

School of History, Philosophy, Political Science and International Relations

**VICTORIA UNIVERSITY OF WELLINGTON**

*Te Whare Wānanga o te Ūpoko o te Ika a Māui*



---

# **The Social Foundations of Normative Judgment**

---

by

**Udayan Mukherjee**

A thesis submitted to Victoria University of Wellington in fulfilment of the  
requirements for the degree of Master in Arts

2014

---

## Acknowledgments

---

My sincere thanks go to my supervisor Richard Joyce. His comments and suggestions throughout the year were instrumental in ensuring that my thoughts were better-formed than they otherwise could have been. In particular, he was extremely supportive of my desire to explore a broad swathe of philosophical territory while simultaneously reigning me in when my ambitions roamed too wide. I am also grateful to Kim Sterelny for useful discussions and leads early on in the year, as well as for teaching the illuminating undergraduate courses that piqued my interest not just in his peculiar brand of empirical philosophy, but in the subject more generally. I am also grateful to the many graduate students in the department who tolerated an interloping economist chiming in on their seminars for three years. The Victoria University of Wellington Philosophy Department has been a fantastic place to think and learn.

Beyond the confines of the Murphy Building, I must thank my friends Ella, Paul, and Jodie, for tolerating my constant flights of conceptual fancy for many, many years. This culminated in their sterling editorial work in the final stretch but their overall influence on these pages is harder to quantify. Lastly, and most importantly, my parents' unyielding support for my eccentric academic pursuits has been valuable beyond words. I don't imagine that they will enjoy reading this thesis, but I have no doubt that they will slog through it nevertheless. I look forward to their feedback, both philosophical and miscellaneous, for decades more.

---

## Abstract

---

Norms suffuse our lives and are a major part of the way that we understand and structure the social world. This thesis provides an account of normative judgment that illuminates the nature of this uniquely human competence. The main argument pursued is that understanding normative judgment requires a direct and sustained understanding of its social functions. Within philosophy, discussion of normativity has often been confined to the moral domain. One major theme of this thesis is the broadening of this focus to include other domains that are rightfully considered normative. Another philosophical shibboleth is the tendency to explain features of human psychology from a conceptual perspective. A second theme of the thesis will be the insistence that empirical research is a useful addition to the project of understanding normativity. I present these ideas in three stages. First, I show why it is plausible to believe in the unity of normative domains and defend a conceptual thesis of Normative Judgment Internalism that sees norms as fundamentally bound up with reasons. Secondly, I outline a puzzle that any theory of normative judgment must answer and then critique orthodox Humean and anti-Humean theories that fail to provide such a solution. Thirdly, I explore empirical research about the nature of normative judgment and tentatively endorse a model of normative cognition that is informed by my earlier arguments.

---

## Table of Contents

---

|   |           |
|---|-----------|
| <b>Introduction.....</b>                              | <b>1</b>  |
| <b>Normative Talk and Thought .....</b>               | <b>6</b>  |
| 1.1 – The Unity of Normative Domains.....             | 7         |
| 1.2 – Norms and Reasons.....                          | 14        |
| 1.3 – Objections & Clarifications.....                | 22        |
| 1.4 – Conclusion.....                                 | 30        |
| <b>Solving the Normative Puzzle.....</b>              | <b>31</b> |
| 2.1 – The Normative Puzzle.....                       | 33        |
| 2.2 – Humean Theories of Normative Reasons.....       | 37        |
| 2.3 – Anti-Humean Theories of Normative Reasons.....  | 42        |
| 2.4 – Evaluating Humean and anti-Humean Theories..... | 47        |
| 2.5 – Shared Norms and Communal Judgment.....         | 53        |
| 2.6 – Naturalising Normativity.....                   | 59        |
| 2.7 – Conclusion.....                                 | 63        |
| <b>The Roots of Normative Cognition.....</b>          | <b>64</b> |
| 3.1 – Norms as a Social-Scientific Concept.....       | 66        |
| 3.2 – Normative Nativism .....                        | 72        |
| 3.3 – Socially Embedded Normative Guidance.....       | 77        |
| 3.4 – The Modified Sripada & Stich Model.....         | 83        |
| 3.5 – Some Evidence for the Model.....                | 90        |
| 3.6 – Conclusion .....                                | 94        |
| <b>Conclusion.....</b>                                | <b>95</b> |
| <b>References.....</b>                                | <b>99</b> |

---

## Introduction

---

“Every faculty in one man is the measure by which he judges of the like faculty in another. I judge of your sight by my sight, of your ear by my ear, of your reason by my reason, of your resentment by my resentment, of your love by my love. I neither have, nor can have, any other way of judging about them.”

Adam Smith<sup>1</sup>

“Moral philosophy has, indeed, this peculiar disadvantage, which is not found in natural, that in collecting its experiments, it cannot make them purposely, with premeditation, and after such a manner as to satisfy itself concerning every particular difficulty which may be.”

David Hume<sup>2</sup>

An outsize portion of human life is spent engaged in banter and quarrel. Some of this is trivial and domestic, like disputes about the proper arrangement of lounge furniture or who should rightfully clean the dishes on any given weekday. This can shade out into broader arguments about the civility of certain sartorial choices and the wisdom of team selectors in our favoured sporting codes. And the stakes can spiral higher and higher, as in our chronic contests for political power or in the

---

<sup>1</sup> Smith 1759

<sup>2</sup> Hume 1740

questions we ask about the appropriate ways to structure our friendships and families. My thoughts on these issues, thankfully, are not the subject of this thesis. But there is a thread that ties all these diverse enquiries together. They all seek to ask questions about the value of some forms of living, questions that can't be asked or answered without noting the way that they recur for all those around us. It is norms that structure this form of social interaction and frame the way that we understand these types of issues. Or so I will claim. What I really want to explore in this thesis is how norms operate and how we deploy them in our lives.

Within philosophy, this type of exploration has often been traced back to its dawn and framed as a Socratic Question.<sup>3</sup> The question in its most general form is put in terms of 'How one should live?', a formulation that suggests a particularly intellectualised way that it should be answered. But the continuity of this question with the everyday discussion that it grows out of is striking. I will therefore argue that the type of explanation we need isn't particularly 'philosophical', at least in a traditional sense. What I am more specifically interested in is the process of *normative judgment*. The first issue this raises is just how normative judgment is related to our norms, which of course requires an understanding of how norms relate to our social discourse. The other issue this raises is how normative judgment relates to our thoughts – what kind of psychological processes are engaged when we do think about norms and their operation. The main animating idea of the thesis, in a nutshell, is contained within its title. It is that *Normative Judgments have Social Foundations*. The main purpose of the thesis is to demonstrate how a better understanding of the social functions that underpin norms helps us better understand the nature and operation of normative judgment.

My arguments will be presented in three chapters, whose plan is as follows:

Chapter 1 is about the nature of normative talk and thought. It starts by outlining what norms are

---

<sup>3</sup> See, for example, Williams 1985: 1-6 & Gibbard 1990: 3-6

and fleshes out my contention that all of the various types of discourse I mentioned above really are of a kind. It goes on to offer an analysis of what norms are that treats them as a relational concept that we employ to evaluate human actions, thoughts and feelings. This is used to support a further analysis of normative judgment, which links them intimately with the practice of giving and asking for normative reasons. This is what I call 'Normative Judgment Internalism' and I end the chapter by considering a couple of objections to this position which clarifies it and also sets up the issues for the succeeding chapters.

Chapter 2 is about how normative judgment works. It contains arguments that lie at the heart of the project of this thesis. It starts by outlining a normative puzzle, which is a clash between two commonsense platitudes about normative judgment. This serves to focus the explanatory problem that theories of normative judgment face. Then I go on to present and critique two prominent philosophical theories of normative judgment that seek to solve this explanatory puzzle. These are Humean and anti-Humean theories, which respectively place an individual's desires or beliefs at the centre of their explanations. My contention will be that both of these theories effectively converge in their solutions, and both also fall short in providing a satisfying solution to the normative puzzle. As a result I offer two positive suggestions about the way that theories of normative judgment should proceed in light of this critique. The first is that they re-examine the nature of the commonsense platitudes in light of the social functions they imply. The second is that they embrace empirical research as a supplement and salve to enrich their ability to explain the operation of normative judgment.

Chapter 3 puts these two suggestions into action. It does so by providing a survey of empirical research on normative cognition, and defending a psychological model that emphasises its social foundations. I start by drawing out a working definition of norms as they have been studied within social science and use this to frame the search for a capacity for normative cognition. I then

examine two different accounts of how this operates, one that makes developmental claims and one that makes evolutionary claims. I show that both of these don't entirely capture the nature of normative cognition, although they do draw attention to important elements of its psychological bases. In light of this I outline a model that fares better, and defend a slightly modified version of it as providing a solid explanatory framework for normative cognition. I end by briefly arraying some evidence in favour of this model's plausibility.

Before I begin, there are a couple of points of influence and methodology that can help clarify the approach I adopt. More so than the usual reverence of philosophy towards its past, there are two important themes of my argument that are doubly influenced by David Hume, Adam Smith and their project of establishing a 'science of man' that originated in the 18<sup>th</sup> Century.<sup>4</sup> In modern terms this is often resurrected as a 'sentimentalist revival'<sup>5</sup> and so undoubtedly many of its points of influence have been absorbed by osmosis. The two I mention are merely those that I was conscious of as I proceeded and I will revisit these themes in the conclusion in light of the arguments I will have presented.

The first point of influence is my desire to explore questions about norms in a general fashion, that doesn't focus exclusively on moral philosophy and meta-ethical issues, but on what are more properly described as meta-normative issues. This is crisply distinguished by Smith as the difference between judgments of propriety & impropriety and judgments of merit & demerit.<sup>6</sup> This issue is addressed directly in Sections 1.1 and 3.1, but more generally shapes my tendency to freely adapt many discussions that are made in a moral or meta-ethical context to my purposes. The second point of influence is the empiricism of the Humean tradition and the melding of descriptive

---

<sup>4</sup> In his biography of Adam Smith, Phillipson 2010 presents an enlightening account of the connections between Hume & Smith and the development of their joint project.

