale Dorsey (2012), for example, endorses Authority and rejects Supremacy. I’ll refer to this combination of views as *weak moral rationalism*. So, which (if any) of these combinations has it right? Which of them should the conservative Humean be aiming for?

2.1. Anti-Rationalism

Some philosophers, some Humeans in fact,\(^{54}\) deny moral rationalism. A full discussion of all the anti-rationalist options here, from moral or normative nihilism to moral voluntarism to existentialism to normative pluralism, is beyond the scope of my discussion. I am interested only in the options available to the conservative Humean. Therefore, my methodology will be conservative. I am interested only in accommodating ordinary intuition about moral reasons and their comparative weights, rather than offering a reforming, revisionary account of the reasons we possess (or lack).

This methodological constraint eliminates the more obvious anti-rationalist options—nihilism, for instance. Moral nihilism claims that there are no moral facts. Since an agent’s having a moral reason is a kind of moral fact, moral nihilism denies that there are genuinely normative moral reasons. This entails the falsity of moral rationalism. The same goes for the more general form of nihilism sometimes called *normative nihilism*, which denies there are any normative facts, and hence that there are any genuinely normative reasons for action.

---

\(^{54}\)I am thinking of those Humeans who embrace the apparently revisionary implications of the view such as the early Foot (1972), before her neo-Aristotelean turn, Williams (1981), and Street (2009, 2012). See the previous chapter for discussion.
Both these views are obviously deeply revisionary. We ordinarily think that there are genuine normative reasons for action and that some of these normative reasons are moral reasons.

To the conservative Humean, a more interesting option for rejecting moral rationalism is *normative pluralism*. Normative pluralism claims that there are multiple irreducibly normative practical domains of normativity (Copp 1997, 2009, Tiffany 2007). The reasons and recommendations of various domains sometimes conflict with one another in the practical guidance they provide, and there is no rationally justified all-things-considered normative judgment that can be made which incorporates the reasons of these various domains. There are only *moral* reasons to *phi*, and *aesthetic* reasons to not *phi*. But there is no *all-things-considered* reason to *phi* or refrain from *phi-ing*.

Normative pluralism entails that moral rationalism is false. Pluralists reject the possibility of there being a rationally justified method of weighing moral reasons against other kinds of practical reasons. They claim that there is no available shared standard or scale by which we could perform these intra-domain weight comparisons. In Copp’s words, ‘I reject the position that, in each situation, all the reasons there are determine one overall verdict, the verdict we might call the verdict of ‘Reason’ or ‘Reason-as-such.’ (Copp 1997, 86-87) But this overall verdict is precisely what is needed to weigh moral and non-moral reasons on a common scale. So moral rationalism, a thesis about the intra-domain comparisons of reasons’ weights, doesn’t make sense as positions, assuming normative pluralism.

To explain why normative pluralism contradicts ordinary intuition, it would be helpful to have a more precise account of the view. Tiffany has defined normative pluralism as the conjunction of two theses: contributory pluralism and deliberative pluralism (Tiffany 2007, 233). Contributory pluralism is the view that there are a variety of sources of contributory reasons for action (233) [… and there is] no such thing as (contributory) reasons-simpliciter.
Rather, there are only moral (contributory) reasons, prudential reasons, legal reasons, and so on. [... T]here is not some unique set of properties common to all instances of the favouring relation.’ (239) In other words, contributory pluralism denies that the reason relation is a single, universal favouring relation. The moral-favouring relation might be very different to the legal-favouring relation, for example. In a footnote, Tiffany explains ‘My claim is that the properties of any particular favouring relation are determined by its corresponding normative domain’ (239, fn 15).

Contributory pluralism is compatible with both moral rationalism and, arguably, with ordinary intuition; one can accept contributory pluralism and still think we can weigh different kinds of reasons against one another in a rationally justified manner. One might think, for example, that contributory reasons of these pluralistic normative domains combine into ‘coalitions of different kinds of reasons’ (Audi 2006, 183). On such a view, the most fundamental form of practical rationality, the form that determines what we rationally ought to do, would consist in weighing these various contributory reasons against one another and forming an overall judgment.

It is the second constitutive thesis of normative pluralism—deliberative pluralism—that both rules out moral rationalism and conflicts with ordinary intuition. This is the thesis that ‘there is no unitary standard for determining how to weigh different types of contributory reasons against each other’ (Tiffany 2007, 233). Deliberative pluralism entails that the most we can say is that an agent ought to phi according to the deliberative standpoint of morality and that an agent ought to refrain from phi-ing according to the deliberative standpoint of etiquette; no practical reasons possess ‘ultimate authority’. ‘To say that some consideration is a reason for A to phi in C’ writes Tiffany, ‘is merely to say that there is a normative framework which favours A’s phi-ing in C’ (2007, 233).
First, why does deliberative pluralism rule out moral rationalism? It denies the existence of a deliberative standpoint from which we derive ultimate normative justifications of what we ought to do—a standpoint from which we can compare and weigh up the contributory reasons of various domains into a singular, comprehensive, all-things-considered normative judgment. In the absence of this *all-things-considered ought*, moral rationalism is an ‘[im]possible [thesis] to coherently state’ (Tiffany 2007, 244). Why? Because moral rationalism is a thesis about the relationship between moral reasons and *all-things-considered* reasons.

Second, why does deliberative pluralism contradict ordinary intuition? Simply because it is fundamental to our ordinary practices to think, talk, and behave as though there can be rationally justifiable *all-things-considered* reasons to perform the actions we perform. Normative pluralism denies this.

To see why, let’s take an example. Imagine your friend is killed by an axe-murderer who was let into their workplace by your friend’s colleague. ‘Yes, I knew the axe-murderer was likely to kill someone if I let them in, and I knew I therefore had strong moral reasons for not letting him in’ the colleague explains, ‘But it would have been such bad manners!’ This is not just any failure; it is a failure (a terrible failure) of practical rationality.

Three related considerations support this interpretation. First, part of the ordinary response, besides tears and perhaps an urge for further violence, would be to try to discover if there was any rationality to the act, or if the colleague is incapable of practical reasoning. You might interrogate the colleague about his reasoning process, or try to make him understand that his relative weighing of the moral and etiquette-related reasons was rationally unjustified. None of this would make sense unless what the colleague has most *ultimate* reason to do is subject to rational justification. Further proof that the colleague’s actions would intuitively be taken as a rational failing is that it would be appropriate to adopt
Strawsonian reactive attitudes toward him, such as anger, blame, and resentment. We take the colleague’s poor weighing of these competing reasons to be a failure of practical reason and agency, rather than events caused by non-rational actors. In sum, our ordinary practices involve weighing inter-domain reasons and coming to comprehensive all-things-considered normative judgments. This practice of weighing and coming to all-things-considered judgments is a process which is subject to the standards of rational justification. Third, we often disagree about what agents have all-things-considered reason to do. If someone were to tell you that the colleague deliberated well and that the reasons of etiquette really did have greater weight than the moral reasons, you would presumably disagree. The content of the disagreement would be about the rational justification of a particular weighting of reasons in forming an all-things-considered judgment. Normative pluralism denies all this.

To be clear, Tiffany’s claim is not that we never make these kinds of ‘ultimate’ judgments about what we ought to do. He accepts that we assign some considerations ‘deliberative weight’ and not others. Rather, he rejects the notion that the contributory reasons of various normative domains can possess a “genuine” deliberative weight, understood as something more than a psychological claim about what in fact carries weight in specific instances of practical reasoning.’ (Tiffany 2007, 251) In other words, inter-domain judgments of reasons’ relative weights may occur, as a matter of contingent fact, but they are not under the jurisdiction of the norms of practical rationality. Rather, they can be justified only relative to these ‘normative standpoints’ that agents happen to occupy. Furthermore, there is no objective, rational justification for the normative standpoints we occupy, in Tiffany’s view. He writes:

When these standpoints conflict, e.g., when the most authentic course of action is imprudent, it is not clear to me that there is a single unqualified answer as to what one ought to do. Rather, it seems that the choice between authenticity and prudence is a matter of what I called ‘agential choice.’ (Tiffany 2007, 262)
An agential choice is to be understood as an ‘existential’ matter, a matter of radical choice. Hence, ‘there is no objective standard of correctness’ (Tiffany 2007, 244). It is this claim that I have argued is contrary to ordinary intuition.

Such considerations lead me to conclude that denying moral rationalism on the basis of normative pluralism is no option for the conservative Humean. If the conservative Humean wants to conserve ordinary intuition, he must accept moral rationalism, or find a better ground for rejecting it.

It might be objected that I’ve not given Tiffany’s view a fair shake, that its implications are more intuitive than I am suggesting. Tiffany discusses a thought experiment in which he thinks the normative pluralist is more closely aligned with our ordinary intuitions than the Moral Rationalist. Should Williams’ Gaugin leave his wife and family to become a great artist? Aesthetically, yes. Morally, maybe not (that depends on one’s view of moral luck). At least that is Williams’ reading of the case. But what should Gaugin do overall, *all things considered*? Tiffany claims it is:

55 Stated in more detail, my reading of Williams is as follows. The Gaugin example is part of his critique of the ‘morality system’ which renders moral luck a conceptual impossibility. Williams’ point is that something like the supposedly impossible moral luck exists. To Williams, this points out the limitations of our post-enlightenment conception of morality—that it doesn’t capture everything important about normative thought and practical justification. To put my interpretation of Williams’ point in Williams’ terms, Gaugin is ethically justified in moving to Tahiti for the sake of his art (via the mechanism of moral luck), where ethics encompasses (at least) the moral, the prudential, and the aesthetic. Ethics provides something like the ‘ultimate’ practical normative justification that Tiffany denies exists. Williams claims the ultimate practical normative question, what we might call the ultimate or all-things-considered question, is the broad ethical one—‘What should I do?’
counter-intuitive to hold that the artists who have such standards [prioritising art/aesthetic value over morality in their practical deliberations] are guilty of some deep confusion or normative blindness. While they certainly fall short as moral agents, it is not clear that they eo ipso fall short in some ultimate unqualified sense. In cases like these, I think normative pluralism gives the right answer. (Tiffany 2007, 246-7)

I find Tiffany’s interpretation of the case unconvincing. The Gaugin case is interesting to consider precisely because producing great art and taking care of one’s family are both considerations that deserve serious deliberative weight. The similarity in Gaugin’s reasons’ weights explains why an all-things-considered judgment is difficult to form. But it seems like a mistake to conclude from the lack of intuitive clarity about what Gaugin has all-things-considered reason to do as evidence for the normative pluralist’s claim that there is no possibility of a rationally justified ultimate judgment. In other words, my diagnosis is that the thought experiment has a confounding factor—the fact that each consideration is plausibly of a similar weight.

If normative pluralism were true, then we should expect to have the same intuition—that there is nothing the agent ought to do in the ‘ultimate, unqualified sense’—whenever we have conflicting reasons belonging to different normative domains. But in the vast majority of cases, this expectation is not met. For instance, think back to the two options provided to the overly-polite colleague discussed above. He can either keep the door shut and save your

56 This may seem to cut against moral rationalism. In my discussion of weak moral rationalism, I explain why this does not, in fact, cut against strong moral rationalism. In short, I’m taking morality as Williams took ethics, as the ‘master’ all-things-considered judgment. Both producing great art and taking care of one’s family are morally significant actions. Hence, the difficulty in deciding what to do is a result of a conflict between moral reasons, rather than a conflict between moral and non-moral reasons.
friend’s life or politely invite in the axe murderer inside, resulting in the death of your friend. Contra normative pluralism, we can seemingly make a clear all-things-considered judgment about what the colleague really ought to do, rationally speaking.

Whether normative pluralism is true is beyond the scope of the present discussion. The point I wish to make is merely that it is a revisionary view, and hence that the conservative Humean is entitled to exclude it from the set of options worth considering. Indeed, neither of the three anti-rationalist views considered here (nihilism and normative pluralism) is methodologically conservative.

2.2. Weak Anti-Rationalism (Rejecting Supremacy)

Should the conservative Humean accept Authority without Supremacy? Some Moral Rationalists\(^{57}\) endorse the lesser claim about moral reasons’ weights that (1) we necessarily have pro tanto normative reasons to act morally and (2) these reasons to act morally often though not necessarily have greater weight than other reasons. They claim the greater weight of moral reasons is a general tendency that permits exceptions.\(^{58}\) This is Lord and Plunkett’s

\(^{57}\) My own taxonomy renders these views versions of moral rationalism (following from Lord and Plunkett’s working definition of moral rationalism provided earlier), although some who have defended these views have called them ‘weak anti-rationalism’, e.g. Dale Dorsey (2012, 2016). Whether his view is best labelled weak anti-rationalism or weak rationalism is, of course, not important; I mention it only to avoid potential confusion.

\(^{58}\) Many arguments for such views are provided by consequentialists in response to the demandingness objection. They argue, roughly, that our moral reasons may indeed be extremely demanding, but that they can be outweighed by our non-moral reasons, e.g. our prudential reasons. Hence, we do not have all-things-considered reason to do counter-intuitive,
interpretation of moral rationalism which I have been using as shorthand for our ordinary
intuitions about moral reasons. I’ll call the view weak anti-rationalism.59

Weak anti-rationalism might seem like an escape hatch for the conservative Humean:
an easier target than strong moral rationalism. But I am unconvinced by weak anti-
rationalism. I therefore believe that the conservative Humean, to truly accommodate our
ordinary intuitions about moral reasons’ weights, must explain how strong moral rationalism
is true. In this subsection, I briefly survey weak anti-rationalism and explain why the existing
arguments for it are unconvincing.

I have already noted some of the reasons why philosophers have endorsed Supremacy. So why
might someone reject it? It is typical to arrive at a rejection of Supremacy by
endorsing practical dualisms of various kinds. For instance, Sidgwick suggests that both the
prudential and the moral provide normative reasons and that neither takes clear precedence
over the other (Sidgwick 1981). Similar ideas can be found in the work of Wolf, who hopes
that there is no such thing as a morally perfect agent, a ‘moral saint’ due to the deficits such a
character would have in other, non-moral domains (Wolf 1982). I take Wolf’s condemnation
of such a character to imply that we don’t have all-things-considered reasons to act morally,
or at least that we don’t have such reasons all the time.60 Crisp likewise claims that there are
cases in which it is not rationally justified to promote the good, and endorses a ‘dualism’
about practical reason on that basis (1996, 65).

59 I take this term from Dorsey (2012).

60 Dale Dorsey shares this interpretation of Wolf, and his discussion of these views has
informed my own (2012, 8).
These dualist views about practical normativity can be interpreted in at least two ways. They could either be claiming something similar to the normative pluralists I have already discussed, except that there are only two incommensurable kinds of normativity (rather than many), or they could be claiming merely that there are two sources of normativity.

Let’s take that first interpretation. Dualist views such as these do not just claim there are two sources of normativity. They also deny that there is an all-things-considered practical ought—an ought simpliciter.61 There is, for example, only what we ought to do morally and what we ought to do prudentially. But there is no possibility of an overarching, unifying rational justification of one’s judgment regarding how best to adjudicate the demands of these various normative domains. In other words, such views deny Supremacy, but they also deny Authority. Notice that normative dualism is just a form of the normative pluralism I’ve already discussed and labelled anti-rationalist.

Unsurprisingly, a normative dualism along these lines is liable to produce the same counterintuitive anti-rationalist implications as normative pluralism. I take it to be intuitive that there really is a matter of what one ought to do all things considered, beyond what one ought to do morally and what one ought to do prudentially. That is, while there are various normative domains, practical reason rationally governs our actions. In acting, we are forced to weigh the demands of various domains against one another. And doing so is not, intuitively, an act of ‘radical choice’ (i.e. a matter of rationally arbitrary choice), as some normative pluralists have claimed.

Normative dualists typically think the two kinds of normativity are moral and prudential, so let’s see if this improves anything compared to other kinds of normative

61 At least, this is my understanding of the aforementioned views.
pluralism. All we need to do is modify the axe-murderer case such that some prudential good is obtained by letting the murderer inside. Perhaps the murderer is offering a fantastic bribe to your colleague, a drug that would improve his life, or a lifetime subscription to a great therapist’s services. It still seems insufficient to say, ‘Sure I failed morally, but prudentially I did well, and there’s no way to say what I ought to have done overall!’ Intuitively, there’s a matter-of-fact about what rationality demanded of you in that situation (what morality demanded).

The second, more interesting interpretation of normative dualism yields more plausible results. Normative dualisms of this kind claim that (1) there are two sources of normativity, (2) that the domain-relative reasons they provide can be rationally weighed up against one another, (3) that a rationally justified all-things-considered normative judgment can be made which incorporates the reasons of these various domains, and (4) that it is possible for moral reasons not to rationally determine what we have all-things-considered reason to do—that the reasons of other domains can have greater or equal weight. Moral reasons, in other words, needn’t settle what we ought to do. They aren’t rationally dispositive. Of course, the dualism of normative dualism doesn’t seem important to any of these points. The same kind of view could posit three, five, or fifty sources of normativity. As long as the reasons of various domains are rationally commensurate, in the specific sense that it is a matter under the jurisdiction of practical reason how to weigh them against one another and come to an all-things-considered judgment, then any such view will accept Authority and reject Supremacy. It is this kind of view I call weak anti-rationalism.

Dorsey (2012) defends such a view. Here is his view stated more precisely:

*Weak Anti-Rationalism:* that a norm n is a moral requirement for a to x provides sufficient reason for a to x. However, in some cases, a moral requirement for a to x will not be rationally dispositive (given n, a will have sufficient, but not decisive, reason to x). (Dorsey 2012, 9)
A normative dualist of this flavour can explain why you ought not to let the axe-murderer inside. The moral reasons, in that case, are ‘rationally dispositive’ in Dorsey’s terms, conclusive in mine. They are weightier than the prudential reasons you have to let the murderer in. So weak anti-rationalism does significantly better than strong anti-rationalism at accommodating ordinary intuition about the weights of reasons.

So does it accommodate ordinary intuition? Can the conservative Humean aim to accommodate a Dorsey-style weak anti-rationalism? Let’s first go over the main claims of the view:

1) Sufficiency: ‘one always has sufficient rational justification to behave morally’ (9)
2) ‘[Sometimes,] one will have sufficient reason to violate a moral requirement’ (9-10)
3) ‘Hence, in these cases, moral requirements will provide sufficient rational justification, but will fail to be rationally dispositive’ (10)

Given what I’ll call Sufficiency, the best that non-moral considerations can do, in rationally determining what we ought to do, is a tie with the demands of morality. ‘One always has rational permission,’ writes Dorsey (10), ‘though will occasionally lack a rational obligation, to conform to moral requirements.’

Dorsey goes on to argue that the reasons of prudence may, when they are sufficiently strong, render it rationally permissible to fail to conform to the demands of morality. One attraction of such an account is that it can explain how act consequentialists can answer the demandingness objection. If morality demands too much self-sacrifice, then it becomes rationally permissible to act on reasons of prudence, rather than morality. Of course, it’s also rationally permissible to act on reasons of morality in such circumstances, given Sufficiency. This seems to capture something intuitive.

My intuitions lack the precision required to judge where the point is that two great empires of normativity pull in opposite directions such that the net effect is a rational shrug. Nonetheless, for structural reasons, I am also sceptical that this balancing point can also
function as a limiting point, as it would have to for weak anti-rationalism to be true. Let me explain. Everyone would agree it is irrational to choose a lollipop over mass relief from starvation (Dorsey 2012, 18). But how much self-sacrifice do we need to rationally balance with the strength of this moral demand? And then, and this is the key question, what would happen if we continued to add to the ledger on the side of prudence? Past the point of equal weight? Making our reasons to act prudentially stronger and stronger? Would it really still be a tie with morality, as Dorsey’s weak rationalism claims? Would it still be rationally permissible to go with what morality demands of us?

Putting some arbitrary numbers on things may help clarify my criticism. Let’s say we have prudential reasons to P and moral reasons to M, where P and M are mutually exclusive options—we have to do one or the other. We know that, according to weak anti-rationalism, the demands of these two kinds of normativity can be weighed against one another. So we can assign arbitrary values to the rational weight of the prudential and moral reasons to P and M respectively. Let’s call the rational currency the Rational Dollar. It acts as a reserve currency, in that all the other normative currencies are traded in Rational Dollars which determines their relative values. Moral reasons also have their own relative weights, in that there may be moral reasons on both sides of the ledger when we decide what to do. And the same is true for prudential reasons. So let’s say the weight of moral reasons, traded against other moral reasons, is conducted in the moral currency of the Moral Mark. And let’s say the equivalent prudential currency is the Prudential Pound. One Prudential Pound has half the weight of two Prudential Pounds. Ten Moral Marks have twice the rational weight of five Moral Marks.

The question is, how can the exchange rates of Prudential Pounds and Moral Marks into Rational Dollars be, such that Moral Marks always have at least the same value as Prudential Pounds, and sometimes more? In the lollipop case, let’s say the prudential reasons
to eat the lollypop add up to one Prudential Pound. And our moral reasons for the alternative action, preventing starvation, add up to a million Moral Marks. To figure out which has greater rational weight, we have to convert them into Rational Dollars. Mathematically, it doesn’t matter, but keep everything simple, let’s simply stipulate that the conversion rate for all these currencies is one-to-one. It turns out that preventing starvation is worth a million Rational Dollars and eating the lollypop is worth one Rational Dollar. So preventing starvation rather than getting the lollypop is worth a net of nine-hundred-and-ninety-nine Rational Dollars. So we are rationally obligated to do as morality demands. Then we have the tie case that Dorsey uses to solve the demandingness objection, where the value of Prudential Pounds and Moral Marks is the same, in that they each convert into an equal number of Rational Dollars. In this case, the demands of prudence are equally rationally weighty as the demands of morality. So far so good. But here’s where it gets tricky.

What happens when we have the moral equivalent of a lollipop (a very weak moral reason) and the rational equivalent of preventing starvation (a very strong prudential reason)? Let’s say the choice is between one Moral Mark and one thousand Prudential Pounds. On Dorsey’s account, it can’t be that the stipulated exchange rate of 1:1:1 holds, because he claims that what is rationally demanded is always rationally permissible. So now, suddenly, one thousand Prudential Pounds is of equal rational value to one Moral Mark. This is what would need to be the case for it to still be rationally permissible to do as morality demands (which Dorsey’s weak anti-rationalism insists is always the case). It would mean that adding new prudential considerations could have no bearing on what we rationally ought to do, meaning they have precisely no weight.

Let’s take a step back to see what’s gone wrong. Dorsey’s weak rationalism opens the door to weighing two kinds of reasons against one another, and once that door is open, it seems impossible to stop it from allowing for the theoretical possibility of normative domains
other than morality having greater rational weight. So weak anti-rationalism must be more anti-rationalist than Dorsey wants. It can still accept Authority— that there is always a pro tanto normative reason to behave morally—but Sufficiency seems untenable without also accepting Supremacy. It’s a coherent view that non-moral reasons can outweigh what is collectively demanded by our moral reasons (rejecting Supremacy). It’s a coherent view that moral reasons are decisive, or rationally dispositive (accepting Supremacy). But it’s not plausible that non-moral reasons can rationally balance out the demands of moral reasons but precisely that and no more. Sufficiency isn’t a stable position.

So the last option on the table, short of strong moral rationalism, is even weaker weak anti-rationalism: Authority without Supremacy or Sufficiency. Such a view would claim that the demands of morality always provide pro-tanto normative reasons for action. It would not, in other words, claim that the requirements of morality have any particular, unique position in determining what we ought to do all things considered. This discussion is far from conclusive, but it provides some reason to think that weak moral rationalism does not quite to capture what is special about moral normativity. I’ll say more about these features in the next subsection, and why they have led so many philosophers to endorse Supremacy.

2.3. Strong Moral Rationalism

A popular view has it that the greater normative weight of moral reasons holds without exception—that it is always practically rational to do whatever our moral reasons collectively recommend, no matter what our other, non-moral reasons we have to do otherwise. That is to say that the strong interpretation of moral rationalism involves acceptance of
**Supremacy.** If $A$ is morally required to $\phi$ at $t$, $A$ is normatively required to $\phi$ at $t$.\(^{62}\)

Supremacy seems to have been accepted by a number of philosophers. Perhaps most obviously, Kant claims that the practical demands of morality and *only* the practical demands of morality are categorical, rather than merely hypothetical. But Kant is far from alone in endorsing such ideas. Mill writes

[T]he Art of Life, in its three departments, Morality, Prudence or Policy, and Aesthetics; the Right, the Expedient, and the Beautiful or Noble, in human conduct and works. To this art, (which, in the main, is unfortunately still to be created,) all other arts are subordinate; since its principles are those which must determine whether the special aim of any particular art is worthy and desirable, and what is its place in the scale of desirable things.

How ought we to compare the strength of the demands placed on us by these three departments of the ‘Art of Life’? Mill continues:

Without attempting in this place to justify my opinion, or even to define the kind of justification which it admits of, I merely declare my conviction, that the general principle to which all rules of practice ought to conform, and the test by which they should be tried, is that of conduciveness to the happiness of mankind, or rather, of all sentient beings: in other words, that the promotion of happiness is the ultimate principle of Teleology. (Mill 1974, 949-951)

Supremacy has also received more recent philosophical defence. This is Dorsey’s summary:

Kurt Baier writes that: ‘It is our common conviction that moral reasons are superior to all others.’ (Baier 1958, 308) R. Jay Wallace writes that to conform to moral requirements is a ‘fixed constraint on what may be counted as a good life.’ (Wallace 2006, 134) Sarah Stroud writes that ‘[Supremacy] is genuinely part of our pre-theoretical conception of, or at least aspirations for, morality.’ (Stroud 1997, 176) I could go on. (Dorsey 2016, 16)

---

\(^{62}\) David Brink defines this view (1997).
Stroud provides a particularly clear account of the intuitions that lead philosophers to endorse Supremacy, or, in her terms, the ‘overridingness thesis’ (Stroud 1997, 171). Her prima facie case for the thesis includes the following considerations. Firstly, ‘we think of morality as, at least in part, a constraint on the pursuit of our aims […] Morality’s felt role as a constraint on the pursuit of those aims would be undermined if it were OK—not morally OK, but OK from this more encompassing point of view—to ignore moral injunctions in favour of those projects.’ (176) Secondly, ‘if, in fact, morality is not overriding, a commitment to honouring its demands seems rationally unmotivated.’ (176) And thirdly, ‘we are inclined to take the fact that S was morally required to phi to be sufficient overall justification for her phi-ing, even if phi-ing was prudentially wrong, aesthetically unattractive, and so on.’ (177)

Supremacy can also explain which actions are practically rational in all the examples raised in the discussion so far. If moral requirements determine the thing to do, that explains why we oughtn’t to let the axe-murderer into the building or choose a lollipop over preventing mass starvation, and it can even explain why it was, perhaps, rationally permissible for Gaugin to move to Tahiti to pursue his art, provided we think that creating and sharing great works of art has significant moral value. The only cases it seems to struggle with are the demandingness objections to utilitarianism. But these objections seem like insufficient grounds to reject Supremacy (we might just reject utilitarianism).\(^63\)

\(^{63}\)This is speculative, but I suspect many of the demandingness objections to morality can be accommodated with a more expansive account of what has moral significance, as with the Gaugin case. Wolf’s moral saint, for instance, may be failing to realise significant moral goods, such as friendship, connection, and self-development. My reading of Wolf and Williams’
My discussion has been brief, but I hope it suffices to demonstrate that Supremacy, and therefore strong moral rationalism, is at least an attractive position. So I think the conservative Humean should aim to accommodate it. But strong moral rationalism might initially seem particularly difficult for a Humean to accommodate. Indeed, the intuitive asymmetry between moral and other reasons’ weights in rational judgment has, historically, been seen as a problem for conservative Humeanism. Since Humeanism throws all normative reasons into the same desire-grounded pile, it is unclear how any kind of normative reason could systematically rise to the top as strong moral rationalism entails. Particularly if one thinks that the weight of an agent’s reasons is determined by the strength of his desires, the conservative Humean project looks to be in trouble. It is certainly no coincidence that many of the anti-rationalists mentioned in this chapter, such as Williams and the early Foot, were Humeans.

I hold the unconventional view that strong moral rationalism is, in fact, possible for a Humean constitutivist to accommodate. I provide a full account later in the chapter. But for now, here is a brief indication of my argument. The constitutivist claims that the possession of certain desire-like mental states is constitutive of rational agency and that these desire-like states are pro-moral (i.e. the objects of these desire-like states would be promoted by moral behaviour). It follows that all possible agents would necessarily have reasons to act morally. But why would these reasons have systematically greater weight?

Proportionalism is the traditionally dominant Humean account of reasons’ weights—so dominant, in fact, that it avoided being stuck with a label until its first explicit critic felt obliged to provide one (Schroeder 2007, 6.5). Proportionalism claims that the weight of our critiques, is that they both target an overly narrow conception of what morality involves (in Williams’ terms, it leaves out various ethical concerns), rather than its rational significance.
reasons is a function of the strength of our desires (and the degree to which action promotes the objects of our desires). Schroeder argues that a Humean can accept moral rationalism only if he rejects proportionalism. I think he is wrong. I accept proportionalism, but I reject the standard implicit account of strength-of-desire—that it should be quantified in terms of the felt intensity of our desires. I argue instead that the strength of a desire consists in its counterfactual robustness under ideal deliberation; would we rather retain this desire rather than another, even in the face of more information, therapy, greater imaginative exercise etc.? If so, it is a stronger desire, in the relevant sense, than one that we would abandon in such a counterfactual. Since some of our pro-moral desires are constitutive of rational agency, they have perfect robust counterfactual robustness (and therefore systematically more counterfactual robustness than our contingent desires). So our constitutive desires are always stronger than our non-constitutive desires. Therefore, given proportionalism, the (moral) reasons that these desires provide have systematically greater weight than our other reasons. So the constitutivist Humean can accommodate strong moral rationalism.