<sup>5</sup> Slote 2012

<sup>6</sup> Smith 1759 talks about propriety/impropriety in Part I and then merit/demerit in Part II



psychological analysis and normative insight with little explicit distinction between the tasks.<sup>7</sup> This shapes my concerns in Section 2.5, is argued for directly in Section 2.6, and quite obviously informs my approach through the entirety of Chapter 3. My ambitions aren't quite as broad as Hume's on this count, but the general flavour of his approach is what I aim to replicate. Hopefully the benefit of this unearned loftiness is justified as I proceed.

---

<sup>7</sup> The introduction of Hume 1740 has a clear expression of this position, when he says that we should leave behind “the tedious lingering method” and “march up directly...to human nature itself; which being once masters of, we may everywhere else hope for an easy victory”

---

# 1

## Normative Talk and Thought

---

This chapter's primary function is to set the stage for this thesis by introducing and exploring its primary subject matter: norms. I examine what norms are, the role they play in human discourse, and the way that we can understand the concepts that underpin this role. The aim is to shed some light on the complexity of the ordinary practices of evaluation and deliberation that surround us. By the end of the chapter I will provide an analysis of the role that norms play in our thought that will serve as the focus of my arguments in Chapters 2 and 3.

I will proceed in three stages. First, I will explore the diversity of different normative domains and argue that despite their differences, norms of all sorts share features which unify them. This introduces one of the themes of the thesis and also prepares the way for a unified analysis of normative discourse. So secondly, I outline a conceptual analysis of norms and normative judgment. I argue that the concept of a 'norm' is an evaluative relation, and that to the extent that we accept norms this is intimately linked with reasons. In doing so I will outline and defend a version of 'Normative Judgment Internalism'. Lastly, I will respond to a couple of objections against this position, which will help to clarify my account and frame the problem that I will address in Chapter 2.

## **1.1 – The Unity of Normative Domains**

The use of the term 'normative' to describe some special brand of human discussion has been prolific in recent philosophical enquiry. It has become more of a specialised term of art than a general idea that would be easily understood outside the academy.<sup>8</sup> Of course there are rules, standards, criteria, or guidelines that might be variously referred to as 'norms' when they crop up in everyday life. But there is more to normative life than the mere occurrence of norms – a special way that norms are reflected upon and come to guide our lives. Some domains of discussion are infused with concepts of a far more slippery kind. These are not just 'fraught with ought', in the famous words of Wilfrid Sellars, but also fraught with notions of value, justification, rationality, obligation and goodness. My aim in this section is not to settle this ambiguity, for some of it is surely a consequence of the inherent difficulty that these notions give rise to. Instead, I want to characterise some features of normative talk that allow us to get a handle on the issues that we will be exploring throughout this thesis. My tentative contention here is that different normative domains share enough fundamental features that there is some mileage to be gained by analysing them as one of a kind.

Lets start with some examples. Consider all the following sentences:

- (1) I shouldn't spend too much money on beer this year otherwise it'll take me ages to pay back my student loan
- (2) If you have read all the scientific evidence on evolution then you are justified in thinking that creationism is false
- (3) The Mona Lisa is a beautiful painting

---

<sup>8</sup> Finlay 2010

(4) I really ought to be nicer to my mum given everything that she's done for me

(5) It isn't polite to chew with your mouth open

What do all these sentences have in common, if anything? There is a temptation to suggest that they really have nothing much in common with each other at all. In one sense at least, they are all about different things. (1) is a recommendation that appeals to what is best for me to do if I care about my future, something like a piece of *prudential* advice. (2) is about when it might be appropriate to hold a particular factual belief, something like an *epistemic* standard of justification. (3) is an *aesthetic* fact which aims to deliver a judgment about whether a painting conforms to standards of beauty. (4) is plausibly read as a *moral* belief about the kinds of things that the speaker should do. (5) is a statement about the *conventions* which govern civilised eating habits in Western society. Clearly then, all five of these statements are about different subjects and have different types of content. Perhaps there really is nothing which these types of statements all share.

This is too hasty, for it seems to me that all of these types of statements involve a certain kind of judgment. One of the more striking features of human mental life is the way we reflect upon our thoughts and actions. We don't just do stuff, believe things and have certain feelings, but we ask ourselves *why* so that we can *get it right*. Each of the examples above serves an evaluative function that contributes to this task. Though they might be issued against the backdrop of an implicit social understanding, they leave room for dissent and disagreement that can easily be imagined. They are all similar in the way they invoke standards, often unseen and unspoken, and use them to deliver a verdict on how someone should live. This shared pattern of evaluation is what marks all these types of statement as part of the normative discourse we take part in.

One might readily admit this though, and still resist the idea that all of our normative discourse involves some common ground. Sketching some details of these different domains shows why there

might be a *prima facie* problem with a strategy that aims to analyse them in terms of the same core concept of a norm. Moral norms might be thought to deal with evaluating the rightness and wrongness of actions. Prudential norms are typically about our abilities to discipline our actions and make those choices which are better for us. Aesthetic norms serve to give us some standards against which we decide whether things are pretty or ugly, in the broadest sense of those adjectives. Epistemic norms give us a way to decide whether our beliefs are justified or not. Conventional norms judge whether something is appropriate or accepted relative to some set of social rules. None of these are meant to serve as fully-fledged analyses of the respective domains. Rather, these are a commonsense gloss of what these domains involve when we discuss them in our everyday lives. They suffice to make the point that different normative domains don't obviously appear to share the same function. Perhaps there isn't one shared pattern of normative evaluation that licences us to think that they are species of a single conceptual genus. If this is the case, then a strategy like mine which looks to analyse all normative discourse in terms of a single concept could seem to be misguided.

My first task is to respond to this suspicion and justify the intuition that different domains of normative discourse aren't as dissimilar as they seem. To do so I'll build on some observations made by Terrence Cuneo.<sup>9</sup> He argues that there exists a 'Normative Web' that weaves together the moral and epistemic domains and so defends a framework for normative discourse that emphasises some common functional features. Two key features he highlights are the common targets of normative appraisal and the overlap in the subject matter of normative appraisal. In what follows I will explain these two observations and show how the parallel can be extended to encompass a broader swathe of normative discourse. I don't aim to provide a decisive argument in favour of this, but simply to show that there is a plausible case to be made in favour of thinking of various normative domains as

---

<sup>9</sup> Cuneo 2007

connected at a basic level.

The first observation is that there isn't any deep distinction between the entities that are appraised by different normative domains.<sup>10</sup> There are many that we could possibly normatively assess: actions, intentions, propositional attitudes like belief, non-cognitive attitudes like emotions and feelings. It is easy to suppose that there is a disanalogy between different normative domains in the kinds of entities they are concerned with. For example, it might be tempting to think that moral, prudential and conventional norms are mainly to do with our actions, what is sometimes thought of as the realm of practical rationality. By contrast, epistemic norms and perhaps also aesthetic norms seem to have more to do with what is often called theoretical rationality, governing mainly our thoughts and feelings but not our actions. But a closer glance at some of the contexts in which we use normative discourse suggests that a more ecumenical approach is more plausible.

It is easy to think of examples where norms apply to a variety of entities. If my child comes crying having fallen and grazed her leg, it makes sense to say that there is a moral norm that we ought to react empathetically towards her. If we are trying to become an elite sportsperson, it might be prudent not just to train hard, but to also believe that your competitors are training harder than you so that you can motivate yourself to stick to your regime. If we discount the critical comments of a colleague because it annoys us how perceptively they have identified flaws in our work, then this is an epistemic demerit since this annoyance is intellectually obfuscating and would prevent us from appreciating the merit of their comments. If we praise some piece of music for being particularly beautiful, we might think that there is an aesthetic norm which would hold that we ought to be moved by it and respond emotionally to that music in some way.<sup>11</sup>

These examples could be multiplied, but the basic point they demonstrate is that there is

---

<sup>10</sup> Ibid: 71-76. A similar point is also made in Gibbard 1990: 36-40

<sup>11</sup> Slote 1971 provides an argument that notions of rationality also apply to aesthetic preferences

considerable overlap in the modes of psychological response that norms apply to. This observation doesn't itself imply a similarity between normative domains. Even though there may be a set of entities that are all up for different sorts of normative appraisal, they could be playing different roles in different domains. The more suggestive inference is that the *type of appraisal* seems to be similar when these aspects of our mental life are assessed in normative discourse. In all the examples in the previous paragraph, there is some suggestion that one's actions, beliefs or feelings are all subject to evaluation from different perspectives and this seem to be the characteristic function that norms bring to bear. Granted, different specific normative domains might be more centrally concerned with some types of entities or others. But the more general point is that normative discourse seems to operate by assessing our responses and this is evidence for one basic shared feature.

The second observation relates not to the function of normative discourse but its content. Though I noted above that there are putative distinctions between the content of our different normative domains, these distinctions aren't as sharp as they seem. Notably, we often use what appear to be hybrid norms in situations where the type of appraisal doesn't clearly belong to one normative domain or another.<sup>12</sup> It can be difficult to disentangle the different dimensions of a particular norm, which indicates that there may not be as deep a distinction in normative content as we might suppose.

What are examples of these hybrid norms? Consider norms of *charitable understanding*. It is typically thought to be a failing if someone is presented with some evidence, such as the testimony of a trusted friend, which they then don't take into account when forming their beliefs. This is obviously an epistemic failing since it would mean not responding to a reliable source of evidence. But it also can be evaluated from a moral dimension as a failure to respect the relationship you have with your friend and the trust in their testimony that should bring. Or consider norms of *elegant*

---

<sup>12</sup> Cuneo 2007: 77-80

*descriptions*. We often describe some passages of prose as particularly elegant. In doing so we are commending it in an aesthetic sense, on all the same kinds of grounds that we use to evaluate good prose style, like economy of language, vivid word choice, lyrical phrasing and so on. But at the same time, part of the elegance of a description is the ability of those very same features to clearly communicate an understanding of what is being described in the reader's mind, something which is an epistemic merit.