3. Humeanism and the Weight of Moral Reasons

Now that we have a clearer idea of the intuitive landscape that the conservative Humean ought to accommodate, it is time to examine whether he can pull it off.

3.1. The Special Problem of Moral Reasons’ Weight

I’ll admit that does not look immediately promising for the conservative Humean. Suppose the Common Desires endorsing constitutivist I discussed in the previous chapter is right that every agent has some moral-reason-explaining desire. The view seems to generate both too little and too much weight for reasons.
Take Martha. Martha, has, in addition to some pro-moral desires, some very nasty desires. ‘Sure, I’d mildly like to save the drowning child—it would fulfil a constitutive desire of mine,’ Martha says, ‘but I want to watch him drown much more!’ Intuitively, Martha has a much better reason, a normatively weightier reason, to save the child than to let him drown. We think Martha would be failing by the standards of practical rationality if she were to let the child drown. But the Humean seems forced to say the opposite. After all, Martha desires to watch the child drown more than she desires to save it. The Humean seems to assign either too little weight to the moral reason or too much weight to the non-moral reason.

Before giving to my response, I want to address a potential objection to my description of the example. It might seem Martha has no reason to watch the child drown, not just a weak reason that is easily outweighed. I think we shouldn’t place much confidence in this intuition. The pragmatics of conversation seem to be such that we don’t say people have normative reasons when they have very weak reasons. After all, there’s no value in pointing out when people have reasons which are sure to be outweighed by other reasons to do other things. As Schroeder says, ‘In ordinary contexts we are interested in reasons in order to place weight on them in our deliberations’ (Schroeder 2007b, 122). Indeed, to say they have reasons of such little weight may violate Gricean conversational maxims and make understanding one another difficult or impossible (Schroeder 2007b, 122). Rather, we tend to talk only about fairly weighty reasons. I think these considerations go some way to debunking the intuition that Martha has no reason to watch the child drown, rather than a very weak reason to.\(^\text{64}\) Now let’s return to discussing the weight of Martha’s reasons.

\(^{64}\) There are alternative explanations of what’s going on here that are less friendly to my way of setting things up. For example, Snedegar (2014, 2017) defends a view he calls contrastivism, according to which ‘considerations are fundamentally reasons for things only relative to sets of
3.2. Proportionalism: A Humean Theory of Reasons’ Weights

The Humean needs a theory of normative reasons’ weights. Proportionalism is such a theory. It claims the weightiness of a reason is ‘proportional to how much [the agent] cares about [the object of desire]’ (Schroeder 2007b, 122). The trouble is that Humeanism and proportionalism seem to jointly entail that Martha, who desires to watch the child drown more than she desires to save it, has a weightier reason to watch the child drown than to save him—an unwelcome result. Schroeder’s diagnosis of why Humeanism seems to get the weight of reasons wrong in cases like Martha is that the default Humean theory of reasons’ weight—proportionalism—is faulty (Schroeder 2007, 98).

There’s a second variable which affects reasons’ weight, even according to the Proportionalist—‘how well it is that [phi-ing] promotes [the agent’s] desire’ (98). If you desire red and white wine equally and going to the fridge gives you an 80% chance of finding white wine, but going to the cabinet only gives you a 50% chance of finding red wine, proportionalism says you have a weightier reason to go to the fridge. But this doesn’t seem to help the Humean account for the special weightiness of moral reasons. For it to help, it would need to be true that moral actions always promote agents’ desires much better than amoral/immoral actions. But it’s not at all obvious why this would be the case; Martha is an apparent counter-example, and she is an even clearer counter-example if we stipulate that alternatives’ (2014, 39). In other words, reasons are always reasons to do one thing rather than another thing. So, on Snedegar’s view, it is not, as I say, that you have very weak reason to phi, but that you have no reason to phi relative to the alternative you have. Snedegar’s view effectively claims there are no such things as pro tanto reasons or reasons simpliciter, since all reasons are contrastive.
Martha would *definitely* achieve the object of the relevant desire if she performed either action.

In the face of these challenges, Schroeder rejects proportionalism in favour of his own theory of reasons’ weight. But there are good reasons for the Humean to want to retain some version of the view. It’s strange for a Humean to think an agent’s reasons are explained by her desires but that the weights of an agent’s reasons are not (Sobel 2009b, Rieder 2016). Given that proportionalism seems well motivated from the Humean position, can it be salvaged? I think so. But we would need to explain why our pro-moral desires (our desires that provide reasons to act morally) are systematically *stronger* than our other desires.

3.3. Theories of Strength-Of-Desire

So what theories of strength-of-desire are there, and do any of them accommodate the thesis that pro-moral desires are stronger than other desires?

Perhaps the most obvious sense in which desires can be stronger and weaker is how intensely we feel them. Of course, many of us intensely feel a desire for objects that would be promoted by moral action. But it’s implausible that pro-moral desires are systematically felt more intensely than other desires in all agents. This is what would need to be the case for the Proportionalist to accommodate the Moral Rationalist claim that moral reasons are weightier than other kinds of reasons.

But perhaps we’re stacking the deck against Humeanism by thinking of the strength of desire in terms of felt intensity (Evers 2013, 316). Is what we desire *most* really what we feel most intensely desirous toward? Not clearly. I’d be inclined to say that I desire a good relationship with my wife *much* more than I desire to eat the last slice of her birthday cake

---

65 This is Schroeder’s ‘Recursive View’ (2007, 5.2).
even if, as I walk into the kitchen, I *feel* my desire for the cake more intensely. If this is right, then *felt intensity* cannot fully determine the strength of desire.

Let’s take another character—Augustus. Augustus shares my intensely felt desire for chocolate cake, but he has another (less intensely felt) desire that would be promoted by refraining from eating the cake (he’s committed to a healthy diet). Intuitively, his pro-cake desire isn’t as reason-providing as his anti-cake desire (or, indeed, to his commitment to a diet). That is, Augustus has an intuitively weightier reason to refrain from eating the chocolate cake. For proportionalism to get the intuitive answer, it needs an analysis of strength-of-desire according to which Augustus’s anti-cake desire is stronger than his pro-cake desire. Felt intensity can’t provide this.

Sobel suggests (2009b) and Evers develops (2013) an account of strength-of-desire Evers calls preferentialism:

**Preferentialism.** There is more (i.e. stronger) reason to X than there is to Y for an agent A just in case A prefers X-ing over Y-ing. (Evers 2013, 317)

A few things about this formulation of Preferentialism are worth noting. First, even though Evers calls this a theory of strength of desire, it is, as stated here, strictly speaking, a theory of *reasons’* weights. But this distinction collapses under the assumption of proportionalism. Preferentialism gives a theory of our desires’ strengths, and then proportionalism tells us that our reasons’ weights are a function of our desires’ strengths (and the extent to which actions promote the objects of our desires). Second, Evers takes preferentialism to define the weight of ‘subjective’ reasons—what reasons it is rational for the agent to take himself to have. The weight of the agent’s *objective* reasons, the normative reasons the agent has, my focus, ‘corresponds to the preference ordering s/he would have if her beliefs […] were true.’ (Evers 2013, 317) So here is the relevant *objective* form of Preferentialism:
**Preferentialism: Objective Reasons (henceforth: preferentialism).** There is more (i.e. stronger) objective reason to X than there is to Y for an agent A just in case A would prefer X-ing over Y-ing if she were fully informed about the relevant facts (the strength of the promotion relation being equal).

Third, an agent’s preference for X-ing or Y-ing,\(^{66}\) Evers points out, can be influenced by the degree to which the agent believes that (for a subjective reason) and the degree to which it is actually true (for an objective reason) that X-ing or Y-ing will promote the objects of the agent’s desires. So, how does Preferentialism perform?

Preferentialism has some appeal as a theory of strength-of-desire. It keeps the Humean-friendly link between the weight of agents’ reasons and features of their psychologies but it also accommodates Augustus, since he plausibly prefers refraining from eating the cake to eating it (despite feeling his desire to eat it more intensely than his desire to refrain).\(^{67}\) So preferentialism gets the right verdict that Augustus has a stronger desire to diet than to indulge, and therefore a weightier reason to refrain from eating the cake.

But now think back to immoral Martha. Let’s stipulate that Martha, as well as feeling her desire to watch the child drown *more intensely* than she feels her desire to save the child, also prefers watching the child drown to saving it (even if she has true beliefs). Then preferentialism entails Martha has a *weightier* reason to watch the child drown. But this

\(^{66}\) I am happy with Evers’ gloss on what a preference consists in: ‘what it is to prefer something is to be disposed to choose it or promote it (or something along these lines)’ (Evers 2013, 317).

\(^{67}\) Evers writes ‘This idea (let’s call it preferentialism) is a natural partner of a Humean theory of reasons (much more natural than Schroeder’s obscure alternative […]). For preferentialism makes the weight of reasons, like reasons themselves, depend on psychological properties of the agent.’ Evers (2013, 317)
seems false. Moral reasons prove the difficult case for the Humean, once again. To accommodate intuition, the preferentialist would need to explain why agents all prefer the moral over the immoral or the amoral.

Notice that to get an account of the objective weights of an agent’s normative reasons, we have already had to idealise—to specify that the preferences that matter are those the agent would have if he had true beliefs. The first step to my solution to getting the right weight for moral reasons is to idealise preferentialism a bit more.

Notice that idealising preferentialism is not idealising Humeanism; I am not claiming that what reasons we have are determined by what we would ideally, counterfactually, desire. I am claiming the weight of our reasons is determined by our ideal preference rankings given our actual desires. Here is the idealised form of preferentialism:

**Ideal Preferentialism.** There is more (i.e. stronger) reason to X than there is to Y for an agent A just in case A prefers X-ing over Y-ing *in condition C*.

Condition C could be (and has been, in the context of other idealised desire-based accounts) specified a number of ways. Most idealisations specify that the agent is fully informed about relevant information. Beyond that, our ideal state might require, as Brandt suggested, after the agent undergoes an exhaustive process of ‘cognitive psychotherapy’. Or, as Smith suggests, if he were acting on the advice of an ideal version of himself acting as an advisor (Smith 1994). Or, as Williams suggests, it could be after full exercise of the imagination. I would like to stay as ecumenical as possible, so I’ll simply refer to C as ‘under ideal circumstances’ and use Brandt’s suggestion as a placeholder where more specification is helpful.

We’ve seen how non-ideal preferentialism accommodates Augustus because he has a preference for sticking to his diet over eating cake. This means his desire to stick to the diet (or realise the object of the diet, e.g. health) is stronger than his desire to eat the cake. And
this means Augustus has a weightier reason to refrain. Since this preference ranking would presumably survive ideal circumstances (e.g. cognitive psychotherapy and full information), ideal preferentialism yields the same intuitive conclusions about the relative weights of Augustus’s reasons.

Let’s now introduce a modified Augustus: Augustus*. Augustus* is like Augustus, except as well as desiring to eat cake more intensely than he desires to refrain, he also prefers to eat cake than to refrain and stick to his diet. Non-ideal preferentialism now entails Augustus* desires to eat the cake more strongly than he desires to maintain a healthy diet, and therefore he has a weightier normative reason to eat the cake, which seems like the wrong result. But ideal preferentialism can explain why he has a weightier reason to refrain from eating the cake. In ideal circumstances (e.g. fully informed about the health effects, effects on mood, having undergone cognitive psychotherapy etc.) Augustus’s* preference ranking would diverge from his actual preference ranking. Plausibly, I think, Augustus* would, in ideal circumstances, re-order his preferences such that refraining would trump eating.➊ If this is right, then Augustus’s* desire to refrain is actually a stronger desire (in the relevant sense) than his desire to eat, and, given proportionalism, his reason to refrain is, therefore, also weightier.

As usual, immoral Martha is a harder case. There is a version of Martha, Redeemable Martha, who would re-order her preferences and prefer saving the child to watching it drown

➊ Notice that I am not appealing to a notion of ideal rational deliberation that is substantive. I do not claim Augustus* will prefer refraining because eating healthily is objectively better, or anything similar. I’m assuming that even on a purely procedural notion of what ideal rational deliberation would result in refraining ranking higher in Augustus’s* preferences, given his other preferences presumably include being healthy, happy, and living a long time.
in the specified ideal circumstances. Then we get the intuitive results the same way we did with Augustus*. But we can also conceive of an Irredeemable Martha whose desires and preferences are very nasty but also very stable and coherent in such a way that ideal rational deliberation, or psychotherapy, would leave them unscathed. Ideal preferentialism would seem to entail that Irredeemable Martha has weightier reasons to watch the child drown—the wrong result. We still need to explain why agents’ ideal preference rankings necessarily rank moral actions over amoral and immoral actions.

3.4. Constitutive Commitments and Ideal Preference Rankings

What matters, in determining the weight of an agent’s reasons, is not what desires they feel more intensely, nor what courses of action or commitments they are actually disposed to choose over others. Rather, according to ideal preferentialism, what determines an agent’s reasons’ weights is the preference ranking for actions he would have in C. So let’s think more about what agents’ preference rankings would look like assuming constitutivism is true. Why would agents’ ideal preference rankings reliably exhibit a preference for moral action?

According to the commitment constitutivism I’ve been outlining, here are certain desire-like mental states, practical commitments, the possession of which is partly constitutive of agency. If so, then we have these commitments inescapably, insofar as we have reasons at all. And this is relevant to the structure of our ideal preference rankings. In an ideal process of rational deliberation, or an ideal process of cognitive psychotherapy, one where we resolve practical conflicts between our various ends, desires, projects, commitments etc., we are rationally compelled to order our preferences in a manner that favours the reasons we have given our constitutive desires/commitments over those we have given our contingent desires/commitments.
As Ridge (2017) argues, the fact that our commitment to some ends is *constitutive of our agency* could be highly relevant in determining *which incompatible end* we ought to abandon:

Plausibly, when two ends conflict we are rationally required to abandon at least one of those ends.

However, it is also plausible that the rational ‘ought’ implies ‘can’, and that this constraint on our deliberations is one that quite correctly plays a role in our practical deliberation. In that case, though, the fact that an end is inescapable can and should appear from the practical point of view. (2017, 2961)

In other words, whenever there is a conflict between two ends such that rationality requires we abandon one and one end can be dropped and the other cannot, then we must abandon the abandonable end. So the ends set by constitutive commitments will always come out ahead when they conflict with other ends. This is to say that rationally, an agent will always *prefer* those actions which promote constitutive ends whenever they are in conflict with actions which promote other ends, and hence they desire the objects of constitutive desires more strongly than they desire non-constitutive desires.

Time for step three: feeding Ridge’s point into ideal preferentialism. We are rationally required to abandon a *contingent* end whenever it conflicts with a *constitutive* end. Imagine two ends to which an agent is practically committed, Contingent and Constitutive, are in conflict. Contingent and Constitutive are mutually incompatible states of affairs, or the necessary means to obtain one would preclude the necessary means to obtain the other. In either case, an agent can either realise Contingent or Constitutive. And if any rational agent must abandon Contingent whenever it conflicts with Constitutive that means (analytically) that any rational agent A *prefers* Constitutive to Contingent. If any rational agent *prefers* Constitutive to Contingent, then he will also, as a matter of practical rationality, prefer whatever actions promote constitutive ends to those which promote contingent ends. So,
according to ideal preferentialism, actual agents always have weightier reasons to perform those actions which promote the agent’s constitutive commitments, since the weight of their reasons is determined by their ideal preference rankings.

Recall also that the practical commitments which are partly constitutive of agency are, the constitutivist claims, pro-moral—they provide agents with normative reasons to act morally. If the constitutivist is right about this, then our moral reasons will be weightier than the reasons we have given our contingent desires, just as strong moral rationalism claims.

What emerges from these points is the following picture of our reasons’ weights. An agent’s ideal preference ranking determines the strengths of the actual agent’s desires and, therefore, the weights of his reasons given his desires. Ideal preference rankings will have two tiers—a constitutive tier and a non-constitutive tier. These two tiers exist because of Ridge’s point. Whenever constitutive ends conflict with contingent ends, a rational deliberator is forced to retain the constitutive end. In other words, the rational deliberator systematically prefers constitutive commitments to non-constitutive commitments. Ideal preferentialism therefore entails that there is, systematically, more/weightier reasons for every agent to phi whenever phi-ing is the necessary means to realising a constitutive end.

Now we can (finally) accommodate irredeemably immoral Martha. Here’s how from the top. Actual Martha prefers to watch the child drown. But she is an agent, so, (given commitment-based constitutivism) she has some kind of pro-moral practical commitment that, let’s say, provides her with a reason to save the child. To figure out whether she has more reason to act morally or to watch the child drown, we need to know which of her reasons has greater weight for her. Ideal preferentialism and proportionalism jointly entail that Martha’s reason to save the child is weightier than her reason to watch the child drown as she would prefer saving the child in condition C (e.g. after some ideal rational deliberation, perhaps after undergoing cognitive psychotherapy, when fully informed, etc.). Given that
watching the child drown and saving the child are conflicting ends which demand incompatible actions, she must abandon one of them. But which to abandon? There is an important asymmetry here. She cannot abandon the pro-moral end set by her constitutive commitments but she can abandon the end set by her contingent desire to see the child drown. So she is rationally required to abandon the latter. Martha’s rational counterpart, Martha in C, would, therefore, prefer to save the child than to watch it drown. This means that Martha’s reason for saving the child has greater weight, other things being equal, than her reason for watching it drown. This is the right result.

4. The Nature of Constitutive Commitments

I’ve been vague about the nature of constitutive commitments. It might even be objected that my proposed constitutivist solution to reconciling moral rationalism and Humeanism is playing on a problematic ambiguity about constitutive commitments. Are they some kind of desire-like mental state? Or are they something more like intentions? Given that I’m defending a broadly Humean account of normative reasons, the obvious kind of mental state for a constitutive commitment to be is desire, and I’ve previously labelled constitutive commitments desire-like mental states. But one of the steps in my reconciliation of Humeanism and moral rationalism makes use of Ridge’s argument, which explicitly deals with the rational implications of an agent having constitutive ends, which are generally taken to be the objects of intentions or plans rather than the objects of desires. In this section, I’ll briefly explain why I think constitutive commitments are plausibly understood as ‘self-governing policies’ in Bratman’s terms, i.e. general plan-like commitments we can undertake to give weight to certain considerations in practical reasoning.

Ridge’s argument relies on the premise that it is practically irrational to simultaneously possess two incompatible ends and that a rational agent would in such
circumstances, abandon one of their ends. If constitutive commitments are just simple desires, then there is no apparent irrationality for having ‘incompatible’ desires.

To see more clearly why constitutive commitments being simple desires would be a problem, let’s attempt to run Ridge’s argument with desires substituted for commitments to ends. We won’t get too far. An agent has two ‘conflicting’ desires. The desires conflict in the sense that promoting the object of one desire would make it impossible to promote the fulfilment of the other desire. A conflicting desire-pair like this might be something like (A) the desire to have a productive morning and (B) the desire to sleep in. Let’s stipulate that A is a constitutive desire, whereas B is not. Desire A gives the agent reasons to set an alarm, get up on time, and so on. Desire B gives her reasons to refrain from getting up, opening the curtains etc. But all these reasons are merely *pro tanto*. The presence of these desires doesn’t seem to settle what she must do, rationally speaking. As such, there’s no reason to think that there is rational pressure on the agent to abandon her non-constitutive desire. She can, rationally, go on desiring both these incompatible states of affairs. But, since she is not forced to abandon the non-constitutive desire, we don’t have a reason for thinking the reasons that follow from constitutive desires take precedence over those that follow from non-constitutive desires. And this is what we need for my proposed reconciliation of Humeanism and moral rationalism to go through.

If these desires were instead *intentions*, or ‘commitments to ends’ as Ridge puts it, things would be different. Having an intention doesn’t just generate *pro tanto* reasons; it seems to rationally settle the thing to do. Specifically, rationality demands (on the wide-scope interpretation) that you must either take the necessary means to achieve your intended end *or* you must abandon the end. Ridge’s point is that if one of your ends that is in conflict with another is constitutive of agency then you *can’t* rationally abandon it. And, since there’s rational pressure on you to abandon one of the conflicting ends, that means you ought to
abandon the non-constitutive end. This would allow for the conclusion that our practical reasons which follow from our constitutive intentions and commitments to ends take precedence over those that follow from our non-constitutive

So it is clear that Ridge is thinking of these constitutive commitments as more like intentions than desire-like mental states. At least, they must be mental states that are rationally subject to consistency requirements. These requirements generate the rational pressure to abandon one of the ends—as Ridge notes, the contingently adopted end, rather than the constitutively adopted end.

But the Humean needn’t restrict himself to simple desires in order to ground normative reasons. Indeed, most Humean theories of reasons appeal to desire in a technical, broad sense as any conative mental state—any mental state with a mind-to-world ‘direction of fit’. This could include a wide range of mental states, e.g. valuings, plans, commitments, projects, and intentions. Arguably all of these are rationally subject to consistency requirements. If I plan to walk the dog at 9 am, and I also plan to start work at 9 am, then my plans are inconsistent and there is rational pressure on me to abandon or modify at least one plan.

So which of these desire-like mental states with consistency requirements do I think constitutive commitments are? For reasons I will explain in the final substantive chapter of this thesis, I think they are best understood as what Bratman calls self-governing policies, i.e. general plan-like commitments to give weight to certain things in practical deliberation. For now, what matters is that a self-governing policy is a mental state with the right kind of structure to resolve the worry about desire-like states not generating any rational pressure to abandon contingent ends/goals when they are inconsistent with constitutive ones; they are rationally subject to consistency requirements. Specifically, our self-governing policies must be consistent with our other plans and intentions. As Bratman writes:
The norms of rationality at issue [with self-governing policies] are “wide-scope” norms on sets of attitudes: they are, roughly, norms that enjoin or reject certain combinations of attitudes [...] these include norms of practical rationality that apply to intentions. In particular, there are norms of consistency and of means-end coherence of intentions and plans. There is a rational demand that one’s intentions, taken together with one’s beliefs, fit together into a consistent model. And there is a rational demand that one’s intentions be means end coherent in the sense, roughly, that it not be true that one intends E, believes that E requires that one now intends necessary means M, and yet not now intends M. In both respects intentions are subject to rationality norms that do not apply in the same way to ordinary desires. (Bratman 2009, 412)

And, since they are general plans, plans that apply all the time (rather than being activated in particular contexts and times), they are always ‘on duty’. Hence, whenever we act intentionally, we are rationally obliged to ensure the intention guiding our action is consistent with our general plan (our constitutive commitment). And whenever they come into conflict, the wide-scope norms of consistency demand we abandon the non-constitutive intention. This is why a rational agent will always prefer actions recommended by our constitutive commitments to those recommended by contingent mental states. I hope this brief discussion is sufficient to satisfy the reader that the objection at hand can be met if we really are committed to general plans in this sense, as a constitutive matter.

5. Conclusion and Stock-Taking

5.1. Conclusion

Conservative Humeanism needs to account for moral rationalism, including the intuitions we have about reasons’ weights and moral reasons’ particular weightiness. I suggested the ‘gold standard’ for the Humean would be to accommodate strong moral rationalism. In the previous chapter, I argued that constitutivism (of a Humean flavour) has
the resources to explain how the Humean can accommodate the intuitive extension of moral reasons (their universality, in particular) and, perhaps more interestingly, their modal status.

In this chapter, I’ve argued that constitutivism can also provide the Humean with an intuitive account of moral reasons’ special weight—why moral requirements seem to settle what we ought to do, rationally speaking.

Let me briefly summarise the four main pieces of the puzzle. These are:

1. Humeanism
2. Proportionalism
3. Ideal Preferentialism
4. Commitment constitutivism

We begin with (1) Humeanism, which claims roughly that an agent’s reasons are explained by his desire-like mental states. We then (2) agree with the proportionalist that reasons’ weights are (partly) a function of the strength of the desires which explain them. Finally, (3) we give an account of strength-of-desire in terms of an agent’s (ideal) preference ranking. Ideal Preferentialism claims a desire D1 is stronger than D2 just in case an agent, in ideal condition C, would have a disposition to perform an action which promotes D1 rather than one that promotes D2, assuming D1 and D2 promote the objects of their respective desires equally strongly. These ideal conditions C may include being fully informed about relevant facts or having undergone a comprehensive process of cognitive psychotherapy, etc. I have stayed uncommitted on these specifics. The thought here is under ideal deliberative conditions, whatever they might be, the agent will be confronted with a choice. Say X-ing is the necessary means to, or even merely promotes, a constitutive end and Y-ing is the necessary means to, or even merely promotes, a contingent end. What does the agent prefer? X-ing or Y-ing? The agent will, given enough hours of psychotherapy or the like, notice that he can abandon the contingent end, but he cannot rationally abandon the constitutive end. So
he will realise that he is rationally required to retain the constitutive end and abandon the contingence one. This means he prefers the constitutive desire over the contingent desire in condition C, entailing his constitutive desire provides weightier reasons than his contingent desire. Finally, (4) we draw on commitment constitutivism, the Humean-friendly version of constitutivism which specifies that the possession of a certain practical commitment (or commitments) is partly constitutive of agency and this constitutive commitment is pro-moral in the sense that it provides reasons to act morally. (1-4) jointly entail that agents’ pro-moral reasons have systematically greater weight than their reasons to act amorally or immorally. This accommodates strong moral rationalism.
Before moving on to the next chapter, let’s briefly take stock. In the previous three chapters, I have played the role of philosophical matchmaker, seeking to couple two views: (1) metaethical Humeanism and (2) a modest form of constitutivism I have referred to as commitment constitutivism (or, more broadly, Constitutive Common Desires). Humeanism claims our normative practical reasons are explained (in part) by our pro-attitudes. Commitment constitutivism claims that the possession of certain pro-attitudes, which I have referred to as practical commitments, is partly constitutive rational agency.

The wedding of these two views would, I argued, make both partners happy. Humeanism provides the necessary metaethical home for constitutivism. This enables the constitutivist to sidestep the influential shmagency objection faced by more ambitious freestanding interpretations of constitutivism. Constitutivism (of the right kind) helps the Humean deal with problems with the extension of ‘reason’, particularly with the amoralism objection. Humeanism is commonly thought to be incapable of explaining the normative reasons we all seem to have to do what morality demands of us. I argued in Chapter Two that if Constitutive Common Desires is true then there are some pro-attitudes that are shared by all rational agents. This would allow the Humean to accommodate universal moral reasons with plausibly robust modality. They would be reasons that are explained by our constitutive pro-attitudes. In Chapter Three, I argued that Humean constitutivism can also account for the intuitive weight of moral reasons.

My account of Humean constitutivism leaves two major gaps. The first gap, explored the next chapter, is the derivation of moral norms from agents’ constitutive commitments. I fill this gap by appealing to something like Darwall’s second-person constitutivism.

According to Darwall’s arguments, all those agents who are second-personally competent
must, as a constitutive matter, have commitments to positions which rationally entail the adoption of a distinctively Kantian attitude of respect for other people. The content of these commitments provides a convincing basis for providing a vindication of recognisable morality (or at least much of it—most plausibly the Scanlonian chunk of morality) the morality of ‘what we owe to one another’.

The second major gap is an ontological one: the existence of these constitutive commitments. For all I say in other chapters, my conclusions are mostly hypothetical. They are of the form: if there are constitutive commitments and they are the way I describe, then they could solve a variety of metaethical problems, both for the Humean and for all those concerned with the normative authority of morality. This is the main topic of the fifth and final substantive chapter.
Chapter Four: Second-Person Constitutivism and the Optionality Challenge

[A] human being regarded as a person, that is, as the subject of a morally practical reason, […] possesses a dignity by which he exacts respect for himself from all other rational beings in the world. He can measure himself with every other beings of this kind and value himself on a footing of equality with them.


1. Introduction

So far, this thesis has argued that if there are pro-moral constitutive commitments, then Humean constitutivism would be an attractive theoretical package. But the account is incomplete. It has yet to provide three key things: (1) an account of the moral reasons/norms which follow from constitutive commitments, (2) an account of the objects of those constitutive commitments and (3) an argument that such constitutive commitments actually exist. In this chapter, I fill the first of those gaps by piggybacking on Darwall’s work, particularly his derivation of Scanlonian moral norms from the constitutive features of the second-person standpoint. In the next chapter, I argue that rational agency necessarily involves, as a constitutive matter, occupying the second-person standpoint and undertaking second-personal commitments, and hence the normativity of central moral norms and reasons follow from the constitutive features of agency.

So what is Darwall’s view? In his influential book The Second Person Standpoint (2006), he argues that insofar as we occupy the second-person standpoint—the standpoint of a person who is engaged with others in an interpersonal way partly characterised by the Strawsonian (1962) ‘reactive attitudes’ of blame, praise, etc.—we necessarily presuppose certain attitudes of mutual accountability and respect between persons. These
presuppositions, or ‘normative felicity conditions’, are partly constitutive of the second-person standpoint. These presuppositions are also, Darwall argues, the necessary raw materials to rationally derive a contractualist account of our moral obligations (at least the substantial portion of morality that concerns ‘what we owe one another’ in Scanlon’s terms). Hence, in Darwall’s view, the normativity of (at least some of) our moral reasons follows from our occupying the second-person standpoint. I rehearse these arguments in section 2.