It might be thought that these examples are merely situations where norms of two different kinds both apply, rather than a single hybrid norm. Maybe the two dimensions of appraisal can come apart. So perhaps in cases where you know your friend is unreliable you ought only to listen to them insofar as that respects the moral weight of your relationship, but you shouldn't incorporate their evidence into the beliefs you form. But even if this indicates that the norms of charitable understanding are sensitive to particular circumstances, it remains the case that there aren't separate aspects of a situation, some of which are being evaluated from one normative perspective and some from another. So even in this modified case, the very same action – listening to a friend's testimony – is still being evaluated epistemically, except now the verdict is opposite. But the interplay between the moral and epistemic dimensions is complex and being charitable involves calibrating the two dimensions against each other when deciding what to do. This is a more natural way to think of things than supposing that the situation is being independently appraised from two different perspectives. In this case, moral and epistemic norms mutually form the basis of our evaluations. And hybrid norms like these aren't entirely remarkable either. Given the complexity of human life it would be odd if the objects of our concern could always be neatly subdivided and classified. Sometimes the content of normative talk will overlap, in those very cases when the subject matter we are evaluating straddles the boundary between different normative domains.<sup>13</sup>

---

<sup>13</sup> On this point Cuneo talks of a continuum of hybrid norms. But since his discussion is restricted to moral and epistemic norms a better metaphor might be a multidimensional space of hybrid norms, where each axis represents a



To recap, in this section I have presented a couple of reasons why it is at least plausible to suppose that our normative domains are unified in some way. They were both related to the nature of appraisal that takes place in our normative talk, the first being the evaluative function and the second being the potential overlap in evaluative content. I don't intend to have proven the case for such a unity, but highlighting these areas of affinity helps us to see why there is some common phenomenon we are getting at when we speak of what it means to be 'normative'.

---

different normative domain.

## **1.2 – Norms and Reasons**

Up till now I have been trying to characterise norms and normativity in a somewhat rough way. Partly, this has been an attempt to bring the primary subject matter of this thesis into sharper focus. But also I wanted to soften the ground for my next task, which is to provide a more rigorous analysis of norms and the role they play in human thought. The previous section was about normative talk, and the unity that I believe it possesses. This section is about normative thought, how it operates and the role that it plays in our deliberations. The main contention I will defend is that there is a necessary connection between normative judgments and normative reasons, and that is a very natural consequence of the way we might understand norms themselves.

### **1.2.1 – An Analysis of Norms**

The first issue to address then, is how to understand the concept of a 'norm'. To do so I will offer an analysis of the concept that more precisely captures some of the notions I was discussing in the previous section, and that I believe provides a strong basis for investigating what is going on when we engage in normative discourse. The analysis is that a norm is a structural relation, that delivers a verdict about the correctness of a response relative to an evaluative framework. Here is a statement of the analysis:

*Norm:* N is a norm, iff it gives a correctness value C to some response R relative to some framework of evaluation F.

This is a fairly bare analysis, but this is deliberately the case. The relation aims to specify just how norms function, and the essential elements that are involved in the appraisal that is inherent in our normative talk. These elements – abbreviated as C, R & F – are presented as open variables to

respect the unity of different normative domains and not to presuppose anything that is specific to some domain or other. To be clearer on what this analysis amounts to, we should consider each element in turn.

The response *R* stands in for the particular mental entity that is evaluated by the norm – the thought, action or feeling which is up for appraisal. Any given norm might only evaluate one or another of these responses, but there needn't be a presumption in favour of associating a particular type of norm with a particular mode of response. But it is surely something that an individual does, feels, or thinks that is at stake in normative discourse. I assume here that acts, thoughts and feelings are different sorts of entities, but this isn't necessary for the purposes of this analysis. All that matters is that I am in the right ballpark when it comes to the way that norms evaluate our mental lives. The role *R* plays in the analysis is simply that it identifies the target of this evaluative function, while leaving it open just what this target is for any particular norm.

The correctness value *C* stands in for the verdict a norm delivers on a response – what might be thought of as the outcome of the normative appraisal. The potential scope that these correctness values can take will vary between and within normative domains. For example, conventional norms of etiquette might judge an action as either polite or not, with no middle ground verdict possible. But in other cases there need not be such bivalence, such as in the delicate shades of judgment we might expect from aesthetic norms about relative beauty. So different normative domains will adopt different vocabularies to express the range of *C*, but the role it plays in the analysis is to capture the way that a response is evaluated. Norms get their bite from actually taking a stand on the appropriateness of our psychological responses and so this element needs to be accounted for in our analysis.

The framework of evaluation *F* is the set of standards against which a response is assessed – what

we can think of as the substantive content of a particular norm. Consistent with my observations above, this might be a content specific to a particular normative domain or content which is hybrid and stands at the intersection of more than one domain. The importance of F in the analysis is that it captures the idea that norms can only evaluate in the context of some rules that apply in this particular instance. This backdrop is what we can consider the framework of evaluation, and it might range between explicit written codes, implicit social understandings, or unconscious biases and tendencies. Just how these frameworks actually operate is an interesting question and one that I will have more to say about later. But here the point is simply that there must be *some* framework involved in norms and so F stands in for that in the analysis.

This analysis is meant to provide an account of a concept that underpins all of the ordinary normative talk that involves language about ought, should, value, goodness, justification and the like. Of course, in any actual sphere of human discourse norms are not discussed in just this way. The aim of providing such a stripped down analysis is that it makes a minimal claim at a high level of abstraction that allows us to focus issues of general interest raised by the phenomenon of human normativity. So with this analysis in hand, we can stipulate the meaning of some of the vocabulary I have been using so far with a bit more precision. The basic normative relation is  $N(R,C,F)$ . A normative statement is any sentence which either explicitly states a normative relation, or implies the existence of one. Normative discourse is the sphere of human talk that involves normative statements, though it needn't be restricted to them. A normative domain is a subset of normative discourse where the evaluative framework is the same, or can also be expressed as the set of normative statements where the value of F in the normative relation is held constant.

### 1.2.2 – Normative Judgment Internalism

This moves us to the second issue that I want to address in this section, which is the nature of our normative thought. Having given an analysis of what norms are, I want to now consider the conceptual role they play in our deliberative lives. As I mentioned above, I will defend a version of Normative Judgment Internalism (NJI). Very broadly construed, NJI makes a claim about an essential – or 'internal' – connection between normative judgment and practical reasoning. My particular version of this claim is that if someone makes a normative judgment, they necessarily commit themselves to the existence of a corresponding normative reason. Right from the start I should make clear that my version of NJI is not concerned directly with questions of motivation and the forming of practical intentions, even though other versions of NJI take this to be a central question. My focus here is on the nature of normative deliberation, not the consequence of that deliberation for practical action. I acknowledge that this further connection is important and in the next chapter I'll explore it directly. So with this issue flagged, I will present my version of NJI, first by clarifying what I mean by normative judgment and normative reasons and then by defending the internal connection of the two in light of my earlier analysis of norms.

I take a 'normative judgment' to be the psychological state that occurs when an individual sincerely accepts a norm. Here I am following the description of this state offered by Allan Gibbard. He calls it a “syndrome of tendencies towards action and avowal” that is characterised by dispositions to defend normative statements in discussion, work out a coherent position in one's mind about these norms and expose oneself to demands for consistency in the way we do so.<sup>14</sup> Accepting a norm in this sense involves a pattern of linguistic and mental behaviour that is typical of normative thought since it incorporates an array of the phenomenon that typically accompanies deliberation about norms. So I am using the term 'normative judgment' to stand in for whatever happens in an agent's

---

<sup>14</sup> Gibbard 1990: 71-75

mind when they sincerely believe normative statements, utter normative sentences to express these thoughts, or do both. Exploring further just how this psychological state works is the main focus of the next chapter, but for now this suffices.

I take a 'normative reason' to be a consideration that justifies our choices as we deliberate about what to do. These are what we appeal to when we reflect about what counts in favour of various actions we could take, feelings we could have, and beliefs we could hold. The terminology surrounding reasons is often muddled in the literature, so a couple of clarifications are in order.<sup>15</sup> First, normative reasons are distinct from motivating reasons which might be invoked to explain why, as a matter of fact, people behave as they do. Secondly, by normative reasons I mean reasons that justify doing something in the objective sense, as distinct from reasons that might happen to be those that an agent has for doing something in the subjective sense. So for example, when someone makes a decision to embezzle some money, they might themselves think the reason they are doing this is because they deserve the money, while the best explanation for their behaviour might be that there weren't enough systems of oversight and punishment to deter them. Nonetheless, we could still claim that there is no reason, in the objective sense, that would justify this behaviour. It is this last sense that I am intending when I make claims about normative reasons.

Given these understandings, NJI should be understood as the claim that if an agent makes a normative judgment (about themselves or about another), then they are making a commitment about a normative reason that they believe should be present in deliberation (of themselves or another). This connection is conceptual, so I am claiming that making a normative judgment *necessarily entails* endorsing a normative reason. An agent can't do one without the other. My account of NJI shares close affinities with the account of 'Inner Moral Judgments' offered by Gilbert Harman.<sup>16</sup> An

<sup>15</sup> I have followed Scanlon 1998 & Schroeder 2007 here, who offer very similar definitions of 'normative reason' despite arguing for very different conclusions about them, as we shall see in the next chapter. The basic distinction between justifying and explanatory reasons can be traced to Frankena 1958

<sup>16</sup> Harman 1975

agent sincerely accepts a normative statement when they judge that a normative relation  $N(R,C,F)$  applies to some subject (or subjects)  $S$ . When they do so they thereby endorse a normative reason in favour of  $S$  having the response  $R$ . Note that  $S$  might refer to the agent making the judgment (I should), it might be someone else (he/she should), it might be everyone (we all should), or it might be some group of people (they should). All of these are possible in different instances of normative judgment. But NJI implies that in all cases, making a normative judgment implies an endorsement of a normative reason for the relevant subject or subjects. So why believe that the analytic connection asserted by NJI true? I believe the best answer is that norms implicate a certain model of human psychology, which in turn is intimately connected with a deliberative perspective. I will trace this argument in the next couple of paragraphs.

The definition of normative judgment leans on a description of the conceptual role that norms play in our thought. But recall that I analysed norms themselves as being constituted by a relation that already involves elements of human thought, namely the responses  $R$  that are the targets of the evaluative function of norms. These responses are the actions, feelings and thoughts of human agents. A folk psychology that includes these elements inherently involves a commitment to interpreting human behaviour in terms of these mental entities. This is sometimes explained as adopting “the intentional stance” in explaining the behaviour of an agent, which means that by deciding to treat an agent as a rational creature we are assuming that they make decisions in light of these elements of their mental life.<sup>17</sup> The significance of this point is that on the line of reasoning advanced earlier, normative judgments are wedded to a certain interpretive perspective in their appraisal of human behaviour, one that adheres to a certain model of intentional psychology. But so far this doesn't bring out the connection between norms and deliberation that is at the heart of the NJI thesis.