Darwall argues that the second-person standpoint provides ‘a foundation for contractualism’ (Darwall 2006, 35-38). But even foundations need to lie on solid ground. Much of this chapter examines that ground. Or, to put it another way, I examine the foundations of the second-person standpoint itself. The problem second-person constitutivism faces can be drawn out with a (now familiar) question, the topic of section 3; how can the view accommodate moral rationalism—the view that if an agent has a moral reason to phi she has a normative reason to phi (with some weight)? The specific challenge second-person constitutivism faces I call the optionality objection. This objection states that it is possible to ‘opt out’ from the second-person standpoint and, therefore, that it is also possible to ‘opt out’ from moral normativity (even assuming Darwall’s derivation of contractualist moral norms succeeds). The existence of this possibility would violate moral rationalism. I shall argue that neither Darwall’s (section 4) nor Korsgaard’s (section 5) proposed solutions to the optionality objection are adequate, but that a third solution is possible. I briefly outline my proposed solution at the end of this chapter and argue for it in the next.

2. The Promise of Second-Person Constitutivism

Constitutivism can come in either a first- or second-personal form. The former is more common. According to first-personal constitutivist arguments, I must accept moral norms insofar as I conceive myself as a full-blooded agent acting for reasons (see e.g.
Korsgaard 1996). The reality or ‘robustness’ of moral normativity is something the individual deliberator discovers for themselves as they perform actions. Such arguments are also generally instances of what I’ve called ‘agential constitutivism’: versions of constitutivism which claim moral norms follow from what is constitutive of agency or action.

Second-personal constitutivists argue that moral normativity is grounded in the constitutive features of interpersonal interaction of some kind. Hence, second-person constitutivism is not (at least generally) agential—it does not claim that moral norms are partly constitutive of agency.69

Darwall’s *The Second Person Standpoint* (2006) is likely the clearest and best-developed second-personal constitutivist account, though he does not describe his view in those terms.70 For that reason, it will be my focus. Roughly, Darwall extracts a Scanlonian contractualist account of moral norms, reasons, and duties based on mutual respect of persons. He extracts all this from the necessary presuppositions we make in occupying the second-person standpoint. In this section, I’ll provide a brief summary of Darwall’s derivation of contractualist norms.71

---

69 In the following chapter, I’ll argue that second-person constitutivism can be considered an agential form of constitutivism because rational agency requires that we think second-personally.

70 Other views that might fall into the ‘second-person constitutivist’ camp include Habermas (1990), Apel (1980), O’Neill (1986), and perhaps certain passages from Korsgaard (1996, 2009). I’ll discuss most of these views in more detail in the next chapter.

71 I’ll largely skip over one of Darwall’s major concerns, namely the particular conception of morality (morality as equal accountability (MEA)) or, more precisely, equal accountability
The second-person standpoint is the standpoint we occupy whenever we make or acknowledge second-personal claims (Darwall 2006, 3). Second-personal claims can generally be understood as claims on each other’s wills, and presuppose second-personal authority which makes the claims valid. In each case in which the second-person stance appears, there is an ‘I-you-me structure of reciprocal address’ (Darwall 2006, 3). In some cases the presence of the second-personal mode of address is clear and explicit, perhaps even linguistically explicit. For example, I might verbally order. ‘I order you to do the laundry!’

In Darwall’s view, second-personal claims purport to provide second-personal reasons. In telling you to do the laundry, I am purporting to give you a reason (a second-personal one) to do so. I am not, in other words, merely trying to make you aware of a first-personal reason you already have. Hence, in occupying the second-personal stance, we must acknowledge one another as sources of reasons.

Darwall illustrates this distinction between first and second-personal reasons and claims with a discussion of two ways in which you might try to get someone to ‘stop causing you pain, say, to remove his foot from on top of yours.’ (2006, 5) Firstly, you could try to ‘get him to feel sympathetic concern […] thereby leading him to want you to be free of pain.’ (5) The story, in this case, is that you’re trying to instil in him the desire to free you from pain, or a belief that doing so is good, and thereby give him a reason to lift his foot:

Qua this form of reason-giving, you would be asking him to agree, as it were, that there is a reason for him to do something rather than asking him to agree to do it. Any claims you might make would thus be on his beliefs about practical reasons and not directly on his will. Second,
your being able to give him the reason would not depend in any way on his seeing you as trying to give it to him or as having any competence or authority to do so. (6-7)

But there is a second, way you might try to give the foot-stepper a reason to remove his foot:

Alternatively, you might lay a claim or address a purportedly valid demand. You might say something that asserts or implies your authority to claim or demand that he move his foot and that simultaneously expresses this demand. You might demand this as the person whose foot he is stepping on, or as a member of the moral community, whose members understand themselves as demanding that people not step on another’s feet, or both. (7)

In other words, we seem to presuppose certain interpersonal authority relations when we make these kinds of second-personal claims on one another. And whether one succeeds in providing a second-personal reason as one makes a second-personal claim depends on whether certain ‘presuppositions’, or ‘normative felicity conditions’, are met:

[A]n addressee attempts to give an addressee a reason for acting that is grounded in normative relations that, as I argue, he presupposes the addressee can be expected to accept. What is built into this presupposition? Quite a lot. (Darwall 2006, 20)

The first thing we necessarily presuppose in second-personal address is that both the addressee and the addressee possess:

[A] common competence, authority, and, therefore, responsibility as free and rational, a mutual second-personality that addressee [and addressee] share and that is appropriately recognized reciprocally. (Darwall 2006, 21)

For an addressee to be capable of accepting the second-personal claim as reason-giving, she must possess ‘second-personal competence’, i.e. the ability to recognise second-personal claims and to act upon them (i.e. to conform one’s will to them). Unlike acting on a first-personal reason, being intimidated or coerced into an action could not be considered acting on a second-personal reason. Insofar as coercion can be understood as reason-
providing, it can only be understood as providing agents with first-personal reasons to comply (e.g., out of a desire to protect one’s life). The act of addressing another second-personally, however, presupposes that they could freely and autonomously choose to act on the reason one is attempting to provide them. It also presupposes that they are accountable to you (and to themselves) for their success or failure in acting according to the second-personal reason you purport to provide.

What emerges from these various assumptions which seem built-in to second-personal address is a picture of participants in a standing of reciprocal authority toward one another and possessing equal second-personal competence. Since the very authority which the addresser purports to possess in making a claim on another (second-personal authority) is the very same authority as that which he presupposes the addressee possesses as a second-personally competent agent, the second-personal mode of address also entails a fundamental equality which underlies even seemingly hierarchical, authority-based relations. Suppose, to take a more challenging case, you are a ranking officer and order a lower-ranked soldier to perform some action. Surely, in this apparently second-personal case, you do not accord equal respect and autonomy? Indeed, you are claiming the authority to determine the soldier’s

---

72 In some sense, you are still ‘free’ if you’re freely acting out of coercion. I take it that claims made in the second-person standpoint presuppose freedom in the richer sense that our Strawsonian reactive attitudes toward an agent who we take to be capable of responding to second-personal claims will be those we adopt toward autonomous agents, rather than coerced ones. We will think of agents we try to provide with second-personal reasons the appropriate objects of praise/blame, for example. In contrast, we might not think that agents we are trying to give first-personal reasons are praiseworthy and blameworthy for conforming to the reasons we are trying to provide them.
will. But even in difficult cases like this, it remains plausible that second-personal address involves the proto-moral presuppositions Darwall points out. In giving the order, you assume that the soldier has the second-personal competence to acknowledge your second-personal authority and to act in accordance with your legitimate claim on my will. To recognise him as having this autonomy and competence is to recognise him as an equal in the sense of possessing ‘equal authority to make claims of one another as free and rational’ and therefore to have an attitude of respect toward him (Darwall 2006, 246). Throughout even this hierarchical, authority-based interaction, we therefore presuppose a relation of mutual recognition, ‘There is thus, I argue, a form of reciprocal respect that is built into all second-personal reason-giving (Darwall 2006, 21)

If all this is correct, we must take ourselves, insofar as we occupy the second-personal standpoint, as having equal normative authority (and therefore as being sources of normative claims on each other) and being autonomous agents.

Darwall states all these claims slightly more formally, as follows:

1. Addressing second-personal reasons always presupposes that addressees can freely and rationally determine themselves by the addressed reasons. (269)

2. Addressing second-personal demands or claims always presupposes that the addressee is responsible for compliance. (269)

3. Addressing second-personal reasons always presupposes that addresses can rationally accept these reasons (and the authority relations in which they are grounded). Thus, although second-personal reasons are frequently addressed to someone standing in some more specific normative relation, the addressee is more properly conceived of as a person-who-happens-to-stand-in-that-normative-relation. Second-personal justification is always ultimately justification to another as free and rational. (269-71)

4. Addressing a demand always presupposes a distinction between legitimate (second-personally justified) ways of relating to someone directly (“demandingly”) that respects I’m
as a free and rational person, on the one hand, and illegitimately coercing him, on the other. It follows that one demand that anyone has the authority to make is that he not be subject to demanding (coercive) conduct that cannot be justified by second-personal reasons. (271-4)

5. Addressing second-personal demands always presupposes, therefore, that the addressee has a second-personal authority as free and rational and, consequently, that the addressee share a common authority to make claims on one another. (274-5)

6. Addressing second-personal reasons always, therefore, presupposes a common basic dignity of persons. (275)

7. Addressing second-personal reasons always presupposes autonomy of the will. (275-6)

Darwall then goes on to argue (2006, ch.12) that a contractualist set of moral norms and obligations (morality as equal accountability) similar to Scanlon’s contractualism follows from the normative felicity conditions of the second-personal standpoint.

The derivation of some kind of contractualist norms from the second-person standpoint runs as follows. In the necessary reciprocal recognition of the second-person standpoint, persons are implicitly:

[C]ommitted alike to their mutual accountability as equal free and rational persons, and this commits them […] to a constraint of reasonableness of any demands they address: they must be able to expect their addresses to accept, or not reasonable reject, their demands as free and rational persons. (Darwall 2006, 320)

Hence, the second-person standpoint vindicates contractualism’s central notion of reasonableness and its idea of a ‘community of mutually accountable free and rational

73 Darwall himself does not commit to a particular substantive contractualist theory. He merely claims that the ‘core’ of all contractualist theories, the idea of moral norms as governed by reasonableness and the centrality of a moral community of equals, can be derived from the second-person standpoint.
persons’ (Darwall 2006, 320). It thereby makes sense why moral norms would rule out acts (i.e. give them the status of ‘morally wrong’) which, in Scanlon’s formulation, would be disallowed by any set of principles that no one could reasonably reject (Scanlon 1998, 153). Hence, the content of these contractualist principles defines our moral obligations to one another as equal and free persons and, therefore, as occupiers of the second-person standpoint.

I’m going to assume that this argument, or one like it, can be made to work.74 If it can, we would be entitled to the conclusion that occupants of the second-person standpoint are committed to the substantive contents of contractualist moral principles and norms. But before we even think about enlisting this argument for my own purposes, a particular challenge must be answered. The remainder of this chapter will deal with this challenge and the solutions which have been proposed.

3. A Problem: The Optionality of the Second-Person Standpoint

I have argued that constitutivism needs to accommodate:

Moral Rationalism. Necessarily, if there is a moral reason for A to phi, then there is a normative reason for A to phi and, at least sometimes, normative reasons provided by moral reasons have significant weight. (Lord and Plunkett 2017, 330)

The constitutivist strategy for accommodating moral rationalism is something like the following. The first step is to give a description of the constitutive features/requirements of some practice, activity, or capacity. The second step is to provide an explanation of why the existence of these constitutive features entails that agents’ moral reasons are genuinely normative. According to the most common version of this strategy, which I’ve called agential

74 Whether it can or not would require a dissertation of its own.
constitutivism, the relevant constitutive features are constitutive of agency. If the moral-reason-explaining features are constitutive of agency, it is relatively easy to see how the universality of moral rationalism can be accommodated. Every agent has a normative reason to \textit{phi} when she has a moral reason to \textit{phi} because every agent has the constitutive features of agency.

But second-person constitutivism is not agential, or at least it is not obviously agential.\textsuperscript{75} The relevant moral-normativity-grounding constitutive features are, in Darwall’s account, those constitutive presuppositions we must make to occupy the second-person standpoint (rather than agency). But this approach of grounding morality in interpersonal interaction generates a problem; it seems perfectly possible for a rational agent to fail to make or acknowledge second-personal claims—to ‘opt out’ of the second-person standpoint. Therefore, it seems possible for a rational agent to lack moral reasons (even assuming moral normativity is indeed grounded in the second-person standpoint).\textsuperscript{76} As Korsgaard asks.

\begin{quote}
\textsuperscript{75} In the next chapter, I will argue that second-person constitutivism is agential. But I accept that this is a non-obvious claim.
\end{quote}

\begin{quote}
\textsuperscript{76} A separate worry is if whether unavoidability (a.k.a. inescapability or non-optionality) is the right kind of property to vindicate the normativity of whatever follows from it (see Enoch 2006 esp. 187-192). Consider Hookway’s case in which a psychoanalyst tells you that you are incapable of giving up certain beliefs (Hookway 1999, 178). The unavoidability of the beliefs does not make them justified. In the practical case, Enoch provides the case of a game of chess which you cannot quit: ‘And assume that sacrificing a pawn is the thing you have most chess-related reason to do […] Well, do you have a reason to sacrifice a pawn? Not it seems to me, if you don’t have a normative reason to play or win the game, and this even if you can’t quit.’ (Enoch 2006, 189) See the first chapter for more discussion of this kind of objection.
\end{quote}
‘Suppose someone does not hold herself or anyone else responsible for meeting demands, or resent it when they do not?’ (Korsgaard 2007, 20) Such a person seems at least possible, and that’s enough for the objection to get off the ground. Let’s call this the optionality challenge. If a person could ‘opt out’ of the second-person standpoint, then it seems equally possible to opt out of moral normativity, which violates moral rationalism’s modally robust universalism.

To be clear, the optionality challenge is not a challenge to the main thrust of Darwall’s argument (that Kantian, contractualist morality is implicit in second-personal address). Rather, it is a challenge to the capacity of his view to address sceptical challenges to morality’s normativity. It is a challenge, in other words, to second-person constitutivism as a strategy for answering sceptical challenges to morality’s practical authority. Accept his main claims, that: (1) the second-person standpoint involves certain normative felicity conditions and (2) moral normativity follows from these conditions. Accept that a particular concept of morality—morality as (second-personally grounded) equal accountability—is necessarily presupposed from the second-person standpoint. Morality could still be, in Kant’s words, a ‘figment of the mind’ (Kant GW S2). Even accepting Darwall’s claims, it might still be that there is no normative reason for an agent to act morally.

Darwall acknowledges the worry:

A natural response to this argument may be to agree that if the presupposition that addressees can freely acknowledge and make authoritative demands of themselves from such a common standpoint is part of the very idea of second-personal address, then perhaps it does commit us to the equal dignity of free and rational persona, but then to deny that second-personal address, so understood, is anything we need have much of a stake in. (Darwall 2006, 138)

And, elsewhere:

What can assure us that the notion of a second-personal reason and its companion ideas of valid demand, accountability, and the authority to claim and demand are not all themselves empty?
Even if taking up the second-person stance commits us to equal dignity and autonomy, that is consistent with that standpoint and its associated commitments being no more than rationally optional, or worse, illusory. (Darwall 2006, 277)

It is worth acknowledging that this objection does not apply only to Darwall’s account. It seems equally worrying for second-person constitutivism generally. Occupying the second-person standpoint means engaging in interpersonal interaction. But, as de Maagt has recently argued, any second-personal interaction that plausibly has ‘a principle of interpersonal morality among its necessary conditions of possibility’ appears to be ‘merely optional’ (de Maagt 2018, 2767). For the remainder of this chapter, I shall explain Darwall’s and Korsgaard’s proposed solutions to such challenges and argue that they are inadequate. I’ll also briefly outline my proposed solution as I conclude the chapter.

4. Darwall’s Solutions

Darwall offers three lines of argument which could be understood as responses to the optionality objection. He claims that we must occupy the second-person standpoint because (1) there is great psychological pressure on us to do so and (2) that doing so is indispensable. He also offers a companions-in-guilt style argument in favour of the genuine normativity of second-personal reasons. I’ll explain why none of these points provides a satisfying answer to the objection.

4.1. Psychological Pressure

A first line of response to the optionality objection in *The Second Person Standpoint* is that just about all of us seem to fall into the second-person standpoint because it’s incredibly difficult for us to avoid it completely. Darwall says:
The reactive attitudes, along with their second-personal presuppositions, are aspects of the human condition that are not optional in any realistic psychological sense. Recognizable human life without second-personal thought and experience is simply unimaginable (Darwall 2006, 278).

Korsgaard finds this response unconvincing:

```
[T]hese [psychological] pressures [to occupy the second-person standpoint] do not show that there is any incoherence in avoiding second-personal address. […] If we do not have to enter the second-person standpoint, why must we have second-personal reasons at all?’ (Korsgaard 2007, 21)
```

Korsgaard is right that the existence of strong psychological currents that pull us toward the second-person standpoint does not adequately address the optionality challenge. That something is difficult to avoid does not make it incoherent (and therefore practically irrational) nor impossible to avoid. And, if we mustn’t occupy the second-person standpoint, then we mustn’t have second-personal reasons. So, our moral reasons wouldn’t be universal in the modally robust sense specified by moral rationalism.

In Korsgaard’s reading, Darwall’s statements amount to the weak claim that there is a great deal of ‘psychological pressure’ on us to occupy the second-person standpoint. Darwall makes some further claims but none of them seem to invalidate Korsgaard’s criticism. One of these further claims is that human life would be ‘unimaginable’ without us occupying the second-person point of view and that opting out is an option in no psychologically realistic sense. But if, as seems probable, this is simply a claim about the limits of our imaginative capacities, then Korsgaard’s criticism stands; that we can’t imagine a human failing to occupy the second-person standpoint does not entail it is impossible. Or, perhaps Darwall means to claim that human life in the collective sense of functional human communities could not exist without our occupying the second-person standpoint. That point can be conceded by Korsgaard too. The possibility of opting out of the second-person standpoint for
an individual agent is enough for her objection to stand, even if it means that a particular agent is incapable of navigating a community.

Before moving on to Darwall’s other responses, I want to briefly flag a possible modification of the psychological pressure response. Resisting the psychological pressure on us to occupy the second-person standpoint does not seem impossible, strictly speaking. It may seem implausible that someone would ever choose that option, but so long as there is the option, it seems possible for an autonomous agent to choose it. We think of ourselves as free-willed even while under massive psychological pressure. So, perhaps the psychological pressure response should be taken to claim something far stronger than Korsgaard takes Darwall to be claiming.

The stronger claim that Darwall might make is that opting out of the second-person standpoint is unimaginable because it is impossible given the nature of rational agency. In other words, he might be making a further constitutive claim. Not only is contractualist morality constitutive of the second-person standpoint, but occupying the second-person standpoint is constitutive of rational agency. Our capacity to exercise rational agency is partly explained by our occupying the second-person standpoint. This move would make second-person constitutivism a form of agential constitutivism. It would also mean that every rational agent would have normative moral reasons, accommodating moral rationalism. This proposal is admittedly not a particularly plausible reading of Darwall’s argument, especially given that Darwall does not have strong rationalist tendencies. I’ll develop this kind of response in more depth in the next chapter.

4.2. Indispensability

Darwall’s second response to the optionality objection is that a conception of our practical reason which leaves out the second-person would be impoverished:
I argue that the second-person standpoint gives us a perspective on our own agency that enables us to appreciate a fundamental difference between theoretical and practical reason and so improves our grasp of reasons for acting. Were someone somehow to avoid taking it up, consequently, she would fail to appreciate what we, who have taken it up, can validate as reasons from a more comprehensive view that includes it. (Darwall 2006, 277)

I understand Darwall to be offering a kind of indispensability argument here. The major premise is that the cost of failing to occupy the second-person standpoint—the damage it does to our practical agency—is too high a price to accept. The conclusion is that we must occupy the second-person standpoint. My analysis will focus on Darwall’s argument, but it’s noteworthy that other second-person constitutivists make similar indispensability arguments. Karl-Otto Apel calls the cost of failing to occupy the second-person standpoint ‘paranoid-autistic loss of self’ while Onora O’Neill threatens being ‘left in solitary and thoughtless silence’ (Apel 1975, 268, O’Neill 1986, 539, as quoted in de Maagt 2018, 2775).

For the sake of argument, accept that occupying the second-person standpoint is important to our agency and, therefore, that failing to occupy it would impose costs on us. How can this premise support the conclusion that we must occupy the second-person standpoint? Here’s a critique from Korsgaard:

[T]he person who declines to take up the second-person standpoint fails to know something intimate and important about his own agency. But unless he has a reason to take up the second-person standpoint and its presuppositions, it is possible that he will never know. (Korsgaard 2007, 22)

I think there are at least two plausible readings of Korsgaard’s critique. Each presupposes Darwall’s argument works a different way, depending on whether the indispensability of the second person is taken to entail that one must, as a metaphysical matter, occupy the second-person standpoint, or that one must, as a normative matter, occupy the second-person standpoint.
4.2.1. The Metaphysical Reading

Let’s begin with what I’ll call the *metaphysical* reading of the indispensability argument. In this reading, the conclusion—that we *must* occupy the second-person standpoint—is understood to be a claim about the *metaphysical* necessity of our occupation of the second-person standpoint given the kinds of things we are.

This is a highly desirable conclusion. It would straightforwardly accommodate moral rationalism. If all agents necessarily occupy the second-person standpoint, and moral conclusions follow from the necessary presuppositions we make in order to occupy the second-person standpoint, then the normativity of those moral conclusions will be secured. If it is impossible for us to opt out from the second-person standpoint (and if morality is partly constitutive of the second-person standpoint) then it is also impossible for us to opt out from morality’s demands.

The problem with an indispensability argument, on the metaphysical reading, is that it’s the wrong kind of argument to yield a metaphysical conclusion. The major premise of the argument is an evaluative claim about the costs of failing to occupy the second-person standpoint. Combined with some minor implicit premises—principles which take these evaluative facts and derive normative facts about what we ought to do—we may be able to reach a normative conclusion. (This is the normative reading of the indispensability argument I will discuss shortly; we *must* occupy the second-person standpoint in the *normative* sense of must, i.e. *ought*, because of the costs of failing to do so.) The metaphysical reading claims that that it is *impossible* to fail to occupy the second-person standpoint, because of the costs of failing to do so. The problem is that this would mean we have to derive metaphysical conclusions from evaluative premises. But how could incurring certain costs have a bearing on what is metaphysically possible? The world does not ban certain costs from being paid. To channel H.P. Lovecraft, there are no evaluative or normative guardrails that intervene to
curtail nature from unfolding in horrifying directions. So the premises of the indispensability argument, in this reading, are orthogonal to the conclusion. The reason for this is the general point that an indispensability argument of this kind seems incapable of providing insight into what is metaphysically possible.

In sum, the metaphysical reading of the indispensability argument yields the ideal conclusion, but it does so at the cost of argumentative validity.

4.2.2. The Normative Reading

What I’ll call the ‘normative’ reading of Darwall’s argument assumes that if a person has reason to take up the second-person standpoint then she also has reason to take up its (moral) presuppositions. In this reading, we don’t need to actually occupy the second-person standpoint in order to secure normative moral reasons; all we need is reason to occupy the second-person standpoint. This reason-to-occupy then automatically gives us reasons to (1) make the pro-moral presuppositions that are constitutive of the second-person standpoint and (2) act morally. In other words, the conclusion of Darwall’s indispensability argument, on this interpretation, is that we must occupy the second-person standpoint in the normative sense that we ought to, that rationality demands it.

A possible objection to the normative reading, and a potential reason for favouring the metaphysical reading, is that its conclusion seems insufficient to solve the optionality objection. Recall that the reason we care about the second-person standpoint being something we must occupy is that we want to accommodate moral rationalism. Darwall argues that we can derive moral conclusions from the constitutive presuppositions of the second-person standpoint. If that argument succeeds, then we can guarantee those occupying the second-person standpoint have normative reasons to act morally. To accommodate moral rationalism, we also need it to be the case that necessary we all occupy the second-person standpoint. The
mere fact that you ought to occupy the second-person standpoint seems insufficient to
guarantee you have normative reasons to act morally; it does not guarantee that you actually
take on the constitutive presuppositions of the second-person standpoint. To guarantee the
normativity of the second-person standpoint’s moral entailments, you must guarantee all
agents actually have the constitutive presuppositions of the second-person standpoint. If we
want the modally robust universality of moral reasons that moral rationalism suggests, we
need whatever we derive those reasons from to be similarly modally robust and universal.

But I think this objection is a bit too downbeat. If Darwall is right, on the normative
reading of the indispensability argument, then we have reasons to occupy the second-person
standpoint. And once we are in the second-person standpoint, we have reasons to stay and
reasons to act morally. So perhaps Darwall can accommodate many moral rationalist
intuitions. He can at least explain why an agent is necessarily irrational if he is acting
immorally. Either the agent is occupying the second-person standpoint or not. If he is, then it
is (on Darwall’s account) irrational to act immorally. And if he is not, then the agent is being
irrational in failing to act on his reasons to occupy the second-person standpoint in the first
place.

This seems like a good result, but there are two significant limitations to this response
worth noting. The first is that it does not entail that all possible agents always have normative
reasons to comply with their moral reasons. It only says they are irrational in some sense
whenever they fail to do so. If a cold-blooded, non-second-person engaging murderer kills
someone, it doesn’t seem like we’d be satisfied to condemn them for failing to act on their
reasons to occupy the second-person standpoint; we’d want to condemn them for failing to
act on their moral reasons. But this response doesn’t entitle us to do so.

The second limitation is that, at best, it shows that we have some pro tanto reasons to
occupy the second-person standpoint. But nothing in Darwall’s argument elucidates how
strong these reasons are, nor what reasons we might have to refrain from occupying the second-person standpoint (and how strong they might be). There might also be certain benefits for failing to occupy the second-person standpoint (e.g. a lack of interpersonal responsibility). So, to show it is always, for every possible agent, practically irrational not to occupy the second-person standpoint, Darwall would additionally need to argue that the balance of reasons is (universally) in favour of occupying. And, if some agents may lack conclusive reason to occupy the second-person standpoint, then it is possible for agents to act immorally or amoral and be perfectly rational. The fact that there are some (inconclusive) reasons to occupy the second-person standpoint, appreciable from within that standpoint, does not (neither normatively nor metaphysically) entail that we must occupy the standpoint.

What we need to resolve this dispute is an account of when agents have objective reasons to engage in practices, or occupy standpoints, in order to reap the benefits that are accessible and obtainable only from within those practices and standpoints. And I don’t think that has a straightforward answer.

This brings us to Korsgaard’s critique. If we understand her as critiquing the normative reading of Darwall’s argument, then she is best understood as claiming that it’s possible a person lacks any normative reason to occupy the second-person standpoint. If this possibility exists, then it is also possible for a person to lack second-personal reasons and therefore lack moral reasons, violating moral rationalism.

However, the promise of benefits (or the threat of losing them) is, prima facie, reason-providing. Korsgaard also accepts Darwall’s claims about the benefits to our agency that the second-person standpoint provides. So how could someone lack reason to occupy the second-person standpoint if that’s required for obtaining the relevant benefits? This possibility exists, according to Korsgaard, because it’s possible for an agent who does not participate in the second-person standpoint to lack knowledge of its benefits. The implication is that if a person
lacks knowledge of certain benefits, then they can fail to provide him a reason to participate in a practice that provides access to those benefits. If the benefits of second-personal agency provide reasons to occupy the second-person standpoint only to those who already occupy it, we can’t use those benefits to explain why a non-participant ought to participate. This would mean that all the indispensability argument permits us to conclude is the rather weak claim that those who occupy the second-person standpoint must (in the normative sense) occupy the second-person standpoint. However, this conclusion does not address the optionality objection.

The problem with Korsgaard’s critique is that normative reasons, in this context, are generally thought to be objective reasons. That is, their existence is not hostage to our beliefs or knowledge. Whether you have an objective reason to drink the petrol in the glass in front of you does not, at least ordinarily, depend on whether you are aware that the glass contains petrol (and not gin) or that drinking petrol is a bad idea. This can be contrasted with a subjective reason which does depend on your beliefs. If you believe the petrol is a nice glass of gin (and you like gin), you have a subjective reason to drink it. Since we are interested in objective normative reasons, though, whether an agent has a reason to occupy the second-person standpoint cannot depend on whether or not he has knowledge of the benefits of the second-person. Those benefits are, presumably, reason-providing regardless of the agent’s epistemic state. So the question we need to ask is whether or not agents have objective reasons to participate in activities that provide them with benefits (that are only recognisable from within the activity).

I think the answer to that question depends on the nature of the benefits and activities/practices. Sometimes, I think, agents have no such reasons. There are certain
benefits which are what I’ll call ‘internal’ to a practice. Take chess. Controlling the centre of the board is an obtainable benefit in a chess game. To a chess player, the promise of this benefit (or the threat of losing it) is reason-providing; it provides reasons to perform certain moves and refrain from others. The interesting thing about this benefit is that it seems incapable of providing reasons to a non-chess player. This benefit is ‘internal to chess’ in the sense that the beneficial status it comprises only exists within chess, and it is, therefore, only reason-providing to those who are playing chess. A non-chess player is not provided reasons by the promise of the benefit of controlling the centre of the board. An explanation of that fact modelled on Korsgaard’s explanation of the second-personal case would run like this. A non-player lacks knowledge of the benefits provided by controlling the centre of the board, and therefore she lacks a reason to take the means to achieving that end (which includes playing chess). But this doesn’t seem like the right explanation. It is not a lack of knowledge that prevents the benefit from providing reasons, but the non-player’s lack of engagement that explains her lack of any reason to control the centre of the board (and play chess).

Let’s take one more example of benefits that are ‘internal’ to a particular practice. Some people take the norms of etiquette to be normative whereas others don’t. Those who do have an appreciation for a whole class of reasons for action—the reasons of etiquette. They also seem to enjoy some benefits from doing so, such as the capacity to be thought of highly by others who care about etiquette. They also derive certain internal benefits that exist only within the ‘game’ of etiquette, such as the status of being well-mannered (in the narrow etiquette-related sense). You might even say that non-etiquetters are working with a (slightly) impoverished agency, being largely insensitive to this small fiefdom of reasons. But the fact

77 It is doubtful that this is the sort of thing Darwall himself intended. Regardless, I think it is a direction of argument worth considering.
that etiquetters have better access to etiquette-related reasons and certain internal benefits provides no obvious reason to the non-etiquetters to start practicing etiquette. Likewise, there might be good reasons to occupy the second-person standpoint that are clear from the second-person standpoint. But if you don’t take up that standpoint you don’t have those reasons.

If the benefits that we enjoy by occupying the second-person standpoint are ‘internal’ like this, if they only exist as benefits within the ‘game’ of second-personal interaction and second-personal agency, then Darwall’s indispensability argument fails. The fact that there are second-personal benefits is not reason-providing to those outside of the second-person standpoint, not at all. Someone who does not engage in the second person has no reason to engage (regardless of the internal benefits), and, therefore, has no guarantee of reasons to act morally. To put the point more bluntly, it may be true that there are reasons to take heroin you can appreciate only if you take heroin. But if this is all Darwall is claiming about the second-person standpoint, then he’s not vindicated the normativity of whatever follows from the second-person standpoint.

Is this a conclusive critique of Darwall’s indispensability argument? I don’t feel prepared to claim that. The benefits of the second-person standpoint might not be ‘internal’ like this; they might be reason-providing even to those who are not actually engaged in the second-person standpoint.

Consider, for example, the reasons of prudence and how they intuitively differ to the reasons of chess and etiquette. It’s relatively easy to view the normative domain of etiquette (normative in the sense of rule-based) as ‘low grade’, as not really normative (normative in the sense of reason-providing), at least not necessarily. But it’s much harder to do the same for the domain of prudence/wellbeing. There’s some intuitive sense in which the reasons of prudence apply to you regardless of whether you happen to care about your own wellbeing. A depressed person has a reason to get out and exercise, even if he is so depressed that he has
no interest in increasing his wellbeing. Prudential benefits seem reason-providing regardless of whether an agent cares about his wellbeing, or takes it to be a source of normativity. Perhaps the benefits of the second-person standpoint are like prudential benefits in this way.

The question, of course, is which model second-person reasons follow. Perhaps they are similar to prudential reasons in the sense that they are reason-providing regardless of whether we are already ‘players’ in the second-personal game—regardless of whether we already take second-person reasons to be normative.

So, whether this critique of the normative reading of Darwall’s indispensability argument succeeds depends on whether the benefits of the second-person are reason-providing to those who aren’t currently participating in the second-person standpoint. I must admit, I lack a strong intuition either way, and I have not discovered arguments for either view. I will merely say that the indispensability argument, in its current form, fails to provide such an argument, and hence is importantly incomplete.

In sum, Korsgaard’s knowledge-based critique of the normative reading of Darwall’s argument seems to get things slightly wrong. My suggested ‘internal benefit’ critique of the normative reading raises more serious challenges to the indispensability argument. At the very least, it shows that the indispensability requires further supplementation if it is to be made convincing.

Taking a step back, accommodating moral rationalism would be far straightforward given the conclusion of the metaphysical reading of the indispensability argument. If it is impossible to avoid the second-person standpoint, then it seems similarly impossible to lack the commitment to the moral norms to which we are necessarily committed from that standpoint. The normative reading of the indispensability argument, on the other hand, accepts that we have the possibility of ‘opting out’ of the second-person. So this strategy requires a lot more moving parts to explain why we can’t thereby ‘opt-out’ of moral
normativity. Even if we succeed in showing that we have reasons to occupy the second-person standpoint is sufficient, much more would have to be said about why those reasons are necessarily better than our reasons to refrain from occupying the second-person standpoint. It’s not clear such an answer can be given.

4.3. Companions in Guilt

In his response to Korsgaard, Darwall (2007) presents a different kind of response to the optionality objection:

Moreover, whether someone actually takes up a second-person perspective or not is irrelevant to whether second-personal reasons for him to act actually exist. Even if it is adopting the second-person standpoint that commits us to accepting the existence of second-personal reasons, that does not mean that second-personal reasons themselves depend upon our actually taking up a second-person point of view. […] What’s more, I argue that, as an epistemic matter, we have no better grounds for accepting familiar reasons concerning the desirability of outcomes, self- or other-interest or well-being, or indeed, normative reasons of any other kind, than we do for accepting the second-personal reasons we can appreciate from a second-person standpoint. (Darwall 2007, 60)

This response does not deny the major premise of the optionality objection—that the second-person standpoint is optional. It denies that the objection has force. Even if the second-person standpoint is optional, Darwall argues, second-personal reasons exist and apply to us. And, furthermore, we have equal epistemic grounds for accepting them.

In Chapter 11 of The Second Person Standpoint, Darwall provides a lengthier version of this argument. This has the logical structure of a companions-in-guilt argument and runs roughly as follows:

1. Anti-global-scepticism: ‘[T]o deliberate at all, we must assume that there are some normative reasons for acting.’ (Darwall 2006, 291)
2. Parity: ‘[W]e have no less reason to accept second-personal reasons grounded in equal dignity than we do reasons for acting of other kinds’… ‘those [reasons] relating to prudence, benevolence, or whatever’ (299 and 291)

3. (Conclusion): ‘For practical purposes, at least, second-[personal] reasons grounded in our common dignity really do exist.’ (299)

If second-personal and other normative reasons have parity of acceptability, then scepticism about second-personal reasons, e.g. prudential reasons, entails scepticism about practical reasons generally. But we seem committed to the existence of at least some practical reasons insofar as we deliberate about what to do. So, given the choice is between a radical scepticism about practical normativity generally and accepting second-person reasons as genuinely normative, we should embrace the latter. Again, this companions-in-guilt style response does not dispute the possibility of an agent opting out of the second-person standpoint. Rather, it suggests that the parity of second-personal reasons with other kinds of reasons is sufficient justification of the normativity of second-personal reasons, and hence of moral reasons.

Let’s focus on Parity. Darwall begins with the following justification of the claim:

78 A quick note on anti-global-scepticism. One limitation of Darwall’s companions in guilt argument is that it does not show there are normative reasons for action. It merely claims that insofar as we deliberate we must assume that there are. In other words, the cost of failing to accept the normativity of second-personal reasons is scepticism about practical normativity generally. But this is not an impossible cost to pay. A moral sceptic may be happy to expand her scepticism to prudential reasons, the reasons of benevolence, etc. We are in the business of providing the foundations for moral normativity, but for all the companions in guilt argument shows, we may have just expanded the problem.
People often take it for granted […] that considerations of self-interest or the agent’s own welfare give her reasons for acting. But that this is a substantive normative thesis and not a conceptual truth. Someone might have full command of the concept of a reason for acting but nonetheless be so depressed that he doubts (without incoherence or conceptual confusion) that the fact that something would be for his good is any reason whatsoever for him to seek it. (Darwall 2006, 291-2)

Darwall’s conclusion here is that the normativity of prudential reasons are equally well (or poorly) grounded as the normativity of second-personal reasons. But this argument for Parity depends on some highly controversial claims. For example, analytic naturalist theories claim that certain considerations give agents normative reasons (or just are normative reasons)—and that this is a conceptual truth. Some might claim that it’s a conceptual truth that prudential reasons are normative, but deny that it’s a conceptual truth that second-personal reasons are. A posteriori naturalist theories accept Darwall’s claim that a consideration’s being reason-providing is not a conceptual truth. However, such views maintain that some kinds of consideration provide/are normative reasons and that some do

79 Jean Hampton (Part II, 1998) offers a similar argument against Mackie and other naturalist sceptics of morality, according to which instrumental reasons are just as hard to accept from a naturalistic perspective as moral reasons.

80 E.g. Jackson (1998) and Finlay (2014)

81 Of course, such theories need to answer certain well-known objections such as Moore’s open question argument.

not/are not. Depending on the particular account, second-personal reasons will, or won’t, be classed as genuine normative reasons. So we can’t assume Parity will be true regardless of the metaethical view we endorse.

It looks like we can’t establish Parity ‘on the cheap’, i.e. without looking at specific metaethical views. Darwall takes up this challenge by outlining two metaethical ‘strategies’ that ‘seem to exhaust the field’ (Darwall 2006, 292). He argues that the normativity of second-personal reasons fits in neatly according to both strategies. These are recognitional and constructivist strategies, also known as substantive realism and procedural realism.

On the recognitional/substantive realist account, normative reasons are mind-independent normative facts; their status as normative reasons does not depend on the verdict of any deliberator. On such an account, Parity demands that our reasons for thinking that our reasons-judgments are veridical are equally good for second-personal and other reasons. Darwall asks:

What possible reason could we have for thinking that our object-dependent desires and intuitive judgments of intrinsic values or desirability are evidentiary or veridical that we would not also have for so regarding intuitive judgments of second-personal reasons and second-personal authority? (Darwall 2006, 298)

The idea is to shift the burden of proof onto the sceptic about second-personal normativity. We need some positive case for the view that our intuitive judgments about the

---

83 Such views have to contend with Horgan and Timmons’ (1991) Twin Earth objection.

84 Darwall takes this terminology from Cullity and Gaut (1997, 4).

85 This terminology comes from Korsgaard (1996 36-7).
normativity second-personal reasons are systematically error-ridden, while at the same time, our intuitive judgments about the normativity of e.g. prudential reasons are largely veridical.

There are, however, recognitional accounts of normative reasons which seem at least friendlier to the normativity of, e.g. prudential reasons than they are to the normativity of second-personal reasons. Take the Humean account of normative reasons I’ve been working with, according to which our normative reasons are always (partly) explained by our desires. On this view, it is much easier to explain why an agent has normative reasons to promote her own good (particularly on the assumption that there are strong links between an agent’s good and the ends she desires) than it is to explain why an agent has normative reasons to take others to be a source of authority capable of making claims on her will. The same lack of parity may appear for other recognitional metaethical views. So, assuming a recognitional account, a great deal more needs to be said to establish Parity.

On a constructivist/procedural realist account, on the other hand, normative reasons are not simply ‘out there’ to be recognised. Rather, normative reasons exist as features or entailments of the practical standpoint—the standpoint a rational agent takes up in deliberating about what to do. According to the constructivist, there is some authoritative deliberative procedure, the results of which are correct simply in virtue of being the results of the procedure. If a (perhaps ideal) agent, following the right deliberative procedure, takes second-personal reasons to be genuinely normative, then they are. On this account, Parity demands that we have equal reason to think the results of the constructivist deliberative

86 Admittedly, Darwall has independent reasons for rejecting such views. So presumably he wouldn’t be impressed with this move. But the point is dialectic rather than substantial. Darwall’s companions in guilt argument seems to depend on unacknowledged substantial metaethical commitments.
procedure would include second-personal reasons as it would other reasons. Darwall claims that the normative judgments of a deliberating agent within the second-person standpoint will ‘survive’ the relevant procedure. But this returns us to the central problem raised by the optionality challenge. Unless we have some reason to think an agent must occupy the second-person standpoint, there is no obvious reason to think that their normative judgments will include any second-personal reasons in the first place. So on the constructivist account, the companions-in-guilt argument fails to transcend the optionality challenge.

4.4. Summary of Darwall’s Proposed Solutions

The optionality objection states that if we can ‘opt out’ from the second-person standpoint we can also opt out from morality, which violates moral rationalism. If we can’t accommodate moral rationalism, then we’ve failed to secure the foundations of moral normativity. Darwall’s solutions to the optionality objection are either insufficient or incomplete. The first of Darwall’s solutions fails because the fact it is psychologically difficult to avoid the second-person standpoint does not entail it is impossible. The second solution faces a dilemma. If we interpret it as suggesting that we must occupy the second-person standpoint in a normative sense, then the conclusion is either insufficient or, at the very least, inadequately supported by the existing argument. If we interpret it in a metaphysical sense of must, then the conclusion is not established by the indispensability argument. Darwall’s argument is built on claims about the various costs to our agency that are involved in failing to occupy the second-person standpoint. But evaluative costs do not generate metaphysical necessities. The third solution Darwall offers is that the fact the second-person standpoint is optional is irrelevant because its reasons are normative. This conclusion is reached via a companions-in-guilt argument, according to which we can either abandon practical normativity entirely or accept the normativity of second-personal (and
therefore moral) reasons. But Darwall’s claim that second-personal and other reasons are companions, i.e. Parity is true, is unconvincing and implausible on many metaethical views. We have many reasons to doubt that second-personal reasons are on par with, e.g., prudential reasons. So the optionality objection retains its potency.

5. Korsgaard’s Agential Solution

Perhaps the optionality objection can be overcome by identifying a further practice or activity that (1) necessarily involves occupying the second-person standpoint and (2) is impossible to opt out from. The most obvious and popular candidate for this role, already favoured by agential constitutivists, is agency. Recasting second-person constitutivism as a form of agential constitutivism is Korsgaard’s strategy.

5.1. The Second-Person Within

Korsgaard endorses both that:

[A]ll [moral] reasons are second personal and also […] that ‘moral obligations can be grounded in the constraints of first-personal deliberation alone.’ (Korsgaard 2007, 10 quoting Darwall 2006, 11)

These seemingly paradoxical claims are compatible, argues Korsgaard, because the self-reflective nature of the rational agent’s mind is itself inescapably second-personal. And this, she thinks, solves the optionality objection. The Kantian conception of first-person deliberation involves

[A] relation of authority over ourselves and that we are as a consequence also accountable to ourselves […] Thus I act under my own authority as a lawgiver, and I am accountable to myself if I do not. So my reasons—and indeed practical reasons in general—are grounded in the authority the human mind necessarily has over itself. We might put this by saying that because of the reflective structure of human consciousness, I think that every rational agent stands in
what Darwall would call a second-personal relation to herself—she has a second-personal voice within. (Korsgaard 2007, 10-11)

In other words, the second-person standpoint is implicit in the first-person deliberative standpoint.

How does the second-person standpoint being implicit in the first-person standpoint help us address the optionality objection? Korsgaard claims that ‘the second-person standpoint does not seem to be unavoidable […] the way the standpoint of first-person deliberation is.’ (Korsgaard 2007, 22) In other words, the first-person standpoint is unavoidable for rational agents. So, if morality is normative for any agent who occupies this standpoint, then we can accommodate moral rationalism. The trouble, as I have described throughout this chapter, is that it’s questionable that the second-person standpoint is unavoidable. But, if Korsgaard is correct that the second-person standpoint is implicit in the first-person, then the second-person is unavoidable too. Any person who engages in first-person deliberation will be engaging in second-personal engagement (with himself). And all agents must engage in first-person deliberation about what to do (if they are to count as agents). Additionally, if Darwall is right that we necessarily presuppose the basic norms of contractualist morality insofar as we occupy the second-person standpoint, then it follows that we all have moral reasons in the sense demanded by moral rationalism.

5.2. Darwall’s Response

Darwall rejects Korsgaard’s proposed solution because he thinks the ‘second-personal authority within’ is not second-personal in the relevant, morality-grounding sense:

I don’t yet see how authority of this kind necessarily involves answerability or accountability. […] It gives one no authority to claim or demand prudence of oneself. That imprudence involves acting against reason is one thing; that it involves letting oneself down is another. As I
see it, therefore, the problem is not how to get from accountability to oneself to accountability to others. I don’t yet see how in giving the kind of authority to oneself that reflective consciousness necessarily involves we yet get the idea that I owe this to anyone, either to myself or to others. (Darwall 2007, 56)

Who is right here? Korsgaard is right that the first-person practical standpoint contains something of the ‘I-you-me structure of reciprocal address’ which characterises the second-personal (at least the ‘I-me’ part), but it is not, I think, second-personal in the same way that the paradigmatic cases Darwall discusses are. In particular, as Darwall points out, the sense in which we are answerable and accountable to ourselves, to the extent to which it is possible, seems to differ from the sense in which we are answerable and accountable to others.87

This difference between Korsgaard’s ‘second person within’ and Darwall’s more robustly interpersonal second-person standpoint does not show that Korsgaard’s claim that the second-person is implicit in the first-person is false. But it does present Korsgaard with a dilemma. On one hand, she can’t introduce too much substance to the second-person within, since it must be something that all agents fall into whenever they deliberate alone. So the sense of the second-person she uses must be fairly minimal. But, on the other hand, we still need to derive moral normativity from the constitutive presuppositions of the second-person standpoint. And the less substance we put into that standpoint, the harder the derivation will

87 The extent to which there is a significant difference between being accountable/obliged to ourselves and to others is debated. For example, Singer (1959) argues that we cannot have duties to ourselves at all. Hill claims that you can promise to yourself but that it’s not the same as promising to others in some respects (Hill 1991, Ch10). See Schaab (2021) for an argument that it is coherent to promise to oneself and release oneself from one’s promise.
become. I see no obvious reason to think that this derivation is impossible from Korsgaard’s second person within, but it does seem challenging.

In sum, Korsgaard contends that the optionality challenge can be overcome because the second-person standpoint is already implicit in the first-person practical standpoint of a self-reflective agent. Darwall rejects this ‘solution’ to the optionality challenge because he thinks that the authority of the ‘lawgiver’ within oneself is not properly second-personal in nature, and hence doesn’t account for the distinctive proto-moral relations of equal accountability he thinks are presupposed in second-personal address. Korsgaard threatens to overcome the optionality challenge at the cost of making second-personal interaction too insubstantial to presuppose interpersonal morality. The Korsgaardian solution does not fail, but it faces serious challenges.

6. Conclusion and Outline of my Proposed Solution

Let me briefly summarise. Darwall derives contractualist moral norms and reasons, those of ‘morality as equal accountability’, from the constitutive presuppositions of the second-person standpoint. That is, any person who occupies the second-person standpoint has moral reasons. But to secure the normativity of these reasons, we need to explain why agents must occupy the second-person standpoint. This is what I call the optionality objection.

I examined three possible lines of response to the objection that can be extracted from Darwall’s work: (1) that there is immense psychological pressure to occupy the second-person standpoint, (2) that occupying the second-person standpoint is indispensable, either normatively or metaphysically, and (3) a companions-in-guilt argument which claims that we have no more reason to reject second-personal reasons than we do any other kind of reason. I argued that none of these responses is convincing. I then discussed Korsgaard’s proposed solution: that each rational agent occupies a second-person relationship with himself, and that
occupying the second-person standpoint is, therefore, unavoidable. But, as Darwall points out, this ‘second person within’ seems to be a lesser sense in which we occupy the second-person standpoint. As such, it is dubious whether the kind of mutual accountability and respect that Darwall thinks is required to occupy the second-person standpoint is also required to occupy Korsgaard’s second-person within. So Korsgaard’s response is also lacking in its current form.

I’ll now offer a brief preview of my proposed solution to the optionality objection, which is developed in the next chapter. In structure, I opt for a Korsgaardian solution. I agree with her that, to answer the optionality objection, you need to show that occupying the second-person standpoint is unavoidable (for moral agents). And I also agree that the best strategy for achieving that is to show that there is some kind of practice or capacity that is truly unavoidable for an agent—something such as rational agency itself. But, as should now be clear, I agree with Darwall that, to derive the kind of contractualist ethics of mutual respect that Darwall does from the constitutive presuppositions of the second-person standpoint, we need something more full-bloodedly interpersonal than Korsgaard’s ‘second person within’. The second-person within does not obviously commit individuals to the kind of accountability and respect of persons we need.

So what practice is there, that is agency itself or is constitutive of agency that also commits its participants to adopt the kind of substantial second-personal stance Darwall is interested in? My answer is language. Language is not simply one more of the many abilities humans have, such as the capacity to run, laugh, or digest. Linguistic competence both requires and enables entry to (and navigation of) what Sellars famously called ‘the logical space of reasons’. Only those who dwell in the space of reasons can be rightly thought of as rational agents—the kinds of beings that are responsive to reasons, who can act, consider, and evaluate what they ought to do, rather than merely respond or behave. I argue that rational
and moral agency, at least in its familiar human form, requires us to be ongoing and active participants in the logical space of reasons, and thereby requires us to be competent linguistic beings. I further argue that linguistic competence necessarily involves occupying the second-person standpoint in the substantially interpersonal sense Darwall is after and Korsgaard fails to secure. Competent linguistic activity necessarily demands certain practical commitments of its participants, commitments to second-personal respect and mutual accountability, and hence to standards of reasonableness which underlie contractualist morality.

In sum, I’ll argue that (1) the possession of language is partly constitutive of rational agency, (2) that language-users necessarily occupy the second-person standpoint and thereby undertake second-personal pro-moral commitments, and (3) that, assuming Darwall’s derivation of a contractualist morality from such commitments, then a Kantian tune can be played on a Humean instrument. Humean constitutivism can vindicate moral rationalism.
Chapter Five: Linguistic Constitutivism

I suspect […] there is something assertion-like at the core of the ethically central cases. Responsible agency requires that one stand for something, and this is something more basic than the subjection to substantive moral requirements.

– Gary Watson, “Asserting and Promising” (2004, 60)

1. Introduction

In the first three chapters of this thesis, I developed Humean constitutivism and showed how it can respond to various challenges. But many of my conclusions have been hypothetical. In particular, I have claimed the Humean account of reasons can be reconciled with moral rationalism if there were these special commitments which I’ve labelled constitutive commitments.

But are there any? Or have I merely postulated a philosophical phantom to fill the gaps in Humean constitutivism? In this chapter, I argue that constitutive commitments actually exist. So, my thesis does not just claim that Humean constitutivism would be a great theoretical package if only there were constitutive commitments. Much more interestingly, it claims that Humean constitutivism actually is a great theoretical package! Humeanism really can be reconciled with moral rationalism. The shmagency objection really can be answered. In short, it’s time to give some reason to accept the (non-obvious) fact that these constitutive commitments actually exist.

The prior discussion should have provided some indication of the kinds of entities constitutive commitments need to be. To summarise, I have argued that they should be: (1) pro-moral conative mental states and (2) constitutive of agency—in order for the Humean to
get the universality, modal status, and weight of moral reasons right.\textsuperscript{88} I will address each of these key criteria in this chapter.

I defend (1) and (2) by arguing for a \textit{linguistic} form of constitutivism and a linguistic account of agency respectively. In section 2, I explain linguistic constitutivism, distinguishing its agential form from its non-agential form. Then, in sections 3, and 4, I argue that certain moral norms apply to us—are ‘robustly’ normative for us—merely in virtue of the fact that we are competent linguistic beings, fellow inhabitants of something like what Sellars referred to as ‘the logical space of reasons’ (Sellars 1997, §36). In the latter part of this chapter, section 5, I argue that this form of linguistic competence is partly constitutive of full-blooded rational agency (or, as an attractive backup option, that it is constitutive of our distinctive \textit{human} form of rational agency). The arguments jointly entail an agential form of linguistic constitutivism. In section 6, I consider and respond to a number of potential counterexamples to my proposal, a rogue’s gallery of characters who attempt, in various ways, to opt out from the authority of moral normativity.

Before getting into the arguments, it’s worth mentioning that several existing philosophical accounts posit a close connection between morality and language.\textsuperscript{89} But, for my ultimate conclusion to go through, I don’t need every instance of language use to commit one to the second-person standpoint. Something like this more ambitious claim is notably made

\textsuperscript{88} By pro-moral I mean that having a commitment of this kind should generate some prima facie normative reasons to act morally.

\textsuperscript{89} I am thinking in particular of the discourse ethical theories of Apel (1980) and Habermas (1990), and to a lesser extent the conversational approaches to moral responsibility pursued by Watson (2004) and McKenna (2012). My arguments won’t discuss these views except in passing, with the exception of Watson.
by Cuneo (2014).\textsuperscript{90} I’ll attempt to show that more minimal claims will do. All my argument requires is that counting as a language user means occupying the second-person standpoint and being committed, in a volitional sense, to giving second-personal claims some weight in practical deliberation. Roughly, my account will first establish that being a player in the linguistic game of assertion involves occupying the second-person standpoint and taking on related proto-moral commitments. The argument then proceeds, drawing on Brandom’s development of inferentialism, via the claim that assertion is the fundamental linguistic

\begin{quote}
\textsuperscript{90}Although he does not specify that the normative statues are second-personal, Cuneo does argue that all instances of speech entail moral facts. Specifically, Cuneo endorses the: ‘Normative Theory of Speech: An agent’s performing some locutionary act $\phi$ing count-generates her performing some illocutionary $\psi$ing in virtue of her having the rights, responsibilities and obligations of being a speaker.’ (Cuneo 2014, 32) Why accept the normative theory of speech? Cuneo argues that it seems reasonable to suppose that there is one unified explanation for how locutions come to count as illocutions. Since the normative theory seems to work for promises and assertions that seems like a good prima facie reason to think that it works for illocutions generally. Cuneo offers the supplementary argument that the only existing alternative to the normative theory of speech—the perlocutionary intentional theory of speech—is inferior, and on this basis we should believe in the best theory going: the normative theory of speech. Accounts of language that seem supportive of the normative theory of speech include Searle (1969) and Brandom (1994, 2000). I am reluctant to rely on Cuneo’s argument for establishing the conclusions I am interested in because (a) it relies on stronger premises than I need for my argument to go through and (b) it relies on expanding the plausible case of two speech-act types (promising and assertion) to all acts generally. But it is unclear that all speech-acts have, as their felicity conditions, these rights, responsibilities and obligations, e.g. asking (see esp. Laskowski 2017).
\end{quote}
practice, such that to count as a linguistic being, one must be a participant in the practice of assertion. Finally, I argue that rational agency requires language, and hence that all agents are subject to the norms of language and assertion. I then round off the chapter responding to a ‘rogue’s gallery’ of possible counterexamples to my argument.

2. The Promise of Linguistic Constitutivism

The arguments I make in this chapter are fairly structurally complicated. So I want to begin by giving the reader an indication of what the arguments in this chapter will jointly yield: a kind of constitutivism based on the constitutive commitments of language-users. According to linguistic constitutivism, as I’ll call it, the normativity of some moral norms and reasons is explained by facts about the constitutive features of linguistically competent beings, specifically the practical commitments they necessarily undertake as language-users.91

91 Is this linguistic constitutivism really a species of constitutivism? Yes. Katsafanas is representative of the dominant form of constitutivism in describing constitutivism as the claim that ‘certain normative claims apply to us merely in virtue of the fact that we are agents’ (Katsafanas 2018, 367). On this account, an account is only constitutivist if it is built on the kind ‘agent’ (or ‘action’). Hence, linguistic constitutivism would not qualify, since it is built on the kind ‘competent language-user’. Ultimately, whether linguistic constitutivism falls under the constitutivist umbrella will not matter, since I will also claim that linguistic capacity is partly constitutive of full-blooded rational agency. This latter argument will firmly place linguistic constitutivism under the traditional agential conception of constitutivism. But non-agential linguistic constitutivism would count as a constitutivist view, in the relevantly interesting sense, according to a more accommodating characterisation of constitutivism provided by Michael Smith (2017). On Smith’s characterisation, constitutivism is the view that: ‘normative facts of certain kinds are explained by facts about the constitutive features of
2.1. Non-Agential Linguistic Constitutivism

Linguistic constitutivism, even absent the coming arguments linking linguistic competence with rational agency (section 5), provides an impressive, if partial, vindication of the normativity of moral norms and reasons by reference to facts about the constitutive features of linguistically competent beings. It claims that moral norms and reasons are normative for us merely by virtue of the fact that we are language users.

Let me briefly outline how it does so. I’ll treat myself to Darwall’s derivation of contractualist moral norms from what we necessarily presuppose from the second-person standpoint. To very briefly recap that argument, the derivation of some kind of contractualist morality from the second-person standpoint runs as follows. In the necessary reciprocal recognition of the second-person standpoint, persons are implicitly:

[C]ommitted alike to their mutual accountability as equal free and rational persons, and this commits them […] to a constraint of reasonableness of any demands they address: they must be able to expect their addresses to accept, or not reasonable reject, their demands as free and rational persons. (Darwall 2006, 320)

Hence, the second-person standpoint vindicates contractualism’s central notion of reasonableness and its idea of a ‘community of mutually accountable free and rational persons’ (Darwall 2006, 320). It makes sense of why moral norms would rule out acts (i.e. give us moral reasons not to perform acts) which, in Scanlon’s formulation, would be something […] some person, or some action, or some state of affairs, or something else entirely. Constitutivists may disagree about this.’ (Smith 2017, 371) The relevant constitutive features, according to the linguistic constitutivist, are the practical commitments of language-users.

92 See the previous chapter for a more detailed explanation of this argument.
disallowed by any set of principles that no one could reasonably reject (Scanlon 1998, 153). Hence, the content of these contractualist principles define the content of our moral reasons and moral obligations to one another as equal and free persons and, indeed, as occupiers of the second-person standpoint. This is why a wholesome set of universal moral norms falls out of what is presupposed by the second-person standpoint.

In addition, I’ll argue that language-users (1) necessarily occupy the second-personal standpoint and (2) are committed, as a constitutive matter, to giving the second-personal reasons they encounter some weight in deliberation. It follows that we can get a recognisably moral set of commitments from what is constitutive of linguistic capacity. These are the basics of the linguistic constitutivist strategy.

I’ll mention two important merits of linguistic constitutivism. The first is that it seems better placed to answer the optionality challenge faced by Darwall’s account.93 The second is that it can be combined with the claim that rational agency requires linguistic capacity in order to yield a new, more plausible version of agential, Kantian constitutivism. This would be well placed to defend a strong version of moral rationalism while being better placed than its Kantian constitutivist fellows to pull the moral rabbit out of the hat of rational agency.