---

<sup>17</sup> Dennett 1987

This further connection is best summed up by the slogan of Ralph Wedgwood, who argues that “the intentional is normative”.<sup>18</sup> The idea is that the explanatory perspective of intentional psychology implicitly involves an evaluation of behaviour which is itself normative. Imagine that we make a prediction about an individual's behaviour like 'Sue is going to the beach today'. To explain this prediction in intentional terms we would normally say something like 'Considering that: Sue has packed a bag with sunscreen, a towel, & a beach ball; Sue likes getting a pleasant tan; Sue was pleased by the good weather forecast for today; it makes sense that she is planning a trip out to the beach'. This is still just an explanation, but the reference to what 'makes sense' for Sue is one which implies that she is the kind of agent who has a disposition to make rational decisions in light of their mental states. An adequate account of this causal explanation is thus partly constituted by a normative appraisal – we are asking about whether an agent's behaviour 'makes sense' in two senses.<sup>19</sup> So if norms are wedded to an intentional stance, this also weds them to a normative stance, where agents' decisions are interpreted not only in terms of what best explains them, but also what justifies them. So if it makes sense for Sue to go to the beach today, then this means both that we are predicting that she will *and* that we think she has a reason to do so. Another way to explain this point is that in human mental life the causal structure of explanation is isomorphic with a rational structure of normativity, a notion which supports the claim that there are intimate connections between the two perspectives.<sup>20</sup>

What is the relevance of this brief excursion into philosophy of mind? It provides descriptive support for my analysis of norms, that until now has been merely assumed. The analysis of norms I offered was that they were relations of responses and correctness, against an evaluative framework. The real commitment of this analysis is now clearer: I am saying that norms are the mechanism we

---

<sup>18</sup> Wedgwood 2007

<sup>19</sup> Ibid: 154-165

<sup>20</sup> Blackburn 1998: 51-59 makes this point clearly, leaning particularly on Davidson 1963.



use to make sense of mental life, both when interpreting the behaviour of ourselves and others and evaluating that behaviour. Furthermore, we can use this analysis to explain why NJI is true.

Normative judgments were defined in terms of accepting norms, and norms are conceptually about the evaluation of human mental lives. It follows that normative judgments are necessarily about considerations that play a role in human deliberation, which, by hypothesis, were what I called normative reasons. This means that normative judgments must be connected to commitments about normative reasons. This argument in favour of NJI is an explanation of the fairly mundane thought that when we reflect on the acts, thoughts and feelings of ourselves and others, we are deliberating about the reasons which make them make sense. And norms are the cognitive tool that we use to do this.

In this section, I have proposed an account of the fundamentals of normative thought. I firstly gave a analysis of the concept of a norm that said its purpose in human discourse was at heart a relational one. I then introduced and defended a thesis about normative judgment, which was that it necessarily involved making commitments about reasons that weigh on human deliberation. This has thrown up many questions that deserve more attention. In particular, in the next chapter I will move from describing what normative judgment *is*, and investigate theories of how normative judgment *works*. Before concluding this chapter though, I want to consider some objections to what I have said, and use them to clarify the scope of what I am defending.

## **1.3 – Objections & Clarifications**

This section considers some objections to what I have said so far. In particular I want to consider a couple of objections to the account of normative judgment I outlined in the previous section. The first objection is to straightforwardly deny NJI, and reject the idea that accepting a norm necessarily requires an endorsement of a normative reason. We might call such a view ‘Normative Judgment Externalism’ and I will explain and evaluate this argument to show why it doesn't pose a serious threat. The second objection can be understood as the concern that I haven't shown how or why the objective reasons I am talking about should matter. The potential problem here is that the relativity built into my analysis prevents reasons from applying to our deliberations in the way that reasons actually do. I will acknowledge that this is an important issue, but use it to clarify the nature of the hypothesis I am proposing and to focus in on the issues that are the primary concern of the next chapter.

### **1.3.1 – Against Normative Judgment Externalism**

The denial of an internal connection between normative judgments and practical reasoning has a long history, particularly when it comes to the domain of moral judgment.<sup>21</sup> Part of this though is due to a failure to disambiguate normative reasons from motivating reasons, and arguments which say that normative judgments don't necessarily motivate are not responses to NJI as it I have defended it. But even if we restrict our attention to normative reasons, it is fairly common to encounter arguments which deny a necessary connection between normative judgments and reasons.<sup>22</sup> A popular strategy used by these sorts of arguments is to challenge the analysis of NJI

---

<sup>21</sup> For example, the issue is discussed in Falk 1947; Foot 1972; Gert 2001; Zangwill 2003

<sup>22</sup> Railton 1986 is one prominent example

directly by saying that it is possible to imagine an agent who truly makes a normative judgment but denies that this has any bearing on the normative reasons they are committed to endorsing. This strategy is often allied to a position which I will call 'Normative Judgment Externalism' and it is my goal here to explain this position and then show that it doesn't stick against NJI. David Brink poses the challenge particularly clearly so I will focus on his version as a foil.<sup>23</sup>

Brink poses his challenge in the context of moral normative reasons. His claim is that there is indeed a connection between morality and reasons, but internalism is not the best way to explain this connection. This is because we can imagine an amoralist, who makes moral judgments but simultaneously evinces a certain indifference to deliberative consequence of these judgments. The amoralist simply lacks the extra psychological feature, whatever it might be, that links moral judgments to commitments about moral reasons that do or don't justify one's conduct. The problem for the internalist is that they must regard the amoralist as making a conceptual error. The question 'Do I have any reason to act in accordance with my obligations?' has to be an incoherent one to ask. As Brink says, "the thought that someone might possibly not have good reason to act on his moral obligation need not force us to withdraw our ascription of obligation".<sup>24</sup> Since the amoralist is intelligible, it doesn't matter how realistic the character is. What it serves to demonstrate is that something must have gone awry in the conceptual analysis of internalism for it cannot make sense, even theoretically, of someone posing this simple question.

But couldn't the internalist just deny that the amoralist has really identified her obligation correctly? Brink calls this the inverted-commas response which is that the amoralist is using the language of moral obligation incompletely. On this response, when the amoralist says 'I have a moral obligation to  $\phi$ ' this should be glossed as 'I recognise that others regard me as having a moral obligation to  $\phi$ '.

---

<sup>23</sup> Brink 1989: 57-62

<sup>24</sup> Ibid: 59

But Brink argues that this response merely stipulates that the amoralist must be making this kind of mistake. This isn't an adequate response to the challenge of the amoralist because the internalist is trying to argue that their position is a consequence of the correct analysis of moral obligation. The inverted-commas response merely asserts the challenge away rather than defending the analysis and so is a “facile solution to a traditional philosophical and popular problem”.<sup>25</sup>

Though Brink poses this challenge in the moral domain we can easily reconstruct a version of this argument as it applies to normative judgment more generally. Imagine a 'Normative Knave' who sincerely accepts a normative statement, yet denies that this commits them to the existence of any corresponding reason. They can confidently claim to be making normative judgments in that they sincerely accept norms and are prepared to engage in full-fledged discussion to defend their position. But it is still possible for them to ask ‘Do I have any reason to act in accordance with the norms I accept?’. NJI must regard this knave as being conceptually mistaken, or respond by claiming that they are only accepting norms in an inverted-commas sense. But to insist that this is the case by fiat is a facile response to an interesting challenge. So the intelligibility of the normative knave counts as evidence that NJI has overstated the connection between normative judgment and normative reasons.

I think Brink is right to insist that this is a challenge, one that can't be met by simply assuming that the knave is conceptually confused. But he is wrong to argue that the mere intelligibility of the knave counts against NJI. For what is at stake between both sides in this argument is what it means to have mastery of normative terms. The internalist claims that the knave is not a competent user of normative language, while Brink insists that they are. As Michael Smith points out, Brink is making just the same kind of assumption about the case that he accuses the internalist of.<sup>26</sup> He assumes that

---

<sup>25</sup> Ibid: 59

<sup>26</sup> Smith 1994: 68-71

being able to reliably engage in normative discussion is a sufficient condition for the sincere acceptance of norms, and that being committed to the existence of normative reasons is not a necessary consequence of this. So what we have is a debate about the conditions under which an agent has mastery of normative language. But this isn't an entirely balanced debate. Soren Svavarsdottir points out, rightly I think, that on methodological grounds the burden of proof should fall on the internalist.<sup>27</sup> This is because it is the internalist who is trying to rule out an explanation of the case while the externalist is merely offering one possible hypothesis about it. So to defend NJI we need an argument which meets this burden and would lead us to think that that the normative knave would only be "accepting" normative statements, and not really accepting them.

To meet this challenge we can start by adverting to the argument I offered at the end of the previous section. This was that bound up in the very concept of a norm are two intertwined perspectives for explaining human mental life – the intentional and the deliberative. The possibility of the knave is a problem because it would suggest that these two perspectives can come apart. They would be able to engage in a thorough evaluation of an agent's behaviour from the intentional stance, discussing what we would expect someone to do given their mental states and make fine-grained predictions about these just as competently as anyone else. Let's think about what this would mean in reality though.<sup>28</sup> Say that the knave is talking to Anne about the case of Sue's beach trip, as I discussed above. By hypothesis the knave would agree that Sue is going to beach on the basis of what he knows about her various thoughts, acts and feelings as specified earlier. So in one sense the knave could agree with Anne that Sue should<sub>p</sub> go to the beach today, where this is understood as a should of prediction in the intentional sense. And in another sense they would disagree with Anne that this means that Sue should<sub>p</sub> go to the beach in the deliberative sense that implies a normative reason. But these results are odd. Not because disagreement on these matters is impossible, but because of

---

<sup>27</sup> Svavarsdottir 1999

<sup>28</sup> The following case is adapted from Gibbard 2003: 154-158

the type of disagreement we are positing between Anne and the knave. Anne would surely find it bemusing that the knave agrees entirely with her about what Sue should<sub>p</sub> do but disagree on what she should<sub>D</sub> do. Her natural response would be that the knave just doesn't quite get it, that there isn't any difference between the two shoulds, and the knave must be actually making a mistake in their interpretation of Sue's behaviour.