But putting the agential line of thought to one side for a moment, I want to briefly consider linguistic constitutivism in its non-agential form. Recall that an important motivation for constitutivism is that it promises to explain moral rationalism—the view that all of us have normative reasons to act in accordance with our moral reasons. This is an ambition that Darwall’s account does not clearly fulfil. The problem for Darwall concerns the

93 This objection is the primary subject of the previous chapter.

94 I will offer what I take to be the most compelling arguments for this view in the second half of this chapter.
apparent optionality of the second-person standpoint. Linguistic constitutivism provides a more fundamental explanation of why the second-person standpoint is non-optional or ‘inescapable’ which, in turn, promises to make possible a more satisfying and comprehensive justification of moral rationalism. It broadly follows the pattern of Korsgaard’s proposed solution (Korsgaard 2007, esp. 10-11), namely looking for a less optional practice in which the second-person is implicit. But the second-personal authority relations which we enter into as language-users are far more obviously genuinely second-personal than Korsgaard’s notion of ‘the second-person within’. Indeed, I’ll argue that the kinds of commitments and responsibilities that language users necessary adopt and presuppose do seem genuinely second-personal in Darwall’s sense. So we can agree with Darwall that we should only ground the second-person standpoint in genuinely second-personal practices, and also agree with Korsgaard that there is a less optional practice out there in which this can be done.

Linguistic constitutivism provides a more fundamental explanation of why the second-person standpoint is unavoidable; it’s built into linguistic capacity. This expands the set of agents for whom moral normativity is inescapable, from beings who happen to fall into the second-person standpoint due to contingent factors to all competent language users. So linguistic constitutivism seems well-placed to ground a modest form of moral rationalism. This is an interesting result in itself. Imagine a spectrum of views regarding the universality of moral normativity. At one extreme are the traditional revisionist Humeans, who insist that, given the right psychology agents can lack genuine normative reasons to act morally. At the

95 I am thinking of revisionist Humeans such as Williams, the early Foot, and Harman. Even the more conservatively-minded Schroeder, who once claimed it was plausible that Humeanism generates moral reasons for action for all agents, now claims, more modestly, that
other extreme are Kantians who insist that every possible rational agent always has normative reasons to act morally. Somewhere in the middle are Aristotelians like the later Foot, who claim that agents who are human beings have reasons to act morally, owing to our particular human nature. Linguistic constitutivism would fall somewhere between the Aristotelians and the Kantians, claiming that all possible agents who are language users have at least some genuinely normative reasons to act morally. This might be an appealing position to those who find it implausible that morality can follow from something as insubstantial as ‘bare rational agency’, and yet find it plausible that some kind of moral rationalism is true.

2.2. Agential Linguistic Constitutivism

In the second half of this chapter I argue that being a language-user is partly constitutive of being a full-blooded rational agent. In combination with linguistic constitutivism, this yields an agential form of constitutivism. Linguistic constitutivism is most promising and interesting, in my view, when combined with the claim that rational agency itself requires linguistic competence. What does such a view have to offer? Quite a lot.

While interesting versions of constitutivism can be grounded in the constitutive features of kinds other than agency/action, there are nevertheless powerful reasons to think agential constitutivism is more promising, particularly to those interested in the Kantian project of vindicating the universal ‘robust’ normativity of moral claims (or, in Korsgaard’s

______________

all moral agents, that is, agents with some pro-moral concerns, (the vast majority of actual agents) have (likely overriding) reasons to act morally.

96 Darwall’s (2010) comments on the impossibility of a recognisable human life avoiding the second-person standpoint seems to indicate some support for this kind of claim.
language, in locating the ‘source’ of practical/moral normativity), to be interested in agency-grounded constitutivism. Agential forms of constitutivism are uniquely well-placed to explain the universality of the moral because they explain why every possible agent would have such reasons and why they would be normatively significant.

Another benefit of this ambitious, agential interpretation of linguistic constitutivism is that it offers a new perspective from which we can diagnose issues in the existing literature. We can agree with Darwall that (1) occupying the second-person stance puts us on the hook of moral normativity and that (2) this moral normativity can be understood primarily in terms of contractualist norms reflecting attitudes of equal respect, responsibility, and mutual accountability. But we have a perspective that explains how, contra Darwall, these conclusions are entirely compatible with the traditional Kantian constitutivist ambition of explaining how moral normativity falls out of the features of rational agency, thereby overcoming the optionality challenge and being well-positioned to defend a strong version of moral rationalism.

This enriched understanding of rational agency also offers a new perspective from which to address the common worry that constitutivism is in the business of ‘pulling rabbits out of hats’ (Wiland 2012, 141). Kantian constitutivism’s critics are right that bare rational agency—as it has generally been conceived—is indeed too thin to ‘pull’ morality from. But constitutivism is actually even more like pulling rabbits out of hats than its critics realise. It seems impossible to pull a rabbit out of an empty hat. But then we look more closely at the hat and realise that it is sitting on a particularly sturdy table in which a rabbit could be hidden. In giving an account of agency’s constitutive relations with language use and an account of all that is involved in our linguistic capacities, we can ruin the trick and make it unremarkable that we can ‘pull out’ the basic materials of morality. The idea is that moral
normativity can be grounded in ‘bare rational agency’, but that rational agency involves quite a bit more than has generally been noticed.

3. Language, Morality, and the Second-Person Standpoint

This section examines the relationship between linguistic and moral norms. Specifically, I argue that competent language users necessarily occupy the second-person standpoint. I also argue that language users necessarily adopt commitments to give some weight, in practical deliberation, to second-personal claims.

How will I get there? Roughly, the idea is that counting as a linguistic player, as someone who is making and acknowledging linguistic claims, requires us to enter into distinctive normative relations of mutual commitment and accountability with one another. We cannot understand each other as saying anything, as asserting anything, without entering into such relations of commitment and mutual responsibility. These discursive relations are second-personal and presuppose all the ‘normative felicity conditions’ Darwall describes as characterising the second-person standpoint. Finally, I argue language users are necessarily committed to second-personal norms not merely in the social sense, but in the volitional, attitudinal sense. We don’t just consider ourselves committed to mutual accountability socially, we commit ourselves to mutual accountability in the sense that we necessarily adopt a distinctive pro-attitude.\(^7\)

To set out the key argumentative steps more clearly:

1. Making and acknowledging assertions requires presupposing certain commitments, rights, and responsibilities.

2. These assertive commitments (etc.) are second-personal in nature.

\(^7\)I’ll claim this is plausibly a self-governing policy, in Bratman’s terms, a commitment to giving certain considerations weight in deliberation.
3. Hence, all asserters necessarily occupy the second-person standpoint.

4. A practice can be rightly considered linguistic only if it involves making and acknowledging assertions. This is what Brandom means when he says that assertion is language’s ‘downtown’.

5. (From 4) All language users are necessarily involved in making and acknowledging assertions.

6. (From 3+5) All language users necessarily occupy the second-person standpoint.

If, additionally, Darwall is correct that insofar as one occupies the second-person standpoint, one presupposes the basic materials of contractualist morality (the equal autonomy and accountability of oneself and others), then it follows that moral norms are implicit in linguistic practice.

For reasons which will soon become clear, my focus will fall on one particular linguistic act type: assertion. In making an assertion, one is thereby taking on certain commitments and responsibilities. There is some disagreement about what assertive commitments and responsibilities are commitments and responsibilities to. Candidates include correspondence with the facts and justifiability. But my argument sidesteps this issue and attempts to answer a different question, namely to whom or what are these discursive commitments owed? The answer, I think, is to one another, in a second-personal sense.

Let me begin with the observation that many philosophical accounts of language involve normatively-laden concepts such as licence, commitment, responsibility, and authority. This insight follows from Austin’s (1975 [1962]) demonstration that language is as about doing things like promising and commanding as well as stating propositional facts. It also follows from the late Wittgensteinian conception of language as a rule-governed activity, social and conventional in nature. So it should be unsurprising that a good deal of
philosophical work has been done seeking to articulate the connections between linguistic and moral forms of normativity. I’ll first tease out some of these connections with a discussion promising, before moving on to my main argument, which focuses on assertion.

3.1. Promises and Moral Commitments

I’ll begin with the tempting idea that the linguistic practice of promising (both making and recognising promises) requires promisers and promisees to undertake moral commitments. While I end up rejecting the promise-based approach, it provides an introduction to the structure of the assertion-based approach I pursue in the next subsection.

One way we might try to vindicate moral conclusions from premises about our participation in linguistic practices language is via an account of the felicity conditions of speech acts. Austin called the necessary conditions for particular locutions to count as good illocutions felicity conditions. Austin and his successors were most interested in the felicity conditions for particular speech act types. For example, for my locution ‘You’re hired!’ to successfully appoint someone to a post, I must have the required standing of authority. So having the required position of authority is a felicity condition of the act-type appointing.

According to one view which might be labelled normative pragmatics, at least some speech-act types’ felicity conditions are normative in nature. Let’s use an easy case to

---

98 The idea that promises are inherently normative, indeed that they allow you to bridge the ‘is-ought gap’, has a rich history I won’t be able to fully explore here. Searle’s classic paper ‘How to Derive “Ought” From “Is”’ (1964) argues that the performance of a promise is a counterexample to Hume’s famous dictum that ‘one cannot derive an “ought” from an “is.”’ (Searle 1964, 43) because, Searle argues, making a promise issues a ‘promissory obligation’ (at least to members of the institution of promising).
illustrate the approach. If I promise you that I will meet you at 12 pm tomorrow, I am thereby undertaking a commitment to do so (or at least, a commitment to the fact that I currently intend to do so). I have also entered a normative relationship with you or changed the nature of the existing normative relationship. By promising, I have made it the case that if I fail to meet you tomorrow at 12 pm it would be appropriate for you to blame me or to demand an excuse. What I’ll label a normative theory of promising claims a locution counts as a promise only in virtue of the agent’s having entered these promising-related normative (indeed, moral) statuses and relationships with others. Now imagine that a parrot makes the

99 I don’t intend this example to be an argument for normative accounts of speech-acts (that would take us too far astray)—I merely use it to provide a demonstration of how moral and linguistic normative statuses could be interestingly intertwined. There are other, non-normative accounts of promising which seem equally capable of capturing the parrot case I raise in this paragraph. An account along the lines suggested by Grice would have it that to promise is to assure someone that you’ll do something. And that promissory assurance can ground a normative fact, namely that ‘[the promiser] is under an obligation’ (Grice 1969, 56). See Scanlon (1990, 1998) for an account of this kind (though one which presumes moral obligations not to ‘unfairly manipulate’ others. This non-normative account of the felicity conditions for promises can also handle the parrot case because the parrot’s utterance fails to provide an assurance. There just isn’t the right connection between the parrot’s mental state and its utterance for such an assurance to be provided by the utterance, so we wouldn’t consider the parrot’s the issuance of a promise.

100 This account is a local form of Cuneo’s (2014) normative theory of speech, according to which every illocution has normative felicity conditions like this. But it might not be that the normative account of promising is required. As I’ve mentioned in a previous footnote, the non-normative Gricean account of promising, promises issue assurance (rather than require
same locution. It says, ‘I promise that I’ll meet you at 12 pm tomorrow!’ Has the parrot really promised you anything? No, not plausibly. The normative theory of promising explains the parrot’s failure to promise in the following way. Parrots are incapable of entering into the kinds of normative relationships that speakers do. They do not take on responsibility for the would-be promise to be fulfilled. You cannot rightly blame the parrot if it fails to meet you at 12 pm. And the fact that the parrot hasn’t altered its normative standing, relative to its audience, explains why its locution does not count as a promise.

The mechanics of the normative account of promising aren’t too important. The point is just that on such a view, a locution can only be considered a promise if the promiser is obligated, in some sense, to fulfil the promise. The locution has the social significance of a promise partly in virtue of it shifting the normative landscape in this way. This is the case even if the promiser makes a false promise; they still must be putting themselves ‘on the hook’ in what looks like a moral sense.\footnote{101}

Perhaps, insofar as someone is promising, he is under some kind of moral obligation. Even if this is the case, it doesn’t get us far. The problem is that the kinds of discursive normative felicity conditions. And even taking someone as having issued an assurance grounds a normative fact, namely that the promiser is under certain obligations to the promisee. So it looks like promisers and promisees will fall under normative relations with one another even under such an account.

\footnote{101} This is not to say that if an agent promises to \textit{phi} then he is morally required to \textit{phi}. He might have promised to do an act of great evil, in which case he is morally required to break his promise. The point is merely that his having promised plausibly generates some kind of pro tanto (moral) reason to do as he promises, a reason which may be outweighed in determining the agent’s all-things-considered reasons for action.
commitments we pick up as promisers seem extremely patchy.\footnote{For a brief discussion of the worry (in relation to assertion rather than promising) see Watson (2004). I’ll discuss this worry in more detail shortly.} Does moral normativity really consist merely of doing what you promised? Do we not have any moral obligations outside of our (perhaps implicit) promises? Can we simply avoid morality’s demands by not engaging in the practice of promising? The answer to each of these questions is fairly clear; No. So I don’t think this strategy for vindicating moral norms is a promising one.

3.2. Promises and the Second-Person Standpoint

It seems unlikely that the linguistic practice of promising commits promisers (and promisees) to a significant set of moral norms \textit{directly}. But perhaps those involved in promising necessarily enter into second-personal relations with one another. What then? If Darwall is right that occupying the second-person stand point implicitly commits us to moral norms (at least the morality ‘of what we owe one another’), then it might be that promising \textit{does} ground moral normativity \textit{indirectly} (via the second-person standpoint). This seems like a far more ‘promising’ way of deriving moral commitments and norms from discursive commitments and norms.

As Darwall points out, the normativity which is implicit in acts of promising is distinctly \textit{second-personal} (Darwall 2006, 8). I can only promise \textit{you} something if I take it that you are a second-personally competent being who can hold me liable for my promise. I can’t make a promise to a toaster for this reason (among others). In promising, I also necessarily presuppose that you have the second-personal \textit{authority} to hold me liable for fulfilling or failing to fulfil my promise, and you, as the promisee, to even conceptualise me as having promised, necessarily presuppose that I am second-personally competent in
acknowledging my own locution as a promise. So it looks like we necessarily occupy the second-person standpoint whenever we are in the business of making or acknowledging locutions as promises.

The argumentative structure outlined here, moving from linguistic practices to the second-person standpoint to moral norms, is the one I shall pursue. However, there is a conclusive problem with grounding moral normativity in the linguistic practice of promising which will prevent me from exploring the idea any further. While promising is the speech-act type that most clearly ‘wears its normativity on its sleeve’, as Cuneo says, it doesn’t seem like the right kind of speech-act to make my broader argument go through. There is no apparent necessity for a language user to be a promiser. Promising seems like an optional, avoidable linguistic practice; we might never make promises or acknowledge others’ promises. So, while promising plausibly has the right kind of connection with the second-person standpoint, it doesn’t have the right connection to linguistic capacity more broadly. Let’s now move on to assertion.

3.3. Assertive Commitments as Moral Commitments

Promising offers a helpful model for thinking about how a vindication of moral norms might come out of an analysis of linguistic practices. Let’s now shift focus to the linguistic practice I think actually does ground moral norms—assertion.103

103 By ‘assertion’ I mean a bit more ecumenical than just ‘uttering a declarative sentence’. For instance, Walton points out that ‘Courtroom sketches in newspapers constitute claims about courtroom events’ and I would want to consider them kinds of assertions. (Walton 1990, 82) As will become clear shortly, I’m thinking of assertion along the lines suggested by Brandom (2010) as the basic ‘move’ in the linguistic ‘game of giving and asking for reasons’.
I will adopt a commitment-based approach to thinking about assertion. I want to briefly explain why I think this is a plausible approach, relative to its competitors. Broadly, we can distinguish descriptive ‘Gricean’ accounts of assertion from normative accounts. I won’t provide any kind of comprehensive arguments against the descriptive approach but I will give some account of the appeal of the normative approach, and, in particular, the commitment-based accounts of assertion, which I think are largely compatible.

I’ll begin with the simple observation that assertion, like promising, plausibly involves normative relations including commitment, obligation, and responsibility. To assert that things are one way or another, seems to involve a kind of standing by it, undertaking certain commitments. At least, a significant number of philosophers have argued so. Regarding what these assertoric commitments are commitments to, there is far less agreement. Peirce claims that ‘to assert a proposition is to make oneself responsible for its

104 Descriptive accounts of assertion claim we can understand assertion in terms of the cognitive features of communicators (Stalnaker 1974, 1978), or in terms of what the speaker represents himself as believing in asserting something (Davidson 1984, 268), or in terms of the communicative intentions of the speaker (Grice 1957). Normative and conventional accounts of assertion, on the other hand, understand an utterance’s being an assertion in terms of it meeting conventional standards and thereby having the social significance of an assertion—something that alters the conversational landscape in a distinctive way. While I am not able to provide a comprehensive argument against descriptive accounts of assertion, I will show that normative accounts seem particularly capable of capturing the social dimensions of assertion, which goes some way to justifying the general approach.

105 For a detailed account of how the commitments associated with assertion and promising differ, see Watson (2004).
truth’ (Peirce, 1934, 384). Searle argues that it is a constitutive rule of assertion that one ‘Counts as undertaking to the effect that [the content of some proposition] \( p \) represents an actual state of affairs’ (Searle 1969). Alston similarly claims that a speaker can count as asserting \( p \) if and only if she ‘took responsibility for its being the case that \( p \)’ (Alston 2000, 7). Watson concurs, ‘When you assert that the Earth is older than 5000 years, you commit yourself to something more or less definite, such that you will be correct or incorrect, depending on the age of the earth’ (Watson 2004, 66). Replacing a commitment to truth or correspondence with a commitment to justification, Chisholm and Feehan likewise claim, ‘the concept of assertion is essentially normative. We can explicate it only by reference to justification’ (Chisholm and Feehan 1977, 151). Brandom, whose account I will draw on shortly, offers an inferentialist version of the claim, according to which asserting something involves making a commitment to (1) its fitness as a premise in further inferential reasoning and (2) a responsibility to it being a claim one is entitled to (Brandom 2000, 10).

Putting to one side the specifics of what is undertaken in asserting, all these views converge on what I’ll call a commitment-based theory of assertion. For an agent to count as having asserted anything, he must in some way place himself ‘on the hook’, normatively speaking.\(^{106}\) We cannot interpret a speaker as both having asserted something and simultaneously being normatively ‘off the hook’.\(^{107}\) These thinkers also converge on their primary philosophical interest: figuring out to what asserters are committing themselves, i.e.

\(^{106}\) I’ll disambiguate the senses in which asserters are committed, in asserting, shortly.

\(^{107}\) As Watson notes, ‘To see another as discursively responsible for what she says is not of course to believe her or even to believe that she is sincere. To lie is also to assert (insincerely).’ (Watson 2004, 57 fn. 6) The point is not that we are sincere in assertion but that we are responsible for sincerity (or truth etc.) in assertion.
what feature of the world or of their psychology or of social reality they are tying themselves to in making the assertion.

The obvious first question to ask is whether asserters are implicitly committing themselves to substantive moral norms. If so, we could safely conclude that moral norms are normative for all those who are in the assertion game. Watson can be read as considering this possibility. Assertion, he claims, is governed by two kinds of norm, norms of ‘epistemic competence’ and norms of ‘sincerity’:

If in answer to your question, I tell you that there is a train station at a specific location, I presume my competence to pronounce upon this subject - that I know what a train is, that I am acquainted with the area, or else that I have this information from a reliable source. In addition, my sincerity is presupposed; I am giving the other to believe not only that I am competent and well positioned to comment on these matters but that I believe what I say. (Watson 2004, 59)

Let’s focus on the norms of sincerity. While the norms of sincerity play an epistemic function (‘the credibility that is crucial to conversational practice depends on both competence and sincerity’), they are primarily moral norms (59). If we are necessarily committed to moral norms of sincerity insofar as we are in the business of assertion, then perhaps we can vindicate moral normativity more generally.

Unfortunately for the linguistic constitutivist, it seems unlikely that this can be accomplished, or at least, not directly. As Watson admits:

[T]hat assertion intersects with morality in this way doesn't by itself show that discursive responsibility is moral responsibility. Arguably, moral commitment requires a larger background. The compliance with moral requirements by an individual who assumed only assertoric responsibility (if that is conceivable) would be too partial and intermittent to be usefully regarded as moral. I would like eventually to get to this question, but here I shall have to leave it open. (Watson 2004, 59-60)
While Watson’s discussion is limited, his worry is clear enough. It’s not plausible that the totality of what we think of as moral normativity, nor even a substantial portion of it, can be derived from the norms of discursive sincerity. Commitment to these norms might vindicate the normativity of some moral reasons, but it seems incapable of delivering many. That’s why an agent’s commitment only to ‘assertoric responsibility’ would be ‘too intermittent’ for the agent to be considered committed to moral normativity generally.

3.4. Assertive Commitments and Responsibilities as Second-Personal

Despite his scepticism of assertive responsibility considered as the ground of moral responsibility, I think there is something of extreme value in Watson’s account. I am referring to the ideas Watson puts forth in the following passage:

[T]here is something assertion-like at the core of the ethically central cases. Responsible agency requires that one stand for something, and this is something more basic than the subjection to substantive moral requirements […] an analysis of conversational responsibility is illuminating because it has little or nothing to do with the features that are often taken as essential to the moral case: namely the propriety of reactive attitudes of the retributive variety, or something similar. This is salutary, I think, because it highlights the possibility that those sentiments and attitudes are secondary even in the moral case. What is crucial are the ideas of commitment and answerability. (Watson 2004, 60)

While Watson’s discussion ends there, I think I can fill in the picture by providing an explanation of the relation between assertion and the basic moral (ethical in Watson’s terms) position of standing for something. As Darwall argues, the ideas of commitment and answerability (understood second-personally) are indeed a suitable bedrock to ground substantive moral norms and reasons. And it is, as Watson suspects, something like the attitude of standing for is required for commitment and answerability. We find it at the heart
of the explanation of our linguistic capacities, and I think we also find it at the heart of our explanation of our moral capacities.

Now, why don’t I agree with Watson’s negative conclusion that assertoric responsibility is too ‘partial and intermittent’ to be considered the ground of moral normativity? I think this negative conclusion stems from looking in the wrong place for the connection between assertive commitments and moral ones. Rather than focusing on what assertoric commitments and responsibilities are to, the content of the commitments, I think we would get further by asking to whom or what they are owed. As Darwall notices, ‘The moral sense of “responsible for” is conceptually tied to “responsible to” (whether to individuals or to one another as members of the moral community)’ (Darwall 2006, 68). As was the case with the brief discussion of the commitments associated with promising, I find it most plausible that assertoric commitments and responsibilities are second-personal in nature. Therefore, in making or acknowledging assertions, we occupy the second-person standpoint. And, if Darwall is right, then the presuppositions of this standpoint entail the basic materials of morality, i.e. equal accountability and autonomy.

3.4.1. The Basic Idea

To whom or what are these assertoric rights and responsibilities, however understood, owed? The most obvious and plausible answer is that we owe these commitments to one another. I assert ‘There is a bomb on this bus’ only if I am thereby ‘liable to correction, blame, or reproach if things are not as I present them.’ (Cuneo 2014, 35) But all liabilities are owed to something or someone. In the bus case, I am liable to my addressees. And in taking myself as liable to them, I necessarily presuppose that my addressees are the kinds of beings who have the kind of second-personal authority (and therefore competence) to hold me liable. Indeed, members of my audience can only take me to have made an assertion if they think
that I am liable to them. By making my assertion, and by taking my location to be an assertion, we all must enter into what Darwall calls the ‘reciprocal address’ of the second-person. So assertion inescapably involves occupying the second-person standpoint.

To someone like Darwall who claims that commitment and accountability can only really be understood second-personally, this is an obvious point. The kind of mutual, intersubjective accountability that we find in assertive discursive practice has the ‘I-you-me structure of reciprocal address’ which Darwall describes as the mark of the second-personal and the necessary presuppositions of which ground a contractualist set of moral norms, since, as Darwall argues:

Whenever we take up the second-person standpoint at all and address any claim or demand whatsoever, we are committed to the shared second-personal authority of any second-personally competent being (moral agent or person) on which the ideas of moral obligation and rights depend. (Darwall 2010, 224).

3.4.2. Is Assertoric Responsibility Too Partial?

With the basic idea in mind, let’s consider a reformulated version of Watson’s worry about deriving moral normativity from assertoric normativity. Watson’s original worry is that an account of moral normativity grounded in the commitments we make when we assert would leave ‘compliance with moral requirements […] too partial and intermittent’ because the moral commitments and obligations we accrue as asserters are too few (Watson 2004, 60). I’ve suggested an alternative way of moving from assertoric norms to moral ones, by focusing not on what we are committed to when we assert, but to whom we are committed insofar as we can be understood as players in the game of assertion. But doesn’t this mean that we can opt out from moral normativity, and avoid taking on moral obligations, by simply refusing to talk to certain people? If asserting something to someone, or taking someone as
having asserted something, involves taking on moral obligations, then can’t we just not assert with those with whom we don’t want to be morally entangled? If so, then the second-person standpoint we necessarily occupy when we’re involved in making and acknowledging assertions would indeed be too partial and intermittent to ground moral normativity.

I think this objection is unconvincing because we can’t be selective like this when we make assertions. Let’s take an example. What if I were not heard by my intended audience but were instead overheard by someone else on the bus, by an unintended hearer? Let’s say I have not consciously sought to enter a second-personal mode of address with this unintended hearer, and yet I have still asserted that there is a bomb on the bus. And the over-hearer would rightfully take my locution to be an assertion, even while recognising that he wasn’t the intended audience. So what does this mean? Must the over-hearer enter second-personal proto-moral relations of mutual accountability with me? Need I not reciprocate? And, more broadly, isn’t this a counterexample to the idea that the assertoric game necessarily involves occupying the second-person standpoint?

Here’s my understanding of what’s going on in the case where someone overhears an assertion not intended for their uptake. I still have asserted something. My over-hearer is still justified in taking me to have asserted something. And I also still seem assertorically responsible in the full second-personal sense, even to the unintended hearer. As evidence of this, say the over-hearer jumped out of the window to escape the bomb on the bus, only to find I was lying. He would be entitled to blame me for my assertion if it were insincere. And, following Darwall, I understand taking on the attitude of blame toward someone as requiring the blamer to relate to the blamed in a distinctly second-personal way. It is not that ‘X is blamed’ and you happen to be X, it is that you blame me. You are holding me accountable for failing to act on my second-personal reasons to assert sincerely. When the over-hearer acknowledges my locution as an assertion, he and I enter into second-personal relations of
mutual accountability etc. Hence, in asserting, I should take myself as being second-personally liable to a wider audience than just my *intended* audience. The commitments we make in asserting render us liable to *anyone* capable of recognising our locutions as assertions. This means that we can’t be partial and selective about who we are assertorically committed to in the problematic way the objection picks out.\(^{108}\)

3.4.3. How Far Do Assertoric Responsibilities Reach?

The prior discussion raises an important question; if asserters are second-personally liable to people other than their immediate audience, just how far does assertoric responsibility reach? I have suggested that it reaches from the asserter to everyone who can acknowledge the assertion *as an assertion*—every competent member of the linguistic community. Let me explain why this is with a clearer case than that of the over-hearer.\(^{109}\)

Imagine the incident were reported in the newspaper and readers learned I had lied. The newspaper readers can acknowledge my reported locution as an assertion. So, the objection goes, it can’t be that being in the game of assertion necessarily requires that we occupy the second-person standpoint. The readers needn’t take themselves to have a second-personal relationship with me, in taking my locution to be an assertion, and I needn’t take myself to be in a second-personal relationship with them, even though I recognise that they acknowledge my locution as an assertion. So the objection is that I needn’t have a second-

---

\(^{108}\) Darwall makes a similar move in moving from specific instances of second-personal interaction to a more general recognition of all persons as potential representatives of the moral community (Darwall 2006, esp. 66-70).

\(^{109}\) The over-hearer may be too directly involved in the bus situation to test the right intuitions.
personal relationship with the newspaper readers. If I did have this second-personal relationship with the newspaper readers, then it would be fitting for the newspaper readers to adopt distinctly second-personal reactive attitudes such as blame toward me (if what I asserted were insincere and false, a lie). But I don’t think that blame would be an entirely fitting attitude for the newspaper readers to hold toward me. They might think poorly of me, be indignant toward me, judge me responsible, etc., but blame would be a slightly odd reaction. Or, if they did blame me, we would probably be inclined to understand them as blaming me on behalf of the bus riders I directly misled and terrified, because they have the right kind of direct involvement with me and my performance of the assertion for attitudes of blame to be fitting. If we do think the newspaper readers are blaming me, we wouldn’t think the newspaper readers are blaming me on their own behalf. Doesn’t this show that being in the game of assertion does not involve taking on second-personal relations?