This might give the impression of a browbeating argument, but it reveals something about the nature of the disagreement. Internalists insist that the knaves don't have complete mastery of normative language because the scenarios like the one described above seems to indicate that they are misinterpreting the conversation that they are a part of. The fact that knaves would insist that the two different interpretive perspectives can come apart indicates that they don't accept the analysis of what a norm is in the first place. So the dispute isn't really about the nature of normative judgment, but the nature of norms themselves. But in this case NJI fares better against the intelligibility of the knave, since I offered a positive argument for this analysis of the concept of norms. In effect this pushes the burden of proof in this dispute back on the externalist. It isn't enough merely to posit the intelligibility of the knave, since this merely elicits two sets of competing intuitions about the case. The externalist owes us an alternative and plausible analysis of norms that can explain how and why the intentional and normative perspectives can come apart in evaluation. This doesn't decisively banish the spectre of the normative knave, but it at least provisionally vindicates my version of NJI against his potential perfidy.

### 1.3.2 – The Relativism of Reasons

Even if one accepts the necessary connection of normative judgments with normative reasons, it is still possible to resist my version of NJI from another direction. This would be by taking issue with

the way that I have cast normative reasons themselves. My definition itself is fairly orthodox, but it could be thought that given the way that I explain the nature of norms, the reasons that fall out of my account are problematic. The source of this worry is that there is a tension between the way that I describe normative reasons as objective while simultaneously basing them on an analysis of norms that involves relativity to standards of evaluation. Put simply, maybe my normative reasons are too wimpy, since I haven't shown how they would actually apply in the deliberations of agents. What I want to do now is to refine this worry and show why it is important. I won't respond directly to this worry though, but use it to tighten my analysis and clarify the scope of the argument of the next chapter.

To introduce the worry that I am considering, we can begin by looking at a classic distinction made by Philippa Foot.<sup>29</sup> The distinction comes in the context of an investigation of the Kantian notion of the categorical imperative, which are those sets of commands that apply to all agents regardless of their ends.<sup>30</sup> Foot argues that these can be understood in two different ways. One interpretation is that categorical imperatives can't fail to apply to an agent, even if they reject that these imperatives have any impact for them. So even though the loutish university student might not care about the rules of etiquette, we don't therefore conclude that etiquette fails to apply to them. A second interpretation though is that categorical imperatives bring about a reason that an agent cannot rationally ignore. The example here, of course, is morality, which ostensibly binds us to follow its dictates even if we don't care about following its rules. The distinction Foot draws attention to is that in both cases, the imperatives might categorically apply to us, but it is only in the latter case that the imperatives are categorically reason-giving. The problem she subsequently raises is that in the latter case there is a "fugitive thought" about just how to make sense of an agent having reasons that mean nothing to them.

---

<sup>29</sup> Foot 1972

<sup>30</sup> Kant 1785

Even though I am not engaged in a discussion about the nature of morality, Foot's observation still has relevance here. Earlier, I stated that by normative reasons I meant *objective* reasons, as distinct from *subjective* reasons. This is similar to the distinction that Foot makes. Objective reasons are considerations that can be used to justify an agent's behaviour, so normative reasons can't fail to apply in one sense. But at the same time, I explicitly disavowed the idea that these necessarily bring an agent subjective reasons that they might respond to, so normative reasons aren't necessarily reason-bringing in another sense. But herein lies the potential worry. If the normative reasons implicated by NJI only apply *to* agents, rather than apply *for* them, then it might be thought that my account suffers as a consequence. If all I have explained is that normative judgments imply reasons that are mere curios but don't need to actually engage the thoughts of the agents they are about, then perhaps my account is too weak in a crucial sense. After all, the phenomenon we are trying to explain is just how norms play a role in human deliberation. It is 'real' reasons that matter in our thoughts, not the 'wimpy' ones that I have focused on.

This potential 'wimpiness' critique is on the mark. Nothing I have said so far would forestall a 'so what' objection from an agent who acknowledged that normative judgments implied reasons, but then found that this left them cold. In short, I haven't given a reason why reasons matter. The nub of this objection is that I have missed something important about the special authority that reasons-talk has in typical human deliberation. Answering the challenge it raises will be the primary focus of the next chapter, where I will investigate various theories of normative judgment that seek to explain just how they account for the way that normative talk and thought engages us. All I want to note here is that this objection has force only insofar as it points out that my account is incomplete, rather than in diagnosing any particular flaws internal to NJI. So to conclude this chapter in light of this critique I want to mention two points that clarify the specific hypothesis I am making about normative judgment.



The first is that a useful way to think of the kind of reasons implied by my account is that they are institutional reasons. This follows a point made by Richard Joyce about the way that we can easily make sense of reasons-talk if we consider that judgments are made from inside a system of normative rules, as if we are a “mouthpiece” for them.<sup>31</sup> When we do this we implicitly prefix our claims about reasons with an 'According to...' modifier. So say we make a normative judgment that implies that S has a reason to refrain from ordering that third martini. This is best understood as if we had begun by saying 'According to the norms of prudence...' before proceeding with the rest of the normative judgment. This implicit modifier is a good model of what 'relative to some framework of evaluation F' means in my original analysis of norms. I am saying that normative judgments are made 'from within' a normative framework, so they deliver us normative reasons that are tied to these institutional frameworks. This means that the normative reasons are clearly relative-to these frameworks. So secondly I want to clarify that in saying that normative reasons are institutional I am not saying that they are relative to something distinctive about the individual making the normative judgment. The 'according-to-F' modifier means that I am committed to saying that normative reasons are end-relational, but this needn't be the ends of the agent making the normative judgment.<sup>32</sup> The ends are standards of evaluation that are independent of any individual. They are features of norms rather than persons, though of course an individual might also share these ends. To be very clear then, normative reasons on my account are relative to institutional frameworks of evaluation. What remains to be explained is just how these reasons can engage our deliberations, rather than merely apply to them.

---

<sup>31</sup> Joyce 2001: 39-42

<sup>32</sup> A point made in Finlay 2009

## **1.4 – Conclusion**

This chapter did three things. First, it explained what norms are and offered some reasons to think that the variety of different domains that all seem to be normative share some core features. This is what legitimises my approach in this thesis to analyse normativity in general. Secondly, I gave a conceptual account of norms which I argued captures those core features well, and then outlined an account of normative judgment in which that analysis plays an important role. The crucial point here was that the notion of a reason is fundamentally implicated in the way we interpret the world. Thirdly, I defended my Normative Judgment Internalism against two objections which clarified that I still hadn't done much to explain just how normative judgment engages the psychology of human agents. I argued that this doesn't undermine the conceptual points I made in this chapter, but it does demand that I answer this concern. The next chapter turns to how we can make sense of this task.

---

## 2

# Solving the Normative Puzzle

---

In the previous chapter I presented my account of what normative judgment *is*. In this chapter I will explore how normative judgment *works*. The aim is to answer the question that I ended the last chapter with – just how do normative reasons actually engage us in our deliberations about what to do? Philosophical answers to this question are thick on the ground, and it would be brash to claim that what I have to say here is conclusive. Nevertheless my position can be stated baldly: philosophical theories of normative judgment need to be informed by an account of the social processes of human psychology. This contention is really the heart of my thesis, and it both follows on from my claims in Chapter 1 and informs what I will say in Chapter 3.

I will argue for this contention in six stages. I begin by outlining what I call the normative puzzle, which is an apparent clash between two intuitive folk theses about the nature of normative judgment. This frames the explanatory task I am interested in contributing to and sets the table for this chapter. In the next two sections, I provide an exposition of two prominent and contrasting theories that have long been thought to be the leading candidates for a solution to the normative puzzle – that are commonly thought of as Humean and anti-Humean strategies. After that, in the fourth section I will argue that despite what appear to be fundamental differences, these two strategies display an interesting convergence which exposes a common flaw: they assert rather than explain an account of rationality. In section five I argue that conceptually, the task of explaining normative judgment would be well served by taking account of the fundamentally communal nature

of normative judgment. Then lastly, I argue that one way to avoid this flaw is for philosophical theories to become more closely acquainted with empirical research.

## **2.1 – The Normative Puzzle**

There is a puzzle which arises when we think about normativity.<sup>33</sup> It arises from the clash of two platitudes which each seem to capture fairly unproblematic ideas about the role that normative reasons play in our discourse. On the one hand, when someone makes a normative judgment we think that they are committing themselves to reasons that are capable of being true or false. On the other hand, we tend to think that these reasons should bring about corresponding practical motivations for those that they apply to. We can call these features the factualist platitude and the practicality platitude.

The factualist platitude is an indication of the apparent semantics of normative talk. When we make a normative judgment in public we appear to be trying to state facts about those judgments. So for example, to say that we judge that it will rain tomorrow is to make a commitment about the truth of that judgment. This is a commitment that might be challenged. For example, a friend might say that our reason for this belief is false because the weather channel we rely on is notoriously inaccurate in its predictions. Engaging in debate about the truth of claims about reasons is an important feature of everyday normative discourse. So assertion seems to be one of the functions of normative judgment and this is why we can say that a factualist platitude holds.

At the same time, it appears that normative reasons have a reliable connection with our attitudes as well. If someone is aware of a reason that applies to them, it normally seems like they are automatically motivated to follow through on these reasons, at least to some extent. This is the practicality platitude. If we have a reason to believe it will rain tomorrow, this induces a certain conviction in that belief, one that we incorporate into our practical decision-making. It would be bizarre if we didn't reassess our picnic plans for tomorrow on the basis of this belief. When we

---

<sup>33</sup> The way that I frame the puzzle in this section owes a great deal to the discussion of the 'Moral Problem' by Smith 1994:1-15. The same issue is also discussed fruitfully in Brink 1997 in response to Smith.

make normative judgments, the implied normative reasons are 'motivationally efficacious' – a way of saying that they make a difference to our plans and intentions about what to do. This too is an important feature of normative judgment and is why we can say that a practicality platitude holds.

So why do these two platitudes clash? It is because they have implications that are in tension with each other. The factualist platitude says that we can be right or wrong when we make normative judgments, since our normative reasons can be true or false. This implies that there is a fact of the matter about these things, and so our normative judgments might seem best interpreted as beliefs about those facts. The practicality platitude says that normative reasons have an impact on our motivations so it is natural to suppose that they are related to the things that we want and the psychological states that are linked with motivation. This implies that normative judgments are about our desires. So for normative judgments to satisfy both of these implications they would have to simultaneously express both beliefs *and* desires. But according to a traditional conception of philosophical psychology, beliefs and desires are very different states.<sup>34</sup> Beliefs are about representing the world and desires are about getting the world to be the way we want it. This point is often explained in terms of a metaphor of “direction of fit” – beliefs are faulty if they don't match the way the world is, but the world is faulty if it doesn't match the way that our desires are.<sup>35</sup> This is why there is at least a *prima facie* tension between the factualist and practicality platitudes. Is this tension really a problem? In one sense not at all, since it is clear that we do behave in ways that vindicate both platitudes simultaneously. After all, that is why we describe them as platitudes in the first place! Obviously the standard conceptual story about human psychology is too simple and needs to be more complex to fit our observations about normative discourse.

The normative puzzle is just how this more complex account of normative psychology ought to be

---

<sup>34</sup> The locus classicus being Hume 1740: 2.3.3 in the section “Of the Influencing Motives of the Will”

<sup>35</sup> Zangwill 1998 discusses the issue well, and the idea traces back to Anscombe 1957 although she never used the term itself.

fleshed out. We can think of each platitude as imposing a corresponding constraint on this analytic task. The factualist platitude can be thought of as implying an advisory constraint, so called because we want to be able to make sense of the idea that people can be wrong about their normative judgments, and others might want to let them know by offering them advice about how to get it right. The practicality platitude can be thought of as implying an explanatory constraint, so called because we should be able to explain an agent's behavior as a result of the normative judgments they make if they are connected with motivation. It isn't necessary that we accommodate each of these constraints when we go about trying to explain normative psychology. In the vast literature about the foundations of morality, many other options are on the table. For example, realists often do away with the explanatory constraint,<sup>36</sup> and expressivists deny the advisory constraint.<sup>37</sup> Alternatively, we could accept the applicability of both constraints in explaining normative psychology, but argue that contrary to common sense one (or both) is never satisfied. An error theorist might say something like this about the advisory constraint.<sup>38</sup>

All these kinds of approaches are revisionary. They take our understanding of the factualist and/or the practicality platitudes and argue that since they don't stand up to philosophical scrutiny we should change the way we think about our normative discourse. In this chapter my aim isn't to assess any of these types of approaches to explaining normative psychology. Instead my aim is to explore solutions that take our intuitions about normative discourse as a given, and so try and navigate a solution to the puzzle under the constraints of the factualist and practicality platitudes. We can label these types of theories as non-revisionary as they attempt to explain normative judgment in a way that vindicates both of the common-sense platitudes about the way normative discourse works. Partly, this is a convenient way to confine my discussion to a manageable scope.

---

<sup>36</sup> Railton 1986; Sturgeon 1985; Brink 1986

<sup>37</sup> Ayer 1936; Blackburn 1993; Gibbard 1990

<sup>38</sup> Mackie 1977; Joyce 2001

But additionally, taking the platitudes as basic data is consistent with my broader approach to explain norms in deference to ordinary observations about the way they operate. At the end of the last chapter I identified that my analysis of normative judgment was incomplete, since I didn't explain just how normative reasons were able to engage us psychologically. Examining non-revisionary theories is the way that I will go about filling in that gap.

In light of this, there are two particular positions I will explore. One of these is a Humean position, which argues that our reasons are dependent on desires and so our normative judgments are about how well certain responses satisfy those desires. The second is anti-Humean, which argues that our normative judgments are best thought of as beliefs about whether or not certain normative reasons are true relative to some standards that determine their truth. Since both of these approaches are non-revisionary it means that there is a descriptive burden that needs to be discharged by either of these approaches. If the conceptual analysis they offer doesn't square with our ordinary intuitions, then this counts against them. This is an important piece of context to keep in mind, as my main contention will be that these approaches do not discharge this descriptive burden. Despite the fact that they offer sophisticated accounts of the operation of normative judgment, they fail to explain the platitudes they set out to. To argue why this is the case, I must first outline the respective argumentative strategies of the Humean and anti-Humean theories, which I do in the next two sections.



## **2.2 – Humean Theories of Normative Reasons**

The first solution to examine is one where normative reasons are explained in terms of the desires of agents.<sup>39</sup> The intuition underlying this position is simple. The idea is that there are basic psychological dispositions which ultimately determine whether an individual has a given reason. So if Anne likes ice-cream and Bob doesn't, then Anne has a reason to go and buy an ice-cream which Bob lacks. If Carl wants to improve his maths ability but Deidre doesn't, then Carl has a reason to believe the teacher's arithmetic shortcuts which Deidre lacks. If Edward cares about the art of film-making and Fran thinks that movies are just a frivolous waste of time, then Edward has a reason to admire the director of a great film which Fran lacks. These simple sorts of cases are meant to serve as a model for how all normative reasons work. The idea is that reasons are to be explained by an agent's psychology in just the same way that the difference between all the preceding pairs of agents is explained. More formally we might say:

There is a normative reason to  $\phi$ , in case  $\phi$ -ing will further the object of some desire of the agent.

What does this analysis amount to? It means that for there to be a normative reason to have some response  $\phi$  (an act, thought or feeling), it must be that  $\phi$ -ing would help satisfy a desire that the agent has. Desire is a label for the psychological state, whatever it is, that is capable of putting our body and mind in motion. Desires involve a disposition to be motivated, where the object of a desire is the particular goal which accompanies the disposition, the 'end-state' which the motive aims at. A theory like this is often called a Humean Theory of Reasons.<sup>40</sup>

<sup>39</sup> In this section the position I describe as 'Humean' is attributable to many thinkers, regardless of the particular terminology they may adopt. So Williams 1981; Lewis 1989; Schroeder 2007 & Finlay 2007 all quite avowedly defend the connection of reasons with desires, while Joyce 2001 defends what he calls a "non-Humean instrumentalism" that is similar in its essentials.

<sup>40</sup> Though it is unclear the extent to which Hume himself would endorse this position, the label has stuck due to the sense that this theory of reasons is vaguely inspired by his arguments that moral principles must be based on sentiments, in Hume 1751

In making sense of the normative puzzle it is very easy to see how this kind of theory accommodates the explanatory constraint. All of the work is done by picking out a psychological state which is intimately connected with motivation. If a normative reason is always connected with our desires, then when we make a normative judgment it is very obvious why we would be motivated by it, since we are essentially making a claim about what we want to do. What is harder to understand is how such a view would make sense of the advisory constraint on normative judgment. For it is odd to say that an agent might be making incorrect normative judgments, if these are just about how to satisfy their own desires. The obvious thought is that people want what they want and others aren't well-placed to pronounce them mistaken on this front. But this seems to imply that when discussing normative reasons there isn't really much that is up for discussion in the first place. This is the serious challenge that a Humean theory faces and to meet this objection, theorists typically refine their theory in two ways.

The first refinement is that normative reasons can fail to motivate us, if we aren't instrumentally rational. The notion of instrumental rationality is meant to cover cases in which  $\phi$ -ing will in fact further the present desires of an agent, but the agent doesn't know why and how this is the case. Adding this stipulation is tantamount to claiming that normative reasons succeed in motivating us insofar as we are rational. The effect of this is that it licenses criticism of an agent, not on the basis that they are mistaken about their desires but on the basis that they are mistaken about the best means to achieve the ends of their desires.

It is clear why a move like this strengthens the position of a defender of a Humean Theory. It avoids the unattractive implication that we can never be wrong about our reasons. It can allow for basic forms of errors that agents might make in their reasoning, like those that might arise from psychological factors like addictions, weakness of the will or unconscious compulsions.<sup>41</sup> Or

---

<sup>41</sup> Stocker 1979

perhaps individuals might have false beliefs about relevant causal facts, which prevent them from seeing that  $\phi$ -ing would be the best way to further the ends of their desires. By outlining some characteristics of what it requires to count as a rational deliberator, the Humean Theory can go some way towards accommodating the advisory constraint. We can be wrong about reasons we think we have, or ignorant of reasons we should have.<sup>42</sup> This occurs when we fail to meet some basic conditions that are stipulated as basic requirements of instrumental rationality. This first refinement can help explain part of the ordinary practice of normative discussion by providing a basis for legitimately criticising an agent's deliberations. As Lewis points out, the connection between our judgments and desires might be a “multifariously iffy connection...but nothing less iffy would be credible”.<sup>43</sup>

But a criterion of instrumental rationality doesn't get us as far as we might want. Normative disagreement isn't restricted to criticising lapses in reasoning, but sometimes seems to be about criticising the very foundation from which an agent is reasoning in the first place. We don't just want to take desires as given, but subject those desires to scrutiny when we engage in normative discussion. For this reason Humean theorists tend to make a second refinement to their basic theory. They broaden the psychological state which is picked out by the term 'desire', accepting that it doesn't have to be a simple brute impulse and can involve a complex mix of disposition and perception which isn't immune from reflective consideration. This opens the door to a much more thorough conception of the deliberative process where agents really can be mistaken about the desires which form the basis of their normative reasons.

How might this be the case? To begin with, expanding the scope of desires means that the elements of our psychology which motivate us can be correspondingly quite expansive. Bernard Williams

---

<sup>42</sup> Williams 1981

<sup>43</sup> Lewis 1989: 116

described the range of these desires as including “dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects”.<sup>44</sup> This outline suggests a much more intellectualised picture of desires where our conscious ideals can play a significant role in determining what we are motivated to pursue. As a consequence reflection can unearth new desires and extinguish existing ones, through a systematic exercise of comparison and imagination about the content which underpins them. Along similar lines, Mark Schroeder develops an account of desires which requires them to have a cognitive component in that they direct our attention to considerations in the world which are related to the objects of our desires.<sup>45</sup> Desires can't merely tug on our wills, but must also confer importance on various facts in the world since they are relevant to the satisfaction of our aims. If they were only functional states of disposition, then they wouldn't be able to rationalise the way that actions are connected with them. In other words, desires must be hybrid psychological states, which engage our motivational and perceptual capacities simultaneously.