To understand why this objection fails, we need to invoke the distinction Strawson makes between two kinds of reactive attitude—‘participant’ and ‘impersonal’ (Strawson 1968, 74). This is not the more famous Strawsonian distinction between the involved and the objective attitudes we can take toward people’s behaviour depending on whether we view them as responsible agents or not, but a distinction within the reactive attitudes we take when we judge them to be acting as responsible agents. Strawson describes participant reactive attitudes are those typically taken up by the agents directly involved in the transaction—those who make and recognise the assertion ‘There is a bomb on this bus’, in this case, such as gratitude (if the warning saves their lives), blame (if it’s a hoax), and resentment (likewise). These kinds of reactive attitudes, I have just claimed, are second-personal in

---

110 As I briefly mentioned, you might also have these attitudes ‘on behalf’ of others, by taking their place in your imagination.
nature. But what attitudes might the newspaper reader me entitled to take toward me and my having made an assertion that there is a bomb on the bus? They might take up *impersonal* reactive attitudes. These are ‘impersonal or disinterested or generalized analogues’ of participant reactive attitudes, felt not from the perspective of those directly involved but as though from an impartial point of view (Strawson 1968, 72). If my assertion is a lie, it would be fitting for my intended audience on the bus to, e.g. blame me—a participant reactive attitude. But it would be fitting for just about anyone who came to understand me to have made that assertion, even the newspaper readers, to feel indignation toward me—an impersonal reactive attitude.\[^{111}\]

As Darwall points out, impersonal indignation and participant-based blame are deeply connected:

> If we come to believe that someone does not deserve blame, say, because he could not possible have known the true character of what he was doing because he was under extreme duress, then this will reduce or even defeat our indignation toward him. (Darwall 2006, 68)

Why do impersonal and participant reactive attitudes have these kinds of close connections? Why is their fittingness dependent on the same factors? I think Darwall gets the right answer when he claims that even impersonal reactive attitudes are second-personal in the relevant sense. They involve the same kinds of second-personal commitments, responsibilities, and obligations, and have the same defeating and mitigating conditions. Hence, claims Darwall, *all*:

\[^{111}\] The attitude of guilt can also be understood as the first-personal form of an impersonal reactive attitude—a kind of internalisation of the reactive attitudes we might take as representatives of the standpoint of the moral community (Darwall 2006, 67).
reactive attitudes [even impersonal ones] are always implicitly second-personal and they thereby invariably carry presuppositions of second-personal address about competence and authority of the individuals who are their targets, as well as about those who have them. (Darwall 2006, 67).

The distinction between participant and impersonal reactive attitudes lies not in whether or not they are second-personal but in how they present to us. We feel participant reactive attitudes ‘as if from the second-person standpoint of a relevant transagent [i.e. an agent directly involved in a transaction]’ (67). So the second-personal nature of such attitudes is more transparent. We feel impersonal reactive attitudes, on the other hand, ‘from the standpoint of members of the moral community’ (67). But, although they feel different, impersonal reactive attitudes carry exactly the same second-personal presuppositions about agents’ mutual competence, authority, and accountability. The newspaper reader can only understand me as having lied (insincerely and falsely asserted) that there is a bomb on the bus if he thinks I have really committed myself to something, in asserting it. And that requires that I am second-personally competent and in relations of mutual authority with others, etc. So while it may not be fitting for the newspaper-readers to adopt participant reactive attitudes such as blame, these are not the only responses which involve occupying and taking on the presuppositions of the second-person standpoint. Impersonal reactive attitudes such as indignation do too. And indignation is a fitting attitude for the newspaper readers to have. The newspaper readers, in recognising my locution as an assertion, really must be occupying the second-person standpoint and relating to me, as the asserter, in a second-personal way.

Let’s put all these points together. In making an assertion, as Watson and others note, we necessarily stand for whatever we assert. We can’t understand someone as having asserted something without thinking they have committed themselves to something. This is what I labelled the commitment view of assertion. But our capacity to generate this commitment requires that we occupy the second-person standpoint (with others). It requires
that we take those who can recognise our locution as an assertion as second-personally competent—they must be if they are to recognise us as having altered the normative landscape, as having taken on a commitment, by having made an assertion. We must also recognise these other players in the assertoric game as having a kind of authority over us, in that they are entitled to hold us responsible for what we assert. Otherwise, we wouldn’t have really committed ourselves to anything. So being in the business of assertion involves occupying the second-person standpoint and all of the related (proto-moral) presuppositions of that standpoint.

I also have argued that we do not merely commit ourselves to our (intended) audience when we assert. In asserting, we are issuing a redeemable token, a general licence, to all those who are assertorically capable, all those who can recognise our assertion as an assertion, to hold us liable via impersonal reactive attitudes and other forms of sanction. This seems like a point of significant difference between promises and assertions. Assertions can be thought of as attempts to update a public ledger of claims, and hence they generate general forms of liability in a way promises don’t. If I promise you I’ll do something, I’m only responsible to you for that promise. An over-hearer isn’t entitled to hold me accountable in a second-personal sense, at least not in the same way you are. If I make a false promise, an over-hearer might be annoyed that my promise has misled them, but I’ve not broken my promise to them. I am responsible for my promises only to those I have promised. Hence, the promisee is in a very special position of authority to hold me responsible for keeping my promise. But I must see myself as second-personally responsible to anyone who could recognise my assertion as an assertion whenever I assert something. In asserting, however, I am standing by something

112 I shall shortly argue this extends to all linguistically competent beings.
in a way that issues a liability that every member of the linguistic community can collect on.\(^{113}\)

To conclude this section, I have argued that making and recognising assertions involves occupying the second-person standpoint—acknowledging and making second-personal claims and taking on second-personal commitments and responsibilities.

3.5. Commitments: Social or Volitional?

In Chapter One, I distinguished two senses in which a person can be committed to something. There are social commitments and volitional commitments (in Calhoun’s terminology, *attitudinal* commitments (Calhoun 2018, 93)). In what sense are those involved in the game of assertion ‘committed’ to the moral norms of the second-person standpoint? Are they simply *taken to be answerable* by the rest of us for failing to adhere to these norms? Or do they also undertake a commitment *themselves* to take these second-personal demands as having real weight? I shall argue for the latter.

3.5.1. Two Senses of Being Committed

I want to return to Watson’s observation that in asserting, we necessarily are *standing for* something. We are committed to something, we are responsible for something. ‘To assert (or attempt to assert), “I am not a responsible being”, would be self-refusing in the sense that it would demonstrate the very capacities that are denied.’ (Watson 2004, 60). I have claimed that we *stand for* things, when we assert, in a second-personal sense. In asserting, we are

\(^{113}\) I recognise that this is a strong claim. I defend it against possible objections in section six of this chapter.
issuing a liability to competent players in the assertion game. But what kind of thing is this attitude, this *standing for*? Is it merely recognising oneself to be performing actions which have statuses in the assertion game? Or does it mean adopting a commitment as one’s own such that one can’t rationally shrug one’s shoulders?

To be committed to something merely in the social sense, as I’ll call it, does not require the possession of any pro-attitude toward the object of the commitment, nor a disposition to act in a certain way. It merely means that the committed individual has the *social* status of being committed, according to the rules of some practice which includes the status ‘A is committed (to X)’. This is the sense in which Ridge (2019) thinks chess players are necessarily *committed* to winning by checkmate. It might seem that making and acknowledging assertions only requires us to undertake commitments in this sense.\(^{114}\)

I think there is a worry, for all Darwall argues and for all I have argued so far, that the sense in which asserters are necessarily committed to the demands and norms of the second-person standpoint is merely social, and hence not capable of providing a robust vindication of the normativity of second-personal reasons. Over the course of this thesis, I have explained

\(^{114}\) That all asserters are committed to second-personal normativity in the social sense might be enough to yield some interesting conclusions for a norm-expressivist like Gibbard. The idea might be that in making a move in the game of assertion, as asserters or acknowledgers, we necessarily endorse of the second-personal norms involved in assertion. We could also follow Brandom in thinking of action as a ‘discursive exit transition’ (Brandom 1998, 130) from deontic scorekeeping, in which case our actions are presumably subject to the same socially generated norms as the rest of the items on our deontic scorecards. And the rules of inference that govern transitions from deontic scorecards to actions may include moral restrictions. This is another sense in which agents might be understood as ‘committed’ to moral norms (in the social sense of commitment).
why I think Humeanism is a promising foundational account of normative reasons for constitutivism to build upon. But Humeanism says that vindications of the normativity of any reason will ultimately have to refer to features of the conative psychologies of agents. Therefore, for constitutive commitments to be of use to the Humean, they had better not be merely social commitments.

If they were merely social commitments, the Humean moral sceptic could respond, ‘Sure I am committed—in the social sense—to moral norms, people will hold me to account for failing to meet up to them, but I don’t care! What does that have to do with what I ought to do?’ Or, to put the worry another way, if assertive commitments are merely social then our second-personal responsibility and accountability, our reasons more generally, could be viewed as providing only ‘low grade’ normativity. Think of etiquette. An agent can recognise they are ‘committed’ to the norms of etiquette in that they are at a dinner party and their behaviour has etiquette-related statuses, such as being ‘poor manners’. But the agent can still claim that there is no rational pressure on him to accept these norms, to give them any weight, or significance, in deciding what he ought to do. To make the reasons of etiquette genuinely normative reasons for the agent, we need him to have adopted some kind of volitional, attitudinal commitment that would give the demands of etiquette some rational weight for him. And we need the same kind of thing to vindicate the normativity of the (moral) norms of assertion.

The volitional/attitudinal sense in which someone can be committed to something, on the other hand, is that their psychology is such that we count them as having a commitment to some end, state of affairs, way of behaving or reasoning, etc. This is the sense of commitment that the Humean thinks has normative significance in deciding what the agent ought to do. So, to answer the Humean moral sceptic, we need to explain those involved in assertion are necessarily committed in the attitudinal, volitional sense.
The question, of course, is whether assertion is like this. Does earning the status of a player in the assertion game, a being whose locutions can be recognised as assertions, a being who can be expected to acknowledge locutions as assertions, require undertaking commitments in the volitional, attitudinal sense?

3.5.2. The Volitional Commitments of Assertion

I think the practice of assertion requires its participants to undertake commitments in both the social and volitional senses. In making an assertion, we commit ourselves in some way to the content of the assertion. This is a commitment merely in the social sense (which is why making insincere assertions is perfectly ordinary). We don’t have to be committed ourselves to p being the case in order to assert p. This is why lying is possible. But we can’t be understood as having asserted p without being taken to be committed, in the social sense, to p being the case, or being responsible for justifying our entitlement to assert p, or whatever the case may be.

However, his isn’t the entire story of assertoric commitments. In counting as a participant in the practice of assertion, a practice in which making assertions is just one possible move (along with, e.g. acknowledging assertions), the price of entry includes undertaking a commitment to give second-personal reasons and claims some weight, some significance, in one’s rational deliberation. Without taking on this commitment, one wouldn’t, I think, be correctly counted as a fellow player in the assertion game.

Let me make first a few brief remarks on volitional constitutive commitments before discussing the assertoric version. In Chapter One, I argued that some activities have constitutive commitments in the volitional sense. To count as a participant in these activities
requires adopting certain volitional attitudes, such as aims, ends, intentions, or goals. More broadly, an individual’s participation in a joint enterprise often seems to necessitate the adoption of commitments. Participation involves recognition of the normative statuses of the enterprise. This, in joint activities, often requires recognising other players (and oneself) as having certain authorities within the practice, and hence recognising that their authoritative practice-related claims have some weight. For instance, playing chess involves recognising the authority of the other player to perform legal moves when it is his turn, moves which can generate game states such as checkmate. In On Social Facts, Gilbert provides the following more general account according to which joint activities involve, in my terms, undertaking constitutive commitments:

However much we may feel forced into participating in conversations [and other joint/social activities] we do put into them, if we enter them at all. We put ourselves into them, in a sense. This is what we have to do, to become members of a plural subject or collectivity. […] This entails taking on or accepting a set of responsibilities and rights: it involves recognizing a new set of constraints on one’s behaviour. (One also accepts certain new entitlements.) The rights and responsibilities most directly in question are neither moral nor legal. They could be referred to as ‘associational’. (Gilbert 1992, 410-1)

Gilbert does not specify whether these associational commitments we must ‘take on’ to be participants in the joint activity of conversation are social or volitional or both. But they are most naturally read as at least partly volitional. It’s not simply that we acknowledge or recognise that our behaviour is permitted or not, according to these associational rights and responsibilities, just as we might acknowledge or recognise that our behaviour is permitted or

---

115 See the earlier chapter on the shmagency objection for my arguments for this view. I claimed that chess and boxing were examples of practices that have constitutive commitments in the attitudinal or volitional sense.
not according to the standards of etiquette. Rather, it’s that we put ourselves into these associational commitments, that we accept them, recognising them not as arbitrary standards but as actual constraints on one’s behaviour. If Gilbert is right that this is what is required, then it certainly sounds like a volitional sense of commitment, something with real rational bite—something we are forced, rationally speaking, to assign some weight to in deciding what to do.

I think this is also what is involved in participating in the ‘game’ of assertion. We have to commit in a volitional sense to the normativity of the second-person standpoint. We have to adopt a general commitment to giving the (potential) second-personal claims of ourselves and others weight in deciding what we are to do.116

But why should we think that we need to adopt an attitude in order to count as players in the game of assertion? I have characterised making and acknowledging assertions as a social practice involving (second-personal) statuses such as commitment, responsibility, and answerability. So participants in the practice of assertion must understand one another as second-personally responsible and answerable to one another. In other words, making and acknowledging assertions involves associating with other players in the same activity, and associating with them as a player among other players.117 And we can’t understand someone

116 To put the point in Bratman’s terminology, we have to adopt a self-governing policy in order to enter into the second-personal relations that make assertion possible. I’ll soon explain why I think this is a helpful way of understanding these commitments.

117 This point may have some parallels with Korsgaard’s idea that reasons are ‘public’ rather than ‘private’, in what she took to be a Wittgensteinian sense. (Korsgaard 1996, 131-45). For critical discussion of Korsgaard’s argument, see Coleman 2005, Gert 2002, Hurley 2001,
as acknowledging their position as a player among other players without understanding them as assigning some weight, some significance, to the second-personal claims and reasons that infuse assertoric practice. This is why counting as a participant in assertion, as a potential maker and ackowledger of assertions, requires the undertaking of, in Bratman’s phrase, a ‘commitment to significance or weight in deliberation.’

Consider a person who lacked this commitment to give these second-personal considerations some deliberative weight, someone who assigns no weight to second-personal reasons and claims, and therefore does not recognise second-personal responsibility or authority. Imagine this person makes a locution that, in the mouth of an ordinary speaker, we would take as an assertion. To be clear, it’s fine that this person has no commitment to the content of what they assert. As I’ve said, to make an assertion does require sincerity. The strangeness of this person’s locution stems from their incapacity to undertake the necessary commitment to give second-personal claims some weight. If he doesn’t place any weight on second-personal reasons and claims, then he can’t see other players in the assertion practice as having the second-personal authority to attach responsibility for the claim to him, to hold him responsible for it. So he can’t view himself as second-personally responsible. He can’t view himself as the appropriate object of reactive attitudes such as blame, etc. if what he says turns out to be insincere, unjustifiable, and/or false. Absent all this, is this (admittedly bizarre) character someone we would be inclined to assign the status of a fellow player in the game of assertion, of making and acknowledging claims to one another? My intuition is he is not. I don’t think our practices are such that we would include him as a fellow player in the

LeBar 2001, and Wallace 2009. I don’t believe these criticisms apply to the point I am making here.
assertion game. We wouldn’t acknowledge his locutions as assertions, and we wouldn’t expect him to acknowledge our assertions as assertions. He wouldn’t be a fellow player.

Intuitions are admittedly tricky here because almost everyone we imagine interacting with is a player in the assertion game and is committed to taking the second-personal claims of others to have some kind of weight (I shall argue that all linguistic beings and, indeed, all agents are). But there are beings that generate utterances which systematically fail to count as assertions. I am thinking of the ‘declarative’ utterances of very young children, robots, and pets. The utterance of someone who lacks any commitment to giving weight to second-personal claims and statuses seems much closer to a parrot’s, in their social significance, than that of a typical speaker.\(^\text{118}\)

But maybe this isn’t strong evidence that constitutive commitments must be volitional rather than merely social. The parrot is committed in neither sense. Let’s stipulate that this uncommitted utterer, presumably unlike the parrot, is taken to be committed by other

\(^{118}\) I do not think it is necessary to claim things are either black or white here. As Brandom points out, while a parrot is a clear non-participant in the assertion game, children may be a middle case. ‘As we grant socially the sorts of responsibility and authority characteristic of adulthood only gradually and in proportion to a child’s mastery of the demands and the skills required to fulfil those demands, so it is with assertion. According to our conventions, the utterance of a declarative sentence […] claims for itself the social status of an assertion. But we accord this status to the utterances of children, madmen, and foreigners only to the degree to which we take them to understand the inferential relations within which the responsibility and authority of asserting arise.’ (Brandom 1983, 644) In relation to my own claims, it may be that we can recognise people as partial players in the assertion game only insofar as we interpret them as having undertaken a general commitment to give some weight to the second-personal claims that underlie assertion.
language users. He is taken to be committed in merely a social sense. To put it in a more Watsonian way, if what he says is unwarranted that would, socially, licence sanction. Or, to use Brandom’s gloss on assertoric commitment, this person falls under a ‘justificatory responsibility’ in the thin sense that his locution might be taken as a ‘move’ in a social practice that generates a responsibility to justify his entitlement to the assertion. But like the sceptic about the normative robustness of etiquette, he doesn’t care about the norms and sanctions of this social practice. He doesn’t care at all about whether or not he can deliver on that justificatory responsibility. He is not himself committed, at all, to the norms of assertion or the inferential rules. He thinks it has no bearing on what he ought to do, rationally speaking. Perhaps it’s somewhat more plausible that such a person is a player in the assertion game. At least, it’s more plausible that his locutions count as assertions than it is that a parrot’s do.

Perhaps we can even go one stage further, and imagine a character who adopts a kind of fictionalist attitude toward the norms presupposed by assertion, the kind of commitment to the scenario and rules one undertakes when playing make-believe games such as Dungeons and Dragons, where we proceed as if those norms and rules are in force, even though he doesn’t really accept or endorse them in the full volitional sense. Would we not count the speaker who is committed in the merely social sense, or the speaker who is committed in the fictionalist sense, as genuinely asserting anything?

I think both these characterisations of the commitments that underlie assertoric practice leave something vital out—namely why we grant the social significance of ‘assertion’ to some locutions and not others, why we count some speakers as claiming things and others as not claiming anything. I think it leaves out the volitional element of assertoric commitment. The associational commitment we adopt as our own, as makers and acknowledgers of assertions, is not just typical of assertive practice. It seems essential for
explaining how distinctive features and normative statuses of assertion are possible. For this reason, the adoption of these commitments is partly constitutive of the practice of assertion.

Consider the fittingness of the reactive attitudes that are ubiquitous in assertoric practice. Someone might assert something unwarranted, such as a deliberate lie, receive some sanction in the form of a reactive attitude, and protest ‘But I undertook no commitment! I am blameless!’ I think we have two options here. We might accept the claim at face value, in which case we would have to reassess our initial acknowledgement of his locution as an assertion, a move seeking to put forth a claim for which the asserter is responsible (for justification, truth, or whatever), or we can continue to acknowledge the speaker as an asserter and reject his excuse.

I’ve already described how asserting something necessarily renders one liable to correction, which will often take the form of second-personal reactive attitudes such as blame and indignation. Such reactive attitudes are an important mechanism through which the norms of assertion are enforced. But to blame someone or be indignant toward them is to judge them as having acted irrationally. We don’t simultaneously think someone has acted as they ought to, rationally speaking, and that they are blameworthy. So, for an asserter’s assertion to have the social significance of an assertion, there must be something volitional in

119 I take this to be the orthodox view. For the unorthodox view, see Williams (1981). Williams argues that our ethical practices ought not to make this assumption that whenever someone has failed ethically, they have failed rationally. A man who beats his wife may be ‘inconsiderate, or cruel, or imprudent […and we needn’t additionally claim] what is particularly wrong with the agent is that he is irrational.’ (Williams 1981, 86) But even Williams thought that blame was a distinctly moral reaction (rather than an ethical one) and built-in these (what he thought were false) moral rationalist assumptions.
his psychology that explains why blame is *prima facie* fitting. There must be something in his psychology that explains why he ought not to have made the claim he did. The existence of a volitional commitment, necessarily undertaken by all asserters, would neatly explain why our reactive attitudes in the context of assertion make sense. By contrast, there is no obvious rational failing involved when someone committed to the norms of assertion *merely in the social sense*, or even *merely in the fictionalist sense* makes an utterance whose content is insincere and false. They aren’t committed *volitionally* to anything which would make such an act irrational.

In sum, to interpret someone as having asserted something involves seeing them as potentially fitting targets of certain reactive attitudes. And these reactive attitudes, in turn, are only fitting if the asserter is subject to rational standards. So there must be some kind of volitional element in the asserter, something that explains why the utterer would be rationally failing by failing to act in accordance with the demands of second-personal reasons, responsibility, and accountability. This is why to count as *in the game of assertion* requires undertaking a commitment to give the second personal some weight in deciding what we do.

3.5.3. Volitional Commitments to Weight in Deliberation

What kind of volitional commitments are these? I have claimed that our assertive practices, and particularly the distinctive ways we sanction unwarranted assertions, presuppose that those in the assertion game have some volitional commitment to give weight, some significance in one’s rational deliberations, to the second-personal reasons, claims, demands, and statuses that assertion involves. This distinctive commitment to weights is a
species, I think, of what Bratman variably describes as ‘personal policies’, ‘self-governing policies’, or ‘commitments to significance or weight in deliberation’. To Bratman develops the idea of self-governing policies as part of his account of rationality under cognitive and time constraints—what he calls ‘bounded rationality’. To I discuss constitutive commitments as self-governing policies in Chapter Three. Most theories of practical rationality are implicitly unbounded. Take a simple rational choice theory. According to this theory, rational agents assign utility values to various possible outcomes of their actions. They then seek to maximise expected utility by selecting the action with the highest expected utility of any of the options. Rational choice theory is generally understood as a descriptive theory; it says we can interpret rational beings as though they are maximising expected utility. But rational choice theory can also be interpreted as a prescriptive, normative theory of practical rationality; on this interpretation, rational choice theory says that rational beings should maximise expected utility whenever they act. But cognition and time are both finite resources. As critics of maximising forms of utilitarianism have long pointed out, aiming to select the optimal choice for every decision we make is extremely difficult, if not impossible. We might do better, that is, generate more wellbeing, if instead of seeking to optimise, we instead merely satisfice (seek a satisfactory solution) or follow general moral rules. A similar argument can be given for practical rationality generally. For earthly creatures like the writer, cognition is a strictly limited resource. As finite beings, we need to be careful with our cognitive purse strings. It does no good to think so intensely about whether or not to reply to that email you’ve been putting off that you forget to look both ways before you cross the road. The commonest strategy for coping with our cognitive limitations is to distribute the cognitive load over a period of time—to think about something for a while. In other words, we can trade time for cognition, which goes some distance to mitigating our limitations on cognition. But time is itself a finite resource. Misuse of time creates its own
deal with our cognitive finitude and our temporal finitude, the fact that we can’t stop the world and take the time to consider all options, we make plans. Plans do the deciding for our future selves. Our future selves can, of course, decide to reconsider plans and abandon them. But the default is that we follow our plans (Bratman 2000, 40). While some plans are highly specific, aimed at realising a particular desired end, many of our plans are quite general. They structure how we live in a recurring fashion. An example might be buying your morning coffee or, even more generally, abstaining from caffeine. Bratman labels these kinds of general plans ‘policies’ (41).

On Bratman’s account, self-governing policies are, like all policies, intentions that are ‘appropriately general’ (2004, 32). What is special about them is that they are policies ‘of giving weight or other forms of significance to certain considerations in practical reasoning and action’ (32-3). So the content of a self-governing policy will be of the form, ‘in one’s relevant practical reasoning, treat C as having this weight.’ (33) Hence, the self-governing policy ‘is a commitment’ (in the volitional sense) to C having that weight in one’s practical costs, whether absolute or in terms of lost opportunity. It can often be better to act at T1 without having considered every option than to consider every option and then act at T2. One’s options might have deteriorated greatly in the meantime. Consider the hesitant romantic, unsure whether to lean in for the kiss and missing the optimal moment, caught up in his ongoing deliberation.

Bratman takes intentions to be conative, desire-like states of mind, albeit special ones that decide what to do in a way desires don’t and are subject to greater demands of consistency. So all this fits nicely with a Humean, desire-based theory of reasons.
deliberations. Unlike merely social commitments, the adoption of a self-governing policy has a rational bite.\textsuperscript{123}

But why are constitutive commitments to weights general commitments to weights? Why isn’t it good enough, to count as an asserter, that we sometimes, depending on the situation, happen to undertake a temporary, specific commitment to assign some weight to the second-personal considerations that assertoric practice involves, without undertaking this general commitment to do so all the time? I think this is the case because, as Bratman observes, adopting policies (general plans of action that resist rational reconsideration to some extent) enables forms of social coordination that would otherwise be impossible. And the assertion game is one of the most complicated social practices there is. Only if I am understood to have a general commitment to be responsive to the second-personal claims that assertion involves can other players in the assertion game have the generic expectation for me to uptake their locutions as assertions. Only if I am committed in this way are they entitled to take my assertion-like locutions as assertions. We can only acknowledge assertions as assertions by default because we adopt this general policy and because we presuppose that everyone else has too. Hence, the adoption of a self-governing policy, a general policy of commitment to giving second-personal considerations weight, seems to enable those in the

\textsuperscript{123} The adoption of such policies arguably plays an important role in making autonomous action possible by helping to manage how we deal with the ‘non-angelic’ parts of our ‘psychic economies’ and coordinate our plans over time. The fact I will assign weight W to C will permit me to generate plans which rely on that self-governing policy. See Bratman 2004.
game of assertion to reliably be recognised and recognise their fellow players as players and for them to respond appropriately.\textsuperscript{124}

In sum, I’ve claimed that the adoption of a particular commitment is partly constitutive of being a player in the game of assertion and that we should understand this constitutive commitment as the self-governing policy of assigning some significance, some weight, to second-personal reasons and claims in their practical deliberation. Participants in the game of assertion, as a constitutive matter, are committed to contractualist moral norms and to taking the (moral) demands of the second-person standpoint seriously.

4. Assertion and Language

I have argued that being a player in the practice of assertion involves (1) entering relations of mutual accountability and answerability with other persons and hence occupying the second-person standpoint, and (2) adopting a volitional commitment to give second-personal considerations some weight in one’s practical deliberation. But it is not merely assertion that requires (1) and (2). It is language use generally. At least, this is the case if assertion is ‘downtown’, in Brandom’s terms: if assertion distinguishes the linguistic from the non-linguistic. The significance is this. If all language users necessarily occupy the second-person standpoint (because all language users are necessarily asserters), then all language users are subject to all the same substantive moral norms and have all the same normative commitments.

\textsuperscript{124} In any case, the generality of these commitments is not essential to my wider argument. All that my argument requires is that those in the game of assertion necessarily adopt commitments to give the demands of the second-person standpoint some weight in deliberation—that the practice of making and acknowledging assertions wouldn’t otherwise be possible.
reasons as asserters. This yields the view I call linguistic constitutivism. Certain moral conclusions follow from the constitutive commitments of language users.

I’ll first explain the plausibility of assertion being the foundational linguistic practice and hence why it is plausible that competent language users are, necessarily, involved in the ‘game’ of making and acknowledging assertions, drawing on the work of Brandom.\(^{125}\)

4.1. Inferentialism

To understand why Brandom thinks assertion is the fundamental linguistic practice, the practice that distinguishes linguistic from non-linguistic practices, and ‘concept-mongering’ practices from non-‘concept-mongering’ practices, it will be helpful to have a rough idea of his broader philosophical commitments in mind. With roots in the work of (among others) Wittgenstein, Pierce, and Sellars, inferentialism is most clearly understood in contrast with representationalist theories of language. In brief, the latter understands linguistic meaning in terms of representation. We discover things and ideas and we make words that stand for them.\(^{126}\) Because of this relation, the words *mean* what they do.

Inferentialism, by contrast, understands meaning not in terms of denoting or representing but

\(^{125}\) An attempt to adjudicate between representationalist and inferentialist approaches to language is beyond the scope of this discussion. But, while I treat myself to Brandom’s inferentialist approach without providing defence against possible objections, I will consider and respond to an inferentialist critique of Brandom’s claim that assertion is downtown.

\(^{126}\) See, for example, Fodor (1981, 1987, 1998).
in terms of the rules of the linguistic practice, and what these (inferential) rules permit.\footnote{A dialectical note: I do not defend inferentialist semantics here, as doing would lead too far adrift from the present scope of discussion. I draw from the view because it is an influential account of language that, I shall argue, yields interesting results for my own purposes. It may be there are alternative, or even superior ways to reach the same kinds of conclusions I am interested in (such as Cuneo’s argument or some of the claims of discourse ethicists). But I think this approach involves the fewest and least ambitious claims, which is why I have pursued it over the alternatives. I explain some further reasons that I don’t pursue these other lines of argument later on in this chapter.} In Brandom’s words:

The standard way is to assume that one has a prior grip on the notion of truth, and use it to explain what good inference consists in. […] Inferentialist pragmatism reverses this order of explanation […] It starts with a practical distinction between good and bad inferences, understood as a distinction between appropriate and inappropriate doings, and goes on to understand talk about truth as talk about what is preserved by the good moves. (Brandom 2000, 12)

A sentence’s meaning, on the inferentialist account (and inferentialism generally takes sentence meaning to be conceptually prior to word meaning) is determined by the commitments\footnote{This is Brandom’s way of defining what we are committed to when we take on an assertoric commitment. Again, I think this sense of commitment is most obviously understood to be a commitment in the merely social sense. I don’t think we need to understand asserters as having an attitudinal commitment to justifying their entitlement to assert what they assert.} and entitlements one would be generating in asserting the sentence. Let’s

\[\text{[Footnotes]}\]
take an example.\textsuperscript{129} The claim ‘grass is green’ means what it does ‘in virtue of the fact that it can be used to undertake a commitment to grass’s being green; this commitment is constituted by the network of inferential entitlements and obligations that the commitment carries.’ In other words, if an agent asserts ‘grass is green’, it has the ‘social significance’ of an assertion because (1) it licences further assertions, or conclusions (according to the inferential rules) and thereby generates entitlements to other assertions and (2) it commits the asserter ‘to vindicate the original claim, showing that one is entitled to make it’. Thus, asserter undertake ‘justificatory responsibility for what is claimed’ (Brandom 1983, 641).