These kinds of characterisations significantly broaden the notion of a desire for the Humean theorist. An implication of this is that there can be standards of rationality that can be applied to our desires themselves. This is because it would be inconsistent to allow that standards of instrumental rationality can be used to pass a verdict on our means-end deliberations while denying that we could judge the adequacy of our ends in the first place.<sup>46</sup> If our desires are really to be able to include our cognitive capacities in the way they are purported to, they can't be insulated from criticism in normative discussion. And of course, this is the result that a Humean should welcome in accommodating the advisory constraint. The challenge was that a theory of reasons based on desires wouldn't be able to explain our ordinary intuitions about the practice of normative disagreement.

---

<sup>44</sup> Williams 1981: 105

<sup>45</sup> Schroeder 2007: 156-157

<sup>46</sup> Scanlon 1998: 363-373

But the refinement of the theory in a way which allows that individuals can be wrong about their desires captures those intuitions by giving an outline method of rationally criticising their desires.

This section has outlined typical features of Humean theories of normative reasons. The essence of this theory is that when we make normative judgments, we are making judgements about how  $\phi$ -ing will further the object of one of our desires. These theories are relatively good at explaining how normative judgments motivate us but face a *prima facie* challenge in explaining the factualist platitude about normativity. I have shown that the typical response is to argue that we can make errors in reasoning and that desires are more complex than they might appear. This is the way that Humean theories tend to navigate a solution to the normative puzzle, while keeping intact our platitudes about how it works.

### **2.3 – Anti-Humean Theories of Normative Reasons**

The second solution to consider looks to explain normative reasons in terms of facts about normative propositions.<sup>47</sup> The basic intuition underlying this position is that reasons are ultimately determined by truths about what counts in favour of certain acts, thoughts or feelings. In order to figure out what reasons we have, we need to make judgments about what we should do. So Paul has a reason to believe that the Beatles represent the pinnacle of pop music if there is an aesthetic standard of value which supports that belief. Olivia should donate to charity if this action is recommended by the best moral theory. Nicola should feel happy when her sports team wins the championship because this is the reaction which makes the most sense for someone in her position. These examples are meant to get at the crucial notion, which is that reasons are ultimately justified by facts about whether they truly constitute reasons, rather than by psychological facts about particular agents. We might say:

There is a reason to  $\phi$ , in case it is true that an agent should  $\phi$  in those circumstances

This is a very broad analysis, and deliberately so. The open question is how we are to explain what it amounts to for it to 'be true that an agent should  $\phi$ '. To flesh this out requires the specification of some sort of more specific decision procedure which we can use to determine what agents should and shouldn't do. Different theorists tell very different stories about just what this decision procedure should look like. Some say that we should consider what an idealised counterpart of ourselves would think.<sup>48</sup> Or that we should use a standard which takes into account which sort of reasons are mutually justifiable.<sup>49</sup> Others still would want to say that truths about reasons are in

---

<sup>47</sup> Just as in the previous section, the position I describe here is an amalgam of widely held views. In addition to those cited below, it is also expressed in Korsgaard 1986; Darwall 1983 & Hampton 1998.

<sup>48</sup> Smith 1994

<sup>49</sup> Scanlon 1998

some way linked to facts about our rational agency.<sup>50</sup> All of these different sorts of views are detailed and complex in their own right. But my characterisation is designed to capture something I think is common to them all. Fundamentally, views like this deny that normative reasons are to be analysed directly in terms of what we want. They reject the central Humean tenet of the previous section, where subjective facts about agents played a pivotal role in analysing normative reasons. So it is fair to describe these theories as anti-Humean. They are all happy to say that an agent can have a normative reason to  $\phi$  even though  $\phi$ -ing might not serve any desire that the agent has.

These theories comfortably make sense of the advisory constraint. On an anti-Humean account, when an agent makes a normative judgment, they form a belief about whether a particular reason truly applies in the circumstance. So disagreement about normative reasons is easily explained in terms of disagreements about whether the agent is applying the right normative standard or whether they are applying that standard correctly. These notions mean that we are licensed to claim that individuals are mistaken about their normative judgments, and so engage them in discussion about how they could get it right. On the other hand, anti-Humean theories don't seem immediately well-placed to make sense of the explanatory constraint. The main objection is that they are vulnerable since an agent's beliefs about normative reasons can be fundamentally disconnected from their motivations about what to do. If normative judgments are linked to our beliefs about whether reasons are justified, then there is a risk that the conclusion of this normative reasoning might leave us cold and make no difference to the practical intentions we form. To respond to this challenge, anti-Humean theories typically make two sorts of refinements.

The first is to say that to the extent that we are rational, beliefs about normative reasons automatically bring about a corresponding motivation to act. If we really believe that there are considerations which count in favour of some intentional response but this makes no difference at

---

<sup>50</sup> Nagel 1970

all to what we want to do then we are evincing a form of irrationality. This position is supported by the intuition that when we engage in normative thought, we are engaging in a deliberative exercise which commits us to taking its output seriously.<sup>51</sup> To suppose that our rational decisions would really make no difference to what we actually do would be absurd in a way. The more natural thought is that someone whose rational normative judgments did nothing to 'get their limbs in motion' is not being rational in the first place. What is essentially happening here is that a criterion of deliberative rationality is being asserted in order to claim that if normative judgments come totally apart from an agent's psychological motivations this is due to errors in reasoning on the part of the agent.

Supporting this refinement is an analysis of the role that desires are traditionally thought to play in human intentional action. There is ostensibly a distinct motivational state which needs to be pressed into service before our decisions are actually translated into action. But anti-Humeans stress that there really isn't any such distinct state in the psychology of the rational agent.<sup>52</sup> Instead, they claim that desires aren't mere urges that tug on our wills in various ways, but need to be attitudes which are directed towards some outcome or end. That is, desires necessarily need to involve seeing things as good or bad reasons in order for them to make sense as an explanation of our actions. So if our judgments about normative reasons fail to motivate us at all, this is sufficient grounds to claim that we are acting irrationally. This gives the anti-Humean theorist the ability to start to explain why their theory of normative judgment can accommodate the explanatory constraint. They say that for agents who meet a basic standard of deliberative rationality, beliefs about what they should do will automatically lead to motivations to follow through. This dispenses with the need to invoke any direct analytic connection between normative reasons and the desires of agents.

---

<sup>51</sup> Smith 1994: 131-133

<sup>52</sup> Scanlon 1998: 37-41



In response to this, a question naturally suggests itself. What is the process by which we come to have the beliefs which are the basis of our normative judgments? By positing such a tight connection between these beliefs and motivations the anti-Humean may have gone some way towards capturing the spirit of the practicality platitude, but as a consequence a lump arises elsewhere in the theory. Specifically, the decision procedures we need to employ to form these beliefs need to be tied to the agents we expect to make these judgments. Otherwise the conclusions about what normative reasons we should act upon wouldn't be able to issue the kinds of practical motivations we need. The risk is that by asserting such a strong conception of what it means to be deliberatively rational, individuals might be alienated from the practice of normative judgment. In short, if the standard for forming conclusive beliefs is beyond the ken of ordinary agents, then we haven't got a successful theory how of ordinary human normative judgment operates.

This challenge shouldn't be overstated. It isn't incumbent on the anti-Humean that agents always flawlessly form beliefs in the way prescribed by the evaluative standards they recommend. What they do need to show is that the decision procedure by which agents ought to form beliefs about reasons is in some way tied to reflective capacities of actual human agents. And this is typically the way that anti-Humeans try to position their theories. They try to give an account of the appropriate standard of evaluation which isn't based on the actual desires of agents but still counts as an ideal of *human* thought. In a sense, we might describe the task as that of finding a standard of evaluation which is just ideal enough to play the task it is assigned. Whether individual instances of this strategy succeed is not directly the issue here; different proposals will have their own strengths and weaknesses. But what any instance of this strategy must avoid is to argue that the correct beliefs about normative reasons are determined by principles which are wholly abstracted from the identity of ordinary agents. This is the second refinement that an anti-Humean theory needs to make in order to explain how normative judgments will motivate those who make them.

This section has outlined typical features of anti-Humean theories of normative judgment. They claim that beliefs about the truth of reasons are the basis of normative thought and talk and this means they comfortably make sense of platitudes about the way we engage in normative discussion and debate. However they face a challenge in explaining how our normative judgments practically motivate us. In response to this challenge, these theories claim that rational agents respond to their conclusive beliefs about reasons as long as these beliefs are evaluated from a standard which is connected to the reflective capacities of typical humans. This is how they explain normative judgment while vindicating our platitudes about it.

## **2.4 – Evaluating Humean and anti-Humean Theories**

Having outlined both Humean and anti-Humean theories, we are now in a position to evaluate these approaches. Recall that I positioned both of these theories as non-revisionary. This means that they are answerable to our intuitions about normative discourse and so must be able to explain how the tension between the factualist and practicality platitudes is resolved. As I mentioned earlier, this means there is a descriptive burden that both theories take on. I want to argue in this section that neither the Humean nor the anti-Humean theory discharge this burden, primarily because both accounts depend on a substantive assertion about the nature of human rationality which is itself unexplained.

The first thing to note is that despite the ostensibly fundamental differences between the two approaches, they end up converging substantially. Desires and beliefs are the central psychological states for the Humean and anti-Humean respectively. But these two states are increasingly massaged into a shape where it becomes difficult to pick out any real difference between the two. Start with the Humean theory. The initial understanding of desires was that they were dispositional states which motivated particular responses when they served to satisfy those dispositions. But it quickly became apparent that a view like this was too crude. Desires couldn't just be the producers of motivational impulses but had to be understood as broader psychological states which might include reflective and perceptual components. This broadening of the notion of desire was necessary in making sense of agents being wrong about the normative reasons they thought they had. Now consider the anti-Humean theory. Beliefs are initially understood as attitudes about the truth or falsity of propositions about normative reasons. But to make sense of the motivational aspects of normative judgments, the process by which these beliefs are formed needs to be tied to the actual reflective identities of agents. Beliefs about normative reasons are perceptual, but these

perceptions can't be entirely divorced from the things we are disposed to want. On one account we end up in a place where our desires can be wrong; on the other we end up saying our beliefs can motivate.