Brandom’s inferentialism is both pragmatic, in taking the practical distinction between good and bad inferences as explanatorily prior to concepts such as truth or correspondence, and normative in taking linguistic meaning to be determined by these normative, inferential rules which are ‘socially instituted’ by communal norms (Brandom, 1983, 640).\textsuperscript{130} So what is the nature and structure of linguistic practice? In Brandom’s view, there is one principal linguistic act type which explains how meaning is possible. This is assertion.

4.2. Assertion as ‘Downtown’

Linguistic practices and speech act types might be thought of as an unstructured, unhierarchical motley of various linguistic games. This is the picture of language the later Wittgenstein seems to have had. Brandom argues that language has a core, a principal

\textsuperscript{129} I take this example from Chrisman (Chrisman 2010, 115).

\textsuperscript{130} See also Gibbard (2012) for a defence of the normativity of linguistic meaning. Gibbard interprets his view as compatible with Brandom’s views (Gibbard 2010).
practice that defines it, ‘the involvement or non-involvement of which distinguishes linguistic activities from non-linguistic activities’. I quote Brandom’s explanation at length:

What makes something a specifically linguistic […] practice is that it accords some performances the force or significance of claimings, of propositionally contentful commitments, which can both serve as and stand in need of reasons. Practices that do not involve reasoning are not linguistic or (therefore) discursive practices. […] By contrast to Wittgenstein, the inferential identification of the conceptual claims that language (discursive practice) has a center; it is not a motley. Inferential practices of producing and consuming reasons are downtown in the region of linguistic practice. Suburban linguistic practices utilize and depend on the conceptual contents forged in the game of giving and asking for reasons, are parasitic on it. Claiming, being able to justify one’s claims, and using one’s claims to justify other claims and actions are not just one among other sets of things one can do with language. They are not on a par with other ‘games’ one can play. They are what in the first place make possible talking, and therefore thinking: sapience in general. Of course we do many other things as concept users besides applying concepts in judgment and action and justifying those applications. But […] those sophisticated, latecoming linguistic and more generally discursive activities are intelligible in principle only against the background of the core practices of inference–and–assertion. (Brandom 2000, 14-5)

The key moves Brandom is making here are as follows. Linguistic practice involves the use of concepts and meanings. From the inferentialist view, as I have explained, conceptual content and linguistic meaning are to be understood in terms of commitments and entitlements governed by normative, inferential rules. Hence, linguistic practice necessarily involves making these commitments, justifying our entitlements, and asking for justifications (giving and asking for reasons). This game of giving and asking for reasons is the ‘defining
core of discursive (concept-mongering) practice’ (Brandom 2000, 189-90). And the assertion game just is the most basic and essential ‘move’ in the broader game of giving and asking for reasons.

Brandom is not claiming that assertion is the only linguistic game there is. ‘Local’ forms of linguistic practice, or ‘suburban’ forms, as Brandom calls them, such as moral, logical, mathematical, or aesthetic discursive practices, have their own vocabularies and concepts. But all these linguistic practices, indeed all possible linguistic practices, ‘avail themselves of the same inferential machinery’ (Price 2013, 33). This is what makes them linguistic. Hence, the ‘one and only speech act that must be present [in a linguistic practice] is asserting’ (Brandom 2010). There can be no linguistic practice in which we are not negotiating transitions between commitments via inferential rules. So every linguistic practice necessarily involves giving and asking for reasons and hence assigning some things the statuses of assertions. So we can demarcate ‘linguistic practices by restricting that term to practices that confer on some performances the significance of claims or assertions.’ (Brandom 2000, 189) This is a conceptual truth about the nature of linguistic practices.

This is the basic idea behind the thesis that assertion is the fundamental linguistic practice. But I should acknowledge that not everyone, even on the pragmatist and inferentialist side of the philosophy of language, is convinced of the linguistic primacy of assertion. Most notably, Kukla and Lance (2009, 2010) argue that this commitment ought

---

131 Brandom uses the terms ‘discursive’ and ‘linguistic’ interchangeably.

132 I can’t consider the full space of theoretical options. This is a necessary limitation of my argument. But I do wish to consider this critique as it shares so many of Brandom’s (and therefore my) assumptions. I feel entitled to help myself to a pre-eminent theory of language, but not to something that turns out to be an esoteric, non-essential feature of that theory.
to be abandoned. I’ll briefly explain why their objection, even if correct, is non-threatening to my argument, before surveying the debate and concluding that the objection fails in any case. Doing so will also further elucidate the role of assertion in the inferentialist picture.

First off, as a dialectical point, even if Kukla and Lance are right and there is no linguistic practice that is ‘downtown’, their alternative perspective seems as attractive (if not more) as Brandom’s for the linguistic constitutivist. Kukla and Lance contend that certain kinds of language use are not built on assertion, but they think these other forms of language necessarily require language users to enter into explicitly ‘second-personal’ relations with one another. Indeed, a significant part of their complaint against Brandom is that it is insufficiently clear to them how the second-personal nature of many discursive commitments and entitlements can be adequately captured (a complaint I find unwarranted). In any case, Kukla and Lance’s less hierarchical account of language seems equally friendly to my ambition to show that language use requires occupying the second-person standpoint and adopting commitments to placing some significance on second-personal claims in practical deliberation.

But I still think it is worthwhile to take a closer look at Kukla and Lance’s critique, in part because they seem to think assertion is not as capable as other kinds of speech acts at accommodating second-personal relations. So, does assertion play, as Brandom claims, ‘an absolutely unique and central role’ in language generally? (Brandom 2010, 316) Or are Kukla and Lance right that this commitment ‘causes a problematic narrowing of [Brandom’s] theoretical vision and resources’ (Kukla and Lance 2010, 116). Kukla and Lance share Brandom’s pragmatism. They concur that language ought to be understood in terms of manipulation of ‘the structure of pragmatic space’ (Kukla and Lance 2010, 117). However, as Austin (1975) famously demonstrated, the pragmatics of language involves much more than propositional claims and semantics. It includes ‘interrogatives, imperatives, performatives,
and so on’ (Kukla and Lance 2010, 116). ‘It would be an odd coincidence,’ they write, if ‘no part of the structure of pragmatic space other than the specifically assertional would be relevant [to demarcating the linguistic]’ (117).

That’s the flavour of the critique. The specific argument they make against the primacy of assertion is that some ‘suburban’ forms of speech cannot be understood as parasitic on, or conceptually reliant on, assertion. Hence, assertion can’t play the essential role in distinguishing the linguistic from the non-linguistic. These utterance types include imperatives such as ‘Shut the door!’ vocatives such as ‘Yo, Emma!’ and perceptual episodes. Let’s take the former. An imperative is ‘always spoken in the second person. An imperative must be issued to someone in order for it to count as an imperative at all.’ (117–8) The relevant feature of these speech-act types, the feature Kukla and Lance contend Brandom’s account can’t accommodate, is that they are agent-relative. They are forms of speech where it matters who the addressee and addressee are. In ordering you to shut the door, I, in particular, am making a command that can only be fulfilled by you, in particular. Or, in the case of perceptual episodes, they seem inherently first-personal:

[M]y perception will only yield new entitlements for anyone if they originate with me – no one else but me can be the first to pass on my perceptual entitlements. But this can only happen if I recognize my perceptual episodes, in a normatively rich sense, as mine (Kukla and Lance 2010, 317)

Why do Kukla and Lance think that the existence of agent-relative speech-act types cannot be accommodated on Brandom’s account? The idea is that an account of language built entirely on assertion is impersonal. Assertions are impersonal in the specific sense that the commitments and entitlements that are recorded on the ‘scorecards’ (which assertions update) are ‘impersonal statuses that can retain their identity regardless of which subject holds them’ (118). There is, claim Kukla and Lance, no difference in the pragmatics, no difference on the deontic scorecards, between asserting ‘I think it will rain’ and asserting ‘X
thinks it will rain (where I am X)’. Since the scorecards of commitments and entitlements read the same, they must be missing something important:

What makes a series of ones and zeros on abstract scorecards into genuine commitments is the way they are taken up in practice, and this taking up is an essentially first-personal activity, just as holding one to a commitment, or recognizing someone’s entitlement, is inherently second-personal. But Brandom’s focus on assertions, and hence on impersonal normative statuses, forecloses his ability to accommodate these structurally first- and second-personal dimensions of normative transactions between agents in discourse. (Kukla and Lance 2010, 119)

I am unconvinced that Brandom’s account lacks the theoretical resources to capture the agent-relativity of certain speech-act types, and hence I am unconvinced that there is any rational pressure to abandon the assertion-first view.

The dispute seems to arise from how we understand the scorekeeping of commitments and entitlements within the inferentialist framework. So let’s first take a step back and clarify the wider project, what Brandom means by assertion, and how the deontic scorekeeping account relates. Brandom’s inferentialist and pragmatist account of semantics in terms of what is asserted by an assertion. And to understand that, he thinks, we need a pragmatic account of what is asserted. Brandom’s account of pragmatics, in turn, is a normative one. In figuring out what is asserted by an assertion, we need to understand how that assertion alters the normative landscape, which he describes in terms of changes to commitments and entitlements. The normative landscape, as I’ve called it, is understood as the scores on the metaphorical scorecards that record commitments and entitlements. So can this deck of scorecards record agent-relative information? Kukla and Lance think not. Hence, they think the game of assertion, which is the game that changes the scoreboard, cannot play the foundational role Brandom claims, of making linguistic practices possible.

But deontic scorecards can capture first- and second-personal information as well as third-personal information (Brandom (2010)). So they can include the kind of information
which needs to be captured if Brandom’s account of language is to include Kukla and Lance’s recognitives and other agent-relative forms of language. This is because we can make the distinction between attributing a commitment to another person and acknowledging it to oneself. ‘It is that basic distinction between deontic attitudes,’ claims Brandom, ‘that fills in the “I–thou” form of sociality that is contrasted with more traditional “I–we” versions.’ (Brandom 2010, 317). In other words, a distinction can be drawn between the attitudes that one takes to ‘one’s own score’ and that one takes to ‘another piece of third-person knowledge’. A lengthier discussion of the importance of this distinction can be found in Brandom’s “Asserting” (1983), in which he uses it to explain how we can take someone’s claim to be justified without taking it to be true. He writes, ‘The social difference between justification and truth is the perspectival difference between those endorsements on the part of others one is prepared to admit as authorized or permitted, and those endorsements one is oneself prepared to undertake.’ (Brandom 1983, 645) To take another’s claim as justified means admitting that person ‘has followed the rules of the asserting game and is entitled to his commitment’, but that does not require one to ‘make that endorsement or accept its consequences oneself [… as] one’s own commitment’, which would constitute taking the claim to be true. This distinction between how we relate to the scorecards of commitments and entitlements (taking some as our own and others as authorized in a third-personal sense) allows for the different perspectival forms of scorekeeping Kukla and Lance deem essential. So Brandom’s account requires no amendment. It has the resources to distinguish e.g. first-personal relations to commitments and entitlements from second- and third-personal relations to commitments and entitlements. So we don’t need to abandon the primacy of assertion in his account of deontic scorekeeping to accommodate the distinction.
4.3. Summary

In this subsection, I argued that if someone is a language-user, then he is an asserter. In the previous subsection, I argued that if someone is an asserter, then he is committed to the moral presuppositions and norms implicit in the second-person standpoint. And he is not merely committed to these presuppositions and norms in the social sense. He is, as a constitutive matter, volitionally committed to giving second-personal claims weight in deliberation. Hence, all language users, as a constitutive matter, have pro-moral commitments. The argument given so far entails non-agential linguistic constitutivism (see section 2).

5. Language and Agency

In this section, I argue that the possession of language is partly constitutive of full-blooded rational agency. I claim that necessarily, A is a rational agent only if A is a linguistic being. But more than this, being a speaker explains how rational thought, deliberation, and action are possible. Hence, being a language user partly constitutes being a rational agent. I first provide my preferred a priori argument for this view and then offer an empirical argument as a kind of backup option.

The significance of this section’s conclusion is this. It would then be true that whatever (moral) commitments language users undertake (as a constitutive matter), are also commitments all possible rational agents have. Hence, all rational agents will, according to the arguments I have made in this thesis, be committed to the equal dignity of all persons, and the contractualist norms of morality as equal accountability.

There is a substantial list of philosophers who claim that language makes rational agency, or even thought, possible. I’ll briefly mention some of the stronger versions of this claim, but the version of it I want to defend is this:
Language requirement on agency. Necessarily, language is required for full-blooded rational agency.

But before I jump into the arguments for this view, I want to first distinguish the sense of rational agency I am interested in from others. While there are a number of interesting lesser senses of agency, by ‘full-blooded rational agency’, I mean the distinctive form of agency that adult humans typically possess and babies, cats, and caterpillars don’t. To have normative practical reasons, for it to be true that you ought to do this, rather than that, requires this form of agency. This is why a normal adult human ought not to torture mice for fun, but the same can’t be said for a cat. We may not like that cats do that, but we don’t think that they ought to stop. This difference in our kinds of agency also explains why we don’t blame cats for doing so, and why we do blame normal adult humans. In short, only full-blooded rational agents are responsive to reasons and subject to rational norms. The moral rationalist has no interest in showing that cats have normative reasons to act morally—only that full-blooded rational agents do.

5.1. The A Priori Argument

It is undeniable that at least some mental state types, some forms of cognition, require language. At the very least, thinking linguistically requires language. Thinking ‘I will say “Yes please”’ requires language. I shall argue that the second-order, reflective form of thought which enables full-blooded rational agency also requires language.

133 The infant may be a ‘rational creature’ in the sense that it is a human and it will typically gain full rationality as the kind of being it is. But it does not currently possess full rational agency. See Davidson (1982, 317) for this distinction.
Full-blooded agency (and action), as I am calling it, is the kind of agency investigated by Anscombe (2000 [1957]) and Davidson (1980, 1982). Davidson labels it intentional agency—the capacity to perform intentional actions, where those are understood in terms of actions being for reasons which is what we do when we understand behaviour as caused by certain (combinations of) mental states, i.e. intentions. In other words, these mental states rationalise action. A key insight of theirs, and the relevant one for my argument, is that we can only understand beings as rational agents, and their behaviour as rational action, if we think they possess ‘propositional attitudes’, some ‘rich pattern of beliefs, desires, and intentions’ (Davidson 1982, 318). We need to understand rational agents and rational action in this way to understand ourselves as e.g. responsible in ways that cats are not. Roughly, the idea is simply that what makes action so different from unintentional behaviour consists of differences in the actors’ minds.

What kinds of thought does language enable? Some philosophers seem to identify thought or reason with language. Others claim that language is required for all and any

---

134 Metaphysically, such views fall under the orthodox ‘event-causal’ framework for understanding agency, according to which an action is an event with a particular causal history (that history: being caused by intentions). For defences of this approach, see Davidson 1980, Goldman 1970, Bratman 1987, Mele 1992.

135 I have in mind a certain German tradition which identified thought with language, developed by Hamann, ‘reason is language, logos’ (Hamann 1784, 5:177 trans. Forster 2010), ‘Reason and language, thought and word, are … essentially one’ (Schlegel 1815, trans. Forster 2010, 40).
I think these are implausible claims. Non-linguistic beings such as animals and infants seem to have some kind of capacity for thought even if it is not the same kind of thought that rational agents can have. A more modest and plausible conclusion, unsurprisingly more widely defended, is that language enables propositional thought—thoughts about things in a conceptually rich way. Beliefs that the oak tree is old. Desires to know how old it is. Intentions to ask the arborist tomorrow. All these views are versions of what Brandom calls differentialism (Brandom 2000, 2). Assimilationism, on the other hand, denies that there is any difference in kind between the cognition of linguistic beings like us and non-linguistic beings. It insists that there are merely differences in degree.

A number of philosophers have endorsed moderate forms of differentialism, according to which propositional attitudes such as belief, desire, and intention, intentional mental states, or conceptual thought, require language. Indeed, Davidson claims not only that language is required for propositional attitudes such as belief, desire, and intention but...

---

136 Hegel, for example, claims, ‘To want to think without words […] is clearly an absurdity.’ (Hegel 1830, par.462 trans. Forster 2011, 151). Schlegel concurs, ‘one cannot think without words’ (Schlegel KFSA 15/2:3 trans. Forster 2011, 40).

137 Some versions of the ‘language-of-thought’ hypothesis would count as differentialists, particularly those that see the social phenomenon of language as prior to the ‘interior’ mental language in which thought occurs. However, there are so many versions of the language-of-thought thesis that it is impossible to assign them clearly to any of the claims mentioned here.

also that, therefore, *rational agency* requires language (as it involves such attitudes). I’ll argue for a moderate form of differentialism that draws on Davidson and Brandom’s work.

Let’s begin with the following case from Malcolm:

Suppose our dog is chasing the neighbor’s cat. The latter runs full tilt toward the oak tree, but suddenly swerves at the last moment and disappears up a nearby maple. The dog doesn’t see this maneuver and on arriving at the oak tree he rears up on his hind feet, paws at the trunk as if trying to scale it, and barks excitedly into the branches above. We who observe this whole episode from a window say, ‘He thinks that the cat went up that oak tree.’ (Malcolm 1973, 13)

Isn’t it clear from this case that non-linguistic animals such as dogs can have beliefs and desires? Not at all. Attributing beliefs, desires, and intentions to animals admittedly ‘works for dogs and frogs much as it does for people’(Davidson 1982, 323), in the sense that treating them *as if* they have such attitudes is predictive and can often be the only way we can make sense of their behaviour. But this clearly does not mean that they actually *have* propositional attitudes. Similarly, someone uninformed about heat-seeking missiles may ‘be justified in attributing a desire and beliefs to the missile, but he would be wrong’ (323). Attributing propositional attitudes to the missile may be successful, predictive, etc., so it may be seen as suitable in conversation, but it would still be an error.

So why think that propositional attitudes are in the exclusive possession of the linguistically blessed? Davidson offers an argument with two premises: (1) that to have a belief, it is necessary to have the concept of belief and (2) in order to have the concept of belief one must have language (324). Why think that belief requires the *concept* of belief? To understand what Davidson means by belief, think of the distinction between ‘thinking’ there is a tree (in the mere sense of being aware of a tree) and thinking *that* there is a tree (in the sense that one is aware *that one is aware* that there is a tree) (Davidson 1982, 324). Only in the latter sense is our thought propositional and hence interpretable as a *belief* that there is a tree. So what is required to have a belief of this propositional kind? In Davidson’s words, ‘in
order to have any propositional attitude at all, it is necessary to have the concept of a belief, to have a belief about some belief” (326). This explains the possibility of surprise when we learn our beliefs are false. We don’t simply cease to have our prior beliefs when the world turns out not to be the way we thought. We notice something, ‘a contrast between what I did believe and what I come to believe. Such awareness, however, is a belief about a belief’ (326). This brings us to the second premise of Davidson’s argument, that language is needed to have concepts (and, particularly, the concept of belief). Being aware that one has a belief, any belief, requires that one understands there is a possibility of error. And we can’t understand ourselves as understanding that possibility unless we understand that beliefs are things which can be true or false, correct or incorrect. This requires the concepts of true and false, correct and incorrect, etc. We need to have beliefs about what beliefs are, i.e. the concept of belief, correctness and incorrectness, etc., in order to have beliefs. Having the concept of a belief involves having many other foundational concepts, such as true and false, and those, in turn, require further concepts.

The particularities of Davidson’s argument are controversial. But the intuitive core of Davidson’s argument is a plausible kind of holism about concept-possession and propositional thought which renders the possession of propositional attitudes and concepts something close to an all-or-nothing affair. To assign one proportional attitude to a being is to assign many. And non-linguistic beings can only fall on the ‘nothing’ side, because it’s implausible they could possess the necessary concepts to have the propositional attitudes. Stitch’s (1979) example of Fido the dog is helpful here:

Fido believes that there is a meaty bone buried in the yard and wants to get it. […] It surely cannot be quite right to say that Fido believes there is a meaty bone buried in the yard. After all, Fido does not even have the concept of a bone, much less the concept of a meaty bone or a yard. […] For Fido does not, it seems safe to assume, have any beliefs about the origin and general anatomical function of bones. Nor would he recognise or exhibit any interest in
chewing atypical bones -- the bones of the middle ear, for example, or the collar bone of a blue whale. Worse yet, Fido does not know the difference between real bones and a variety of actual or imaginable ersatz bones (made of realistic looking plastic, perhaps, and partially covered with textured soy protein suitably flavoured). Nor is there anything that would count as explaining the difference between real and fake bones to the dog. Fido is incapable of understanding that distinction. But given Fido's conceptual and cognitive poverty in matters concerned with bones, it is surely wrong to ascribe to him any belief about a bone. To clinch the point, we need only reflect that we would certainly balk if the same belief were attributed to a human who was as irretrievably ignorant about bones as we take Fido to be. (Stich 1979, 18-9)

The key point is this. Propositional attitudes such as conceptually rich beliefs and desires require a degree of conceptual sophistication. And it is highly plausible that this level of conceptual sophistication requires language.

The theoretical mechanics developed by pragmatists such as Brandom, I think, offer a useful alternative conceptual framework which elucidates conceptual and mental holism and why it would suggest concept possession requires language.\textsuperscript{139} The intentional state of belief,
on the inferentialist picture, can be understood as a kind of claim one attributes to oneself (or another). It’s a commitment on someone’s personal deontic scorecard, so to speak. But the belief that $p$ can only be understood as a belief that $p$, in virtue of its inferential connections with other claims. A belief that that wooden thing is a tree licences the inference that it is a plant. We wouldn’t attribute a belief ‘that the wooden thing is a tree’ to someone whose belief lacked these kinds of inferential connections. But this kind of grasp on a claim’s inferential connections seems sufficient to count as a full-blown language user. So any practice such as belief or desire that involves ‘concept-mongering’, in Brandom’s terminology, necessarily requires language.

In sum, propositional attitudes such as belief, desire, and intention require their possessors to have a degree of conceptual mastery which plausibly involves being a language-user. Non-linguistic beings can be sensibly thought of as if they have propositional beliefs and desires, but the sense in which they have beliefs and desires is both lesser and parasitic on the richer propositional, conceptually rich sense of belief and desire that linguistic beings can possess. The moderate differentialist conclusion follows: only linguistic beings can have propositional attitudes.\(^{140}\) Since propositional attitudes such as belief, desire, the approach, but simply because it involves more ambitious claims than I need for my argument.

\(^{140}\) Glock (2010) formalises Davidson’s differentialist argument into three major premises, an intermediary conclusion, and a final conclusion. These are: (1) Concept thesis. Thinking (having thoughts) requires concept-possession. (2) Language thesis. Concept-possession requires language. (3) Lingualist conclusion. Thinking requires language. (4) Dumbness thesis. Animals lack language. (5) Differentialist conclusion. Animals cannot think. (Glock 2010). I am only interested in the lingualist conclusion, and hence have only defended (1) and (2).
and intention are required for rational agency, we can conclude that the language requirement on agency is true.

Let’s connect this argument with those put forth earlier in the chapter. Intentional, rational agency, the kind we think renders its possessors subject to rational demands and reasons, involves agents possessing the propositional mental states of intention, desire, and belief. And the capacity to have these mental states requires language possession. Hence, all rational agents are linguistic beings and, therefore, rational agents are subject to the same norms as linguistic beings, as asserters. Following from the linguistic constitutivist conclusions drawn in the first half of this chapter, this entails that all rational agents are, as occupiers of the second-person standpoint, subject to contractualist moral norms and have normative reasons to act morally, as explained by their constitutive commitments to give weight to second-personal claims.

5.2. The Empirical Argument

While I find the a priori argument for the view that rational agency requires language compelling, I would also briefly like to present an empirical argument for the language requirement on agency. Consider this an interesting back-up option for the agential linguistic constitutivist.

Now that we’re not restricted to a priori arguments, we can be more targeted. While full-blooded rational agency certainly requires propositional attitudes such as beliefs, desires, and intentions, I think it requires even more specific forms of cognition. In particular, full-blooded rational agency seems to involve what Clark calls ‘second-order cognitive dynamics’, which is roughly thinking about thought (Clark 1996, 177). In humans at least, our capacity for this kind of cognition is explained by our possession of language.
The first thing to say about the idea that some kind of second-order thought is required for rational agency is that it is not a new idea. Almost every constitutivist and many theorists of agency have claimed that the capacity for some distinctively reflective mode of cognition is required for rational agency.\textsuperscript{141} Arruda labels this thesis the self-reflection requirement:

\textbf{Self-Reflection Requirement.} Competent agents engage in self-reflection as part of their self-constitution as agents. (Arruda 2016, 1166)

This formulation is a helpful start but since only Kantian constitutivists tend to use the concept ‘self-constitution’ we might restate the self-reflection requirement more generally as:

\textbf{Self-Reflection Requirement.*} Some distinctively reflective mode of cognition is partly constitutive of rational agency; its possession or exercise explains how rational agency is possible.

Accounts differ over what exact kind of reflective/second-order thought plays a role in making rational agency possible.\textsuperscript{142} But I think we can extract a common intuitive core.


\textsuperscript{142} The two most influential constitutivist accounts are provided by Korsgaard and Velleman. Velleman’s ‘kinda-Kantian’ account of rational agency claims its exercise essentially involves reflexive thought. Korsgaard takes a more distinctively Kantian view that the ‘source of normativity’ is found in the capacity of self-reflection. More specifically, there is a normatively authoritative form of reflection - ‘reflective scrutiny’ - which confers normativity on its conclusions: ‘The capacity for self-conscious reflection about our own actions covers on us a kind of authority over ourselves, and it is this authority which gives normativity to moral claims’ (Korsgaard 1996, 19-20). Harry Frankfurt also endorses a version of the self-reflection
Full-blooded rational agency, the kind of agency we attribute to beings who are responsible for performing actions, and who are subject to the norms of practical rationality, requires the capacity to hold one’s attitudinal states in mind and think about them. A being incapable of thinking about his thoughts could only ever be compelled by his desires, intentions, or plans to act this way or that. But a full-blooded rational agent must be capable of more than this. He must be capable of accepting or rejecting, approving or disapproving, endorsing or condemning, his attitudes. He must be capable of judging his own mental state (the desire for cake, the belief that there is cake in the fridge) to be reason-providing or not. All this must be the case for us to think of rational agents as operating, to put it in Kantian terms, ‘under the idea of freedom’, as beings who are meaningfully autonomous and hence interpretable as

---

requirement from a broadly Humean account, according to which our ability to form and maintain ‘higher order’ desires (or practical commitments, plans etc.) partly explains why we are autonomous agents, the kinds of beings for whom practical normativity exists. On Frankfurt’s account, persons (i.e. human beings/agents) share with non-persons the mental states of desire and belief. These mental states can also play a similar role in choice-making in non-persons as it does in persons. What distinguishes persons from non-persons, in Frankfurt’s view, is the ‘structure of the will’. Persons are distinctive in virtue of possessing ‘second-order desires’, i.e. desires about their first-order, ordinary desires. This crucial difference in structure of persons’ wills allows persons ‘the capacity for reflective self-evaluation’ (Frankfurt 1971, 7). One can adopt pro-attitudes and beliefs about one’s first order desires. And this means we can accept or reject first order desires in a way that affirms them as our own, thus conferring normativity, or rejects them as foreign to us, in some sense. For criticism of Frankfurt’s view, see Watson’s argument that ‘higher-order volitions are just, after all, desires, and nothing about their level gives them any special authority’ (Watson 1987, 149). But the point is just that all these accounts of what rational agency consists in involve second-order forms of cognition.
responsible for their actions. In sum, the capacity for reflective thought, thought about one’s thoughts and attitudes, explains how full-blooded rational agency is possible.

Why think that second-order forms of thought, such as reflective thought, require language? Language is not simply a means of communication. It is also an extremely transformational mental tool. Like any tool, it allows us to do a lot more than we could otherwise. At least, many philosophers and cognitive scientists have thought so. Dennett, for instance, claims that becoming initiated into language profoundly reprograms the brain, resulting in entirely new neural machinery and cognitive capacities and phenomena (Dennett 1991, 218-219 and 1995, 370-373). It is not our relatively minor differences in brain structure that explain our dissimilarity to other apes, but the transformational reprogramming of our brains that language performs.

Dennett’s reasoning for this view runs approximately as follows. First, there are profound differences between the cognitive capacities of humans and non-human animals, despite similar brain structure, particularly with mammals such as dolphins, dogs, and, particularly, other great apes. ‘As researchers regularly discover,’ Dennett notes, ‘the more ingeniously you investigate the competence of nonhuman animals, the more likely you are to discover abrupt gaps in competence.’ The obvious explanation of these profound divergences is that we can think in ways that they can’t. In particular, our advantage seems to be ‘the perspective we gain from our ability to reflect in our special ways’ (Dennett 1996, 116, my emphasis). What is required for this reflective form of cognition? A particular form of intentionality, where intentional systems can adopt the ‘intentional stance’ toward themselves. To reflect, we need a shift from ‘first-order’ intentional systems to ‘second-order’

---

143 Dennett is drawing on the empirical research to make this point, and particularly on Cheney and Seyfarth’s investigation of the capacities of vervet monkeys (Cheney and Seyfarth, 1990).
intentional systems (Dennett 1996, 121). Animal thoughts may be about things, but their thoughts can’t be about animal thoughts. In other words, reflective thought requires second-order thought.

Dennett provides his own account of how language reprograms the mind in such a way that reflective thought and second-order intentionality possible (and, indeed, consciousness (in its reflective human form at least). But as Clark argues, language even conceived of as an ‘external resource’ (1998, 167), can enhance, facilitate, and extend cognition in a number of fundamental ways, though without requiring the computational reprogramming of the brain:144

[W]e use words to focus, clarify, transform, offload, and control our own thinkings. Thus understood, language is not the mere imperfect mirror of our intuitive knowledge. Rather, it is part and parcel of the mechanism of reason itself. (Clark 1998, 207)

More precisely, Clark describes six (non-exhaustive) ways that the possession of language extends and enhances our cognition. Language possession augments memory, allows us to simplify our models of our environment, to coordinate, to ‘transcend path-dependent learning’ (not to be constrained by one’s own paths of cognition), to generate ‘control loops’ (e.g. to write down plans, to respond to instruction), and to manipulate and represent in ways that would ‘quickly baffle the un-[language]-augmented brain’ (Clark 1998, 173). Do any of these six linguistically-enabled augmentations of cognition enable this

144 Both Dennett and Clark accept ‘supra-communicative’ views of language, that is views which claim language does far more to cognition than act as a vehicle for communication. Other thinkers that clearly fall into the supra-communicative camp include Carruthers, who claims that thinking is done via ‘inner speech’ (Carruthers 1996, esp. ch.2), and Herder, who describes language as ‘constitutive of reflection’ (Taylor 2016, 33). This is the same concept as Brandom’s differentialism.
distinctive second-order form of thought? The kind that is required for rational agency? I don’t think so. A massive enhancement in an existing capacity is not a new cognitive capacity. In these six ways, language merely seems to make us better at doing what non-linguistic beings could do to some extent.

However, Clark makes one more suggestion about language which, I think, shows that at least one distinctive form of cognition requires language. The idea is that our grasping of a language enables ‘our ability to display second-order cognitive dynamics’, a.k.a. ‘metacognition’ i.e. all those forms of cognition that require thinking about our thoughts:

[A] cluster of powerful capacities involving self-evaluation, self-criticism and finely honed remedial responses. Examples would include: recognizing a flaw in our own plan or argument, and dedicating further cognitive efforts to fixing it; reflecting on the unreliability of our own initial judgements in certain types of situation and proceeding with special caution as a result; coming to see why we reached a particular conclusion by appreciating the logical transitions in our own thought; thinking about the conditions under which we think best and trying to bring them about. The list could be continued, but the pattern should be clear. In all these cases we are effectively thinking about our own cognitive profiles or about specific thoughts. (Clark 1996, 177)

Many of these are precisely the sorts of capacities which explain the capacity of full-blown rational agency. But how does language deliver them? Clark offers the following rough suggestion:

[A]s soon as we formulate a thought in words (or on paper), it becomes an object both for ourselves and for others. As an object it is the kind of thing we can have thoughts about. In creating the object we need have no thoughts about thoughts—but once it is there, the opportunity immediately exists to attend to it as an object in its own right. The process of linguistic formulation thus creates the stable structure to which subsequent thinkings attach. (Clark 1996, 177)
I think Clark’s speculation that second-order thought and its variety of associated capacities are made possible by the possession of a ‘public language’ is correct. But we need to say a bit more in support of the idea that second-order cognitive dynamics require language. We don’t merely want to claim that language helps beings have second-order cognitive dynamics in the same way that language possession helps augment memory (via conversation, books etc.). In other words, we need to refute the idea that language is helpful but unnecessary for a being to have second-order cognitive dynamics.

An initial line of thought in favour of this view is that second-order thought, thinking about thinking, is plausibly a distinctively human capacity, ‘one not evidently shared by the other, non-language-using animals who share our planet’ (Clark 1996, 177). Given that we are also the only animals that possess language (at least in the relevant rich sense), we might think the best explanation of these joint facts would be that language explains our capacity for second-order thought, and hence that second-order thought is ‘an entire species of thought in which language plays the generative role—a species of thought which is not just reflected in, or extended by, our use of words but is directly dependent upon language for its very existence.’ (177)

This argument by inference to the best explanation might lend some plausibility to its conclusion. But there are stronger reasons for thinking that language enables second-order thought. Second-order thought, thought about thought, requires the object of thought to be another thought. So the question is what is necessary for these first-order thoughts and mental states to be receivable by the mind in such that they can become the objects of further thoughts? And the big idea is that the cognitive tool of language is required. By labelling and fixing an idea in a public language, we both ‘offload’ the content of the cognition to the
social and inferential realm of language and thereby ‘fix’ its meaning and content.145 In other words, a public language permits us to *stabilise* cognition into (potentially expressible) linguistically fixed thoughts.146 And this stabilisation of cognition into thoughts, through the vehicle of public language, is what provides thoughts with the necessary conceptual and inferential stability for them to become the objects of further thought.147 This stabilisation of cognition as linguistically fixed thought ‘enables us to inspect and criticize our own reasoning in ways that no other representational modality allows’ (Clark 1998, 177).148 So

145 Clark summarises: ‘Public language and the inner rehearsal of sentences would, on this model, act like the aerial roots of the Mangrove tree—the words would serve as fixed points capable of attracting and positioning additional intellectual matter, creating the islands of second-order thought so characteristic of the cognitive landscape of homo sapiens.’ (177)

146 Taylor makes much the same point, claiming that language means that ‘Instead of being overcome by the ocean of sensations as they rush by us, we are able to distinguish one wave, and hold it in clear, calm attention, It is this new space of attention, of distance from the immediate instinctual significance of things, of focused awareness.’ (Taylor 2016, 12).

147 Clark is drawing on (and I think improving on) some of the arguments of Jackendoff, who claims ‘language provides a scaffold that makes possible certain varieties of thought that are more complex than are available to nonlinguistic organisms’ (Jackendoff 1996, 2) and elsewhere claims that ‘the mental rehearsal of sentences may be the primary means by which our own thoughts are able to become objects of further attention and reflection’ (Clark 1996, 177).

148 For what I take to be an a priori argument by elimination for this kind of view, see Bermúdez (2003). Bermúdez argues that the ‘vehicles’ that allow the object thoughts of second-order thought to be the objects of further thoughts cannot be anything but a public
it’s not just that language helps us process cognition into processible ‘chunks’ that can be the objects of further thought. It’s that only through language is this kind of cognition possible for us. Hence, language possession is partly constitutive of the self-reflective modes of thought that rational agency requires.

To sum up, I have claimed that full-blooded rational agency involves second-order forms of cognition such as self-reflective thought, and that our capacity to have second-order thoughts is explained by our possession of language. I argued this was the case because language is what makes it possible to ‘anchor’ thoughts, to provide them with the necessary kind of cognitive stability, for us to take them as the objects of further thoughts. Hence, all full-blooded rational agents are necessarily possessors of language. The language requirement on agency is true.

This argument has a limitation worth mentioning. I think it is most plausibly understood to involve empirical claims. So it allows us to conclude that there is no non-linguistic rational agent, given how language and second-personal thought happen to connect. But it does not entitle us to conclude that there is no possible non-linguistic rational agent—some agent who has some other cognitive tool that can sufficiently stabilise thoughts such that they can be the objects of further thoughts. The empirically grounded argument provided here only entitles us to conclude that for beings like us, the kind who can only access second-

language (Bermúdez 2003, 8.2). Hence, ‘thoughts can only be the objects of the type of reflexive thinking in which thoughts are the objects of thought if they have natural language vehicles [...] We think about thoughts through thinking about the sentences through which those thoughts might be expressed.’ (2003, 8.2)

149 Dennett seems not to distinguish his a priori and empirical claims. Clark, I think, takes his argument as empirically based and his conclusion as contingently true.
order cognition because of our grasp of language, language possession is required for rational agency.

Thinking through the implications of this limitation, I think we arrive at a vaguely Aristotelian step away from Kantianism, in the following sense. If linguistic practice involves a commitment to moral norms, as I have claimed in this chapter, then the argument offered in this subsection entails that all rational agents necessarily have normative reasons to act morally. Or, more precisely, it claims that all possible rational agents who are relevantly similar to human beings in the sense that they require language for second-order thought have normative reasons to act morally. It does not entail that all possible agents have normative reasons to act morally. The argument provided here, when combined with speaker constitutivism, accommodates a slightly weakened version of moral rationalism. Nevertheless, I think this would be an appealing and significant result for the Humean constitutivist to achieve.\textsuperscript{150}

6. Objections: Condemning the Rogue’s Gallery

Before concluding the thesis, it is worth drawing together a collection of possible counterexamples to the claims I have made in this chapter.\textsuperscript{151} In the previous chapter, I develop and explore what I dub the ‘optionality challenge’ to second-personal forms of

\textsuperscript{150} Of course, I think we can do better than this with the a priori argument already provided.

But I think this is an interesting alternative for those who are left unconvinced by the a priori argument.

\textsuperscript{151} I thank Matthew Silverstein and Michael Cholbi for challenging me on these points and for their generous and thoughtful discussion of these issues.
constitutivism such as Darwall’s. In this chapter, I promised a strategy which meets this challenge. But have I succeeded?

6.1. Reviving the Optionality Challenge

Recall the nature of the challenge. Second-person constitutivism, as per Darwall’s interpretation, doesn't necessarily ground itself in the constitutive features of agency. Instead, it's based on the inherent assumptions we make when we view things from a second-person perspective, not from the lens of agency. Hence, the second-personal perspective anchors morality in interpersonal dynamics. However, this raises a concern. It's prima facie conceivable for a rational agent to not recognize or assert these second-personal claims, effectively choosing not to engage in this standpoint. This implies that such a person might be devoid of moral reasons, assuming morality is indeed anchored in the second-person perspective. The mere possibility of such an individual highlights a significant criticism. This is, in essence, what I labelled the 'optionality challenge'. In grounding moral normativity in the second-person perspective, we fail to secure the comprehensive universality of moral normativity. We fail to accommodate moral rationalism. The linguistic constitutivism I have developed in this chapter remains second-personal in nature. Hence, the optionality challenge must be answered. But to get more purchase on the force of the objection, let’s consider a few unsavory characters.

6.2. The Strategic Hobbesian

The strategic Hobbesian agrees with everything I have to say in this chapter. He is convinced that, by virtue of his being a rational agent, he must be practically committed to second-personal accountability with his fellow persons, and the contractualist moral norms which follow. But when the reasons he has in virtue of those constitutive pro-moral practical
commitments run up against those reasons he has given his other, contingent desires, he asks himself, ‘What’s in it for me? Why should I give my moral reasons any weight, whenever they run up against my other reasons?’

I refer this character back to my account of the weight of reasons in Chapter 3. There, I endorsed the ‘standard’ Humean account of the weight of normative reasons. According to that account, proportionalism, the weight of one’s normative practical reasons is a function of the strength of one’s reason-providing desires (and other factors such as the degree to which various actions promote the satisfaction of those desires). ‘All the better!’ says the Hobbesian, ‘I happen to desire to serve myself very strongly indeed! I must be acting rationally by systematically sidelining these pesky second-personal reasons (which I acknowledge I have).’

What is less good news for the Hobbesian is that I also endorsed a particular theory of strength-of-desire: ideal preferentialism. According to this view, the strength of the strategic Hobbesian’s desires, strength in the relevant proportionalist sense for determining the weight of reasons, is determined by something like the counterfactual robustness of those desires under ideal conditions. I argued that under such conditions, we would, in cases in which we had incompatible ends, first abandon our ends, or our practical commitments, which we recognised as contingent (and hence abandonable) before we would abandon those practical commitments we recognised were partly constitutive of agency. Indeed, we could never abandon the latter.

The point here is that even if the Hobbesian feels his self-serving desires much more intensely than he feels his constitutive practical commitment to second-personal norms, the latter remains the stronger desire, that is the more reason-providing desire, on my ideal preferentialist account. The strategic Hobbesian is quite right that he has reasons to act both self-servingly and morally, but what he doesn’t realise is that his reasons to act morally have
more rational weight. Hence, if the strategic Hobbesian were to ignore my advice and act on the basis of his contingent self-serving amoral desires whenever they conflicted with the demands of his constitutive commitment, he would be failing to act rationally in those instances. He would be failing to act as he has all-things-considered reasons to act.

6.3. The Extreme Bigot

Let’s try another difficult character. What can we say about a bigot? Not just an ordinary bigot, but an extreme bigot? Someone who doesn’t talk or listen or recognise an entire class of people as persons—as apt for second-personal address? Could he simply opt out of the universal application of morality, rationally speaking? To do so would obviously be a moral failing, but why would it be a rational failing? How does my account rule out the rationality of a selective and non-universal morality of the extreme bigot?

A preliminary point is worth making. It is doubtful that even the most extreme actual bigots meet this standard by which they completely avoid the second-person standpoint in their interactions with members of the groups they despise.

Let’s not think about actual bigots, but instead about the philosophically ideal bigot. Imagine a slave’s master who never makes second-personal claim on the slave’s will, in giving an order, as he refuses to see the slave as second-personally competent, but makes statements such as ‘I will punish you if you fail to X’. Or, if you think as I do that the expectation of being understood already commits one to the second-person standpoint, imagine a bigot who lives thousands of miles from the people he refuses to recognise as apt for second-personal address, such that there is no chance for second-personal interaction to sneak in. My claim is that there is still a rational failing in such cases.

As Abraham Lincoln said, ‘As I would not be a slave, so I would not be a master.’ The thought here is that the extreme bigot seems to be rationally failing is on the basis of
incoherence. Since all persons have the same second-personal capacities that you have, and those capacities make you apt for second-personal address (and explain your linguistic capacities etc.), it is irrational on grounds of inconsistency to simultaneously recognise yourself as apt for second-personal address and fail to recognise others who objectively possess the same capacities. Indeed, even if you don’t recognise yourself as linguistically competent or apt for second-personal address, insofar as you recognise anyone as such, you are rationally obliged to recognise all relevantly similar beings in the same way (including yourself and all others). Rationally, second-personal recognition can’t be done selectively. This is why I claimed, in 3.4.3, that we do not merely commit ourselves to our (intended) audience when we assert. In asserting, we are, and must understand ourselves as, issuing a redeemable token, a general licence, to all those who are assertorically capable, all those who can recognise our assertion as an assertion, to hold us liable via impersonal reactive attitudes and other forms of sanction. Of course, the bigot could escape this incoherence by going the other way, by abandoning the second-person standpoint and its associated practical commitments entirely. This suggestion leads us to even more troublesome character.

6.4. The Psychopath

The traditional villain of Humeanism, the ideally coherent psychopath, is the toughest character in our rogue’s gallery to convict of irrationality. Let’s adapt this character to the

---

152 The philosophical use of psychopathology or sociopathology, going back to Hume’s discussion of the ‘sensible knave,’ is highly idealised and differs significantly from the real condition. For the purposes of this discussion, a psychopath is an agent who has no ‘pro-moral’ desire-like mental states. In other words, none of the ideal psychopath’s desires would be served by acting morally, whereas some may be served by acting amorally or immorally.
linguistic constitutivism I have defended in this chapter. The psychopath agrees with the extreme bigot that his hated group are not conceivable as mutually accountable persons. Indeed, he agrees that it is impossible to even talk to such people, given that doing so would involve second-personal interaction and its associated practical commitments. But the psychopath does not stop there; he extends the bigot’s attitude and universalises it. It is clearly going to be much harder to accuse this character of the kind of incoherence I suspect the bigot is guilty of.

Now, I might try to make the psychopath realise that he takes himself to be the kind of being that can issue second-personal demands to others, even if he does not recognise that he is likewise accountable to them. But let’s make the case harder for ourselves. Perhaps our psychopath even extends his refusal to enter the second-personal standpoint to his treatment of himself. Perhaps he does not even consider himself capable of issuing second-personal demands. ‘Why on earth should anyone lift their foot from mine because I tell them to do so?’ he asks, ‘I have no such authority.’ In fact, he does not even ask this question because he refuses to engage in the second-personally infused practice of language. We can all agree this character is failing morally. But where exactly is the irrationality? Why must he occupy the second-personal standpoint? Why must he be a language user?

To address the ideally coherent psychopath, there are no shortcuts or modest forms of response. We cannot point to a convenient practical incoherence. We need to draw on the full resources of the linguistic constitutivism developed in this chapter. I have defended linguistic constitutivism as a form of agential constitutivism and claimed that in order to even count as a rational agent, one must adopt a pro-moral practical commitment. In other words, if the agential linguistic constitutivism I have developed here is correct, then the psychopath, in lacking this pro-moral commitment, is not a rational agent in the first place. He is not the kind of being who has moral reasons, nor indeed normative practical reasons. Specifically,
the idea is that insofar as he is a rational agent, he has the cognitive capacities of a rational agent. Roughly, he is capable of being responsive to reasons. He’s the kind of thing that can play the normative game.

Section 5 of this chapter argues that the possession of language is partly constitutive of the kind of full-blooded rational agency we attribute to entities insofar as we think they could have moral reasons. I claim that necessarily, A is a rational agent only if A is a linguistic being. But more than this, being a speaker explains how rational thought, deliberation, and action are possible. Hence, being a language user partly constitutes being a rational agent. The claim, then, is that the psychopath cannot avoid language insofar as he is a possible rational agent. Hence, he necessarily undertakes commitments of mutual accountability and is subject to the demands of moral normativity.

Of course, a reader who considered the arguments for these views I have provided and been left unconvinced could see the apparent possibility of the non-linguistic but agential psychopath as a reductio ab absurdum against my arguments and against linguistic constitutivism. If my arguments rule out the non-linguistic but agential psychopath, the critic might have it, so the worse for my arguments. Here, I can only appeal to the strength of those arguments on their own terms, and briefly take the opportunity to remind the reader of the general picture of rational agency which falls out of my account, and how it rules out the possibility of the non-linguistic rational agent.

On my account, defended in detail in section 5 of this chapter, linguistic competence is partly constitutive of rational agency. I am not the first to reach this conclusion (Davidson 1982). But the picture I develop is, roughly, the following. Rational agents are necessarily linguistic beings because the possession of language enables the very cognitive capacities which explain how rational agency is possible. Language enables rational agency because it makes conceptual thought, second-order thought, and reflective thought possible. It gives us
the forms of cognition that we require to enter the ‘space of reasons.’ Must the psychopath actually speak? Need he explicitly issue and recognise second-personal claims via (linguistic) interaction with others? No. There is no metaphysical necessity for any of that to occur. Indeed, this kind of thought is the basis of my optionality objection to Darwall’s second-person constitutivism in Chapter 4. But must he be a language user? Must he be a player in the game of ‘giving and asking for reasons’, if only in his own mind with hypothetical interlocutors? My arguments entail that he must be, insofar as he can also be considered a rational agent and the kind of being to whom we want to attribute moral reasons.

7. Conclusion

To sum up, the basic idea behind linguistic constitutivism is that Darwallian presuppositions of morality as equal accountability are already implicit in the second-personal standpoint that we necessarily occupy as language users. I’ll briefly summarise the key argumentative moves and how they fit together.

In section 2, I characterised linguistic constitutivism as the view that certain normative facts are explained by constitutive features of linguistic beings. I distinguished a non-agential and an agential version of linguistic constitutivism and explained the versions of moral rationalism each would plausibly deliver.

In section 3, I argued that being a player in the linguistic practice of assertion involves occupying the second-person standpoint, in the sense of being committed to the second-personal claims, reasons, and relations which underlie assertoric responsibility. In asserting some claim I am thereby necessarily undertaking certain second-personal responsibilities to others. If what I have asserted is unwarranted, others now possess the normative standing and authority such that it would be appropriate for them to blame me or resent me, for example. We can’t make sense of someone as having asserted anything unless we take them to be
normatively committed and accountable in these ways. The most plausible understanding of the nature of these assertive commitments and ways of being accountable, I argued, is that they are owed to one another: that they are second-personal. We seem, then, to implicitly occupy the second-personal normative relations of equal accountability as speakers. Drawing on Darwall’s arguments, I concluded that, as linguistic beings, we necessarily presuppose the equality and autonomy of ourselves and others, which itself commits us to contractualist moral norms that reflect equal accountability and respect. Assertion also requires undertaking a volitional commitment to give those second-personal claims that are the foundation of contractualist moral norms weight in practical deliberation. We can’t understand people as in the assertion game without them undertaking this commitment. Combined with the Humean theory of reasons, this constitutive commitment also secures the normativity of contractualist moral norms and permits us to answer the optionality objection.

In section 4, I argued that to count as a competent language user, one must be capable of performing ‘downtown’ linguistic acts. Drawing on Brandom’s inferentialism, I argued that what distinguishes linguistic from non-linguistic practices is that some moves in those practices are given the status of an assertion/claim. Hence, one just doesn’t count as a linguistic being unless one is in the game of assertion.

While the arguments provided in sections 3 and 4 collectively yield a non-agential linguistic constitutivism, an interesting view in its own right, I then explained, in section 4, why linguistic constitutivism is plausibly considered an instance of the more philosophically interesting agential form of constitutivism; I argued that rational agency requires language. Drawing primarily on ideas from the work of Davidson, Stich, and Brandom, I provided an a priori argument for this conclusion, according to which (a) intentional thought, having propositional attitudes such as belief, desire, and intention, is required for full-blooded agency, and (b) language is required for this kind of thought. I also provided an empirical
version of the argument, drawing on ideas about how language enables new forms of cognition. I argued in particular that second-order thought, reflective thought, is required for full-blooded rational agency and that the possession of public language allows us to have second-order thought.

Linguistic constitutivism is the view that, as a constitutive matter, all language users are rationally committed to contractualist moral norms. Combined with the language requirement on agency, we get agential linguistic constitutivism. This claims that all possible rational agents are rationally committed to morality as equal accountability, and hence that our moral reasons are always normative.
Concluding Remarks

1. Answering the Moral Sceptic

I introduced this thesis with the moral sceptic’s question: ‘Why should I be moral?’ I will now answer him in the only manner acceptable to an adherent of the Humean theory of reasons: You have reasons to be moral that are explained by the conative features of your psychology. Slightly more prosaically, I answer the sceptic in the way Bernard Williams suggests: ‘You are committed to an ethical life merely because you are an agent’ (Williams 1985, 33, my emphasis). Williams’ answer reflects the two key elements of the Humean constitutivist vindication of moral normativity. The relevance of being ‘committed’ to an ethical life accommodates the Humean idea that an agent’s moral reasons must be explained by desire-like features of his psychology, while the explanation of why you have this commitment—because you are an agent—reflects the constitutivist strategy.

Let me briefly summarise how the arguments in this thesis link together to provide this answer to the moral sceptic. For the question ‘Why should I be moral?’ to be a genuine question, the asker must have reasons to act one way or another—practical normative reasons. So the moral sceptic must be responsive to reasons. He must be a rational agent.

In Chapter Five, I argued for agential linguistic constitutivism. This view claims that merely being a rational agent entails having normative reasons to act morally. Being a rational agent requires that you are a linguistically competent being, and that requires (a) occupying the second-person standpoint, the standpoint of someone in the business of making and recognising second-personal claims and (b) undertaking a volitional commitment to give second-personal claims some weight in your practical deliberation.

To get from this constitutive commitment to give second-personal claims weight in one’s deliberations to specific moral conclusions, linguistic constitutivism piggy-backs on
Darwall’s derivation of the contractualist moral norms of ‘morality as equal accountability’ from the presuppositions of the second-person standpoint. As I outline in Chapter Four, occupying the second-person standpoint, and addressing second-personal reasons, involves a number of presuppositions, including that those who make and acknowledge second-personal claims necessarily take themselves and other persons to be mutually accountable to one another as free and rational beings. We can also make and recognise second-personal claims on one another’s behalf as representatives of ‘the moral community’. But what is demanded by us, representatives of the moral community, on behalf of each other? What is demanded of us from the moral point of view? Compliance with ‘principles that are acceptable, or not reasonably rejectable, to each as free and rational agents […] apt for second-personal address’ (300-1). Hence, the constitutive commitment rational agents necessarily undertake (to give weight to second-personal claims), combined with the necessity of occupying the second-person standpoint as a rational agent, rationally entails giving contractualist moral norms, and reasons normative significance.

So agential linguistic constitutivism explains why moral norms and reasons are genuinely normative for all rational agents—why they are not things we can rationally shrug our shoulders at. This partially answers the moral sceptic.

Can we explain why moral reasons have the universality and robust modal status that they intuitively possess? In other words, can we accommodate moral rationalism assuming the Humean theory of reasons? The intuitive existence and modal status of moral normativity are already entailed by agential linguistic constitutivism. As I argued in Chapter Two, whatever moral norms and reasons we have in virtue of our constitutive commitments will be normative for every rational agent—and modally robust—normative for every possible agent.
The moral sceptic is forced to concede that he has normative reasons to act according to moral norms and reasons. So moral norms and reasons have some normative significance. But why do moral reasons also outweigh amoral and immoral reasons? In Chapter Three, I answered this question by arguing that Humean constitutivism can accommodate strong moral rationalism. In particular, it can accommodate the Supremacy thesis that moral requirements are also rational requirements—that whenever our moral reasons demand we phi, we have most normative reason to phi.

My argument for this conclusion combined, the view that the strength of an agent’s reasons is a function of the strength of the agent’s relevant desires (and the degree to which an action will promote the object of these desires) with ideal preferentialism, a theory of strength-of-desire according to which an agent’s strongest, most reason-giving psychological states are (roughly) those that have the greatest counterfactual survivability under ideal conditions.

Since rational agents cannot abandon their constitutive commitments but can abandon their contingent commitments and desires, I claimed the reasons agents have because of their constitutive commitments always have greater weight than their other reasons. If agential linguistic constitutivism is right, and agents have normative reasons to act according to their moral reasons in virtue of their constitutive commitments, that would explain why moral reasons trump other reasons in deciding what we ought to do all-things-considered. In sum, Humean constitutivism can fully accommodate moral rationalism.

2. Postscript: What about Animals?

I have argued that all rational agents are rationally subject to contractualist moral norms and have normative reasons to act morally. This is, of course, a significant claim, but there is also an important limitation worth mentioning. These contractualist norms and
reasons may not exhaust moral normativity. If they do not, then this thesis may leave the
normativity of a significant portion of morality unsecured. Scanlon, for instance, writes:

Contractualism […] is not meant to characterize everything that be called ‘moral’ but only that
part of the moral sphere that is marked out by certain specific ideas of right and wrong, or
‘what we owe to others.’ The boundary of this part of morality marks an important moral
distinction, but it does not mark the difference between those beings who ‘count morally’ and
the rest. (Scanlon 1999, 178)

What moral reasons and obligations might contractualist morality be leaving out?
Broadly, it might struggle to explain why we have moral obligations to non-rational entities,
such as non-rational animals, the natural environment, and perhaps even non-existent beings
such as future people. Likewise, under Darwall’s contractualist framework, we might struggle
to explain why we have moral obligations to those we don’t think are a part of the ‘moral
community’ of second-personally competent beings.

The potentially limited scope of contractualist morality raises the question of whether
I have responded comprehensively to the moral sceptic. I may have shown he has good
normative reasons to act morally to other rational beings, but I have not shown he has good
normative reasons to act morally in other ways.

This limitation is doubly troubling for my account because of my endorsement of
strong moral rationalism. My account suggests that contractarian moral norms and reasons
trump all others. But if there are significant moral considerations and reasons besides
contractualist ones, then it is much less plausible that contractualist reasons are rationally
dispositive. Intuitively, some vast amount of animal suffering ought to rationally outweigh
whatever ‘we owe to one another’, in determining what we ought to do all-things-considered.

I’ll mention that the contractualist framework may yield a more impressive chunk of
common sense morality than might initially be feared. And the more morality that can be
accommodated under the contractualist framework, the more plausible strong moral rationalism becomes.

We might follow Scanlon’s suggestion that obligations to non-rational moral subjects might be accommodated by contractualism via the concept of *trustees* for non-rational beings:

[I]n deciding which principles could not reasonably be rejected we must take into account objections that could be raised by trustees representing creatures in the group who themselves lack the capacity to assess reasons’ (Scanlon 1999, 183).

Or (I suspect more promisingly) we might think that contractualist obligations and norms (e.g. norms against causing needless pain) also govern our interactions with non-rational beings. As Scanlon himself claims, ‘like the pain of humans, the pain of nonhuman animals is something we have reason to prevent and relieve, and failing to respond to this reason is a moral fault.’ (Scanlon 1999, 182). Likewise, Darwall speculates:

[W]hat we are morally responsible for might include, for example, the protection of cultural treasures, wilderness, and/or the welfare of other sentient beings, quite independently of the reaction any of these have to the interests of free and rational persons. […I]f we have such moral obligations, then these are among the things we free and rational agents have the authority to demand of one another. (Darwall 2006, 28)

The idea is that contractualist principles, norms, and reasons may govern our interaction with nonhuman animals and other non-rational beings. Or, in Scanlon’s terms, it may be that mistreating non-human animals and the environment would be disallowed by moral principles which no one could reasonably reject. I hope to explore these issues in future work.
Bibliography

Kant, I. 2020. Groundwork of the Metaphysic of Morals. Routledge, Section II.


--- 2019. How to play well with others. Games, Sport and Play, pp.74-98.


Scanlon, T. M. 1999. What We Owe to Each Other. Cambridge: Harvard University Press


--- 1998, What We Owe To Each Other, Cambridge, MA: Harvard University Press.


Snedegar, J. 2014. “Contrastive reasons and promotion” Ethics. 125, 1, October p. 39-63
------- 2009b, Review of Slaves of the Passions, Notre Dame Philosophical Reviews.