Though terminological differences remain, it is unclear whether there is any residual distance between these two approaches. It seems that either theory is attempting an explanation from opposite starting points but with a similar result. This observation implies that both theories offer an account of normative judgment which invokes both beliefs and desires. So despite the purported difference between the two approaches we end up with an account of normative psychology which is more ecumenical than it might have initially appeared. The convergence isn't a conceptual one. Rather, it is an indication of an overlap in both approaches prompted by a necessity to make sense of the platitudes about normative discourse that they set out to explain.

A natural suggestion at this point is that this impasse is a consequence of the naïve dichotomy between beliefs and desires that was the starting point of this investigation. The result might tempt us to look for a third way. Along these lines some explore the possibility of a psychological state that is neither a belief nor a desire, but has belief-like and desire-like characteristics – what has been referred to as a “besire”.<sup>53</sup> This kind of state would promise to explain the nature of our normative judgment much more parsimoniously by tying it to a single psychological state that has the capacity to explain both the practicality and advisory platitudes. John McDowell provides a description of such a state as a mode of conceiving of the world that casts certain acts in a favourable light.<sup>54</sup> The idea is that there is a distinctive way of understanding the facts of a situation that disposes us to act, think and feel certain things. The example he uses is that if we come to view someone as shy and sensitive, that particular perceptual understanding will itself bring about a tendency to treat them

---

<sup>53</sup> The term was coined by Altham 1986

<sup>54</sup> McDowell 1978

with care, feel empathy for them in social settings, exhort others to feel the same way and so on.

Importantly the point is that none of these practical effects require the presence of an independent desire – they are the mere consequence of a special way of seeing the world.

This type of besire account, though tempting, should be resisted. Though it might seem like it allows us to evade the problem by rejecting the simple belief/desire psychology at its source, it doesn't fare much better and brings about problems of its own. For starters, a theory that views beliefs and desires as distinct psychological states needn't deny that they often contingently coexist, so we can *both* believe that someone is shy *and* desire to treat in a certain way.<sup>55</sup> This reliable connection is in fact is the target of explanation. On this count, besires don't offer any different conclusions about the tricky cases than the Humean or anti-Humean accounts do. Instead they offer a unitary state to explain the mechanism by which this occurs. The problem with this is that it implies a necessary entanglement of aspects of our normative discourse that we can, at least theoretically, keep separate.<sup>56</sup> Imagine we are criticising someone for 'failing to see the world right' in their perception of their shy friend. This might mean that we are saying that they have been inattentive to facts about their friend's behaviour which suffice to establish that he is shy. Or it might mean that we are saying that they should be more disposed to help their friend deal with this personality trait. These two dimensions of normative discourse are separable in principle, but a besire account would suggest that they aren't, since both are part of the unitary way that we conceive of the world. This is at least *prima facie* implausible. It shows us that we don't have an easy way out of the impasse, since a besire theory would deny that there is a distinction between the factualist and practicality platitudes in the first place.

So where does that leave us? We have seen that the Humean and anti-Humean approaches converge

---

<sup>55</sup> Smith 1994: 118-122

<sup>56</sup> Blackburn 1998: 97-100

in their explanation of normative judgment, a surprising result given their purportedly distant starting points. Obviously, this itself isn't a criticism of either approach. Perhaps we'd even expect some convergence between these two accounts, since they are both trying to explain the same set of intuitive data. But this convergence allows us to see the common flaw that undermines both the Humean and anti-Humean approaches. Though each tries to offer a robust descriptive account of normative psychology, they are vulnerable in this endeavour. In both cases the story about the operation of belief and desires is supplemented by an assertion about the operation of human rationality. Humeans emphasise instrumental rationality, the idea that agents could be held to account for failures to reason from their means to their ends. Anti-Humeans emphasise deliberative rationality, the idea that an agent who believes they have a conclusive reason can't consistently reject that they have a motivation. These assertions were important in allowing either theory to make sense of ordinary platitudes about normative discourse. However the objection that can be levelled against them is that instead of offering a positive explanation of normative psychology, they instead assert a normative claim and try and use this as a foundation of a descriptive account. Essentially, they are trying to derive an 'is from an ought'.

The problem is that in offering a solution to the normative puzzle, we don't want to invoke the platitudes we are attempting to explain as part of our explanation. But this seems to be the effect of the strategy adopted by both the Humean and anti-Humean accounts. The assertions about the operation of human rationality aren't descriptive but are themselves substantively normative. They take a thesis about how human normative thought *should* operate and use this as the basis of a description of how it does operate. Worse though, these assertions don't seem to be very far removed from the platitudes they are invoked to explain. Insisting on a criterion of instrumental rationality draws on the intuition that a proper basis for criticising an agent's deliberations is that they might be wrong about their reasons. This is very close to the same intuition underpinning the

factualist platitude. Similarly, applying a standard of deliberative rationality gains plausibility from the intuition that an agent with conclusive reasons who isn't motivated is evincing a cognitive irregularity. This intuition is close to the one underpinning the practicality platitude. So the nub of this objection is that both the Humean and anti-Humean accounts are taking platitudes about normative discourse as descriptively significant and then are using this as a part of their account of normative psychology. But this doesn't discharge them of the burden they need to meet, which is to give an account of normative psychology which explains but does not assume these platitudes.

What does this mean? Despite a great deal of sophisticated conceptual work, Humean and anti-Humean approaches are deficient in an important respect. They don't provide an explanation of how normative judgment works. We haven't yet seen how our normative discussions actually engage us, the mechanisms by which this discourse is possible, the nuts and bolts of the operation of our ordinary talk and thought. This is a very important defect in this orthodoxy. We are owed an explanation of this sort, and my argument in this section has been that neither of the standard non-revisionary theories provide one. This means the normative puzzle that I began this chapter with remains unsolved in an important respect. The two potential solutions we have examined have brought us closer to the answer, but there is further still to go.

For the rest of this chapter I offer my own ideas about how philosophical theories should proceed in light of this critique. As I see it, there are two avenues that need to be explored in order to provide a more satisfactory explanation of the nature of normative judgment. Both are motivated by the thought that the best approach from here is to pay more direct attention to the factualist and explanatory platitudes that are the real source of the problem. So far they have been taken as platitudinal data, but I will suggest that we can investigate them directly and show how that might help us progress towards a more satisfactory solution. The first suggestion is conceptual and looks to diagnose what philosophical accounts may have under-emphasised in their approach. I argue that

attending to the shared elements of normativity can provide a way to improve the analysis of normative judgment. The second suggestion is that philosophical accounts enrich themselves by looking outwards towards empirical research that can provide insight into the operation of normative judgment. We can think of this as the project of naturalising normativity and I justify this approach in the next sections.



## **2.5 – Shared Norms and Communal Judgment**

This section provides a diagnosis of why non-revisionary theories are vulnerable, and a suggestion about how they can orient themselves better in trying to explain the normative puzzle. The animating idea is the same as the title of the thesis: that normative judgment has social foundations. I will briefly consider how the Humean and anti-Humean theories have overlooked this fact. Then I will examine an argument between Bernard Williams and Christine Korsgaard that illustrates how these traditions have the resources to be enriched from within by considering social conditions of normativity. This will lead me to demonstrate how these lessons can be incorporated into the way that I have framed the normative puzzle throughout this chapter, by expanding our understanding of the practicality and factualist platitudes themselves.

If sociality is so important, then how have standard theories of normative judgment omitted it from their analyses? I suggest that their emphasis has been slightly off target. Both the Humean and anti-Humean approaches seem to be explaining normative judgment from an individualistic standpoint. They each pick a mental state, belief or desire, and then build an account which treats that state as central. Their attempt to vindicate observations about normative discourse is built from within. We are invited to use essentially introspective methods to think about the operation of our beliefs or desires; to reflect upon our own capacities and use these as a secure basis for our account of normative thought. It is clear why this is a tempting and fruitful path. Individuals all have some access to the experiences they have when thinking about how norms play a part in their thoughts and there is a certain authority that these first-person observations intuitively possess. Trying to ground an account of normative discourse in shared individual mental states thus seems like a good way to try and secure footing for a theory of normative psychology which is by its very nature difficult to pin down.

The problem is that the target phenomenon isn't something which is individualistic. Normative discourse necessarily involves a distinctively social dimension. The factualist platitude is an observation about how we come together to discuss norms, and the practicality platitude is an observation about how odd it is if others don't react appropriately to these discussions. There appears to be a sense in which norms importantly mediate our interactions with each other that is elided by theories which focus on the beliefs and desires of individuals without much attention on the social aspects of our normative psychology. My proposal then is that the natural response to my critique in section 2.4 would be for both Humean and anti-Humean theories to acknowledge that normative psychology is a shared enterprise. Even putting aside any turn to empirical investigation, conceptual theories of normative judgment could benefit from more fully embracing this.

This is not an entirely idiosyncratic approach to take. In fact, several authors have alluded to these types of ideas, almost parenthetically, in the midst of defending theories which don't emphasise them. For example, at the end of a long and detailed exposition of an avowedly anti-Humean theory, Michael Smith remarks that the social dimension of normative must have some epistemological consequences. He writes:

“Other things being equal, the individual must therefore have a proper sense of humility when she finds herself in conflict with the group....None of us has any special epistemic gifts that would justify us in privileging our desires and judgments over the desires and judgments of others.”<sup>57</sup>

Similarly, after offering an extended defence of the authority of individual reflection, Christine Korsgaard rejects the idea that reasons could ever be private mental entities in trying to explain the normative origins of obligation. She writes:

---

<sup>57</sup> Smith 1994: 176-177

“If these reasons were essentially private, it would be impossible to exchange them or share them. So their privacy must be incidental or ephemeral; they must be inherently shareable...I take this to be equivalent to another thesis, namely, that what both enables us and forces us to share our reasons is, in a deep sense, our social nature.”<sup>58</sup>

So I don't make any particular claim to novelty. Instead what I will do is pick up the suggestions offered by the quotations above and develop what they imply for philosophical theories of normative judgment. A good way to see how both Humean and anti-Humean theories can adapt themselves is to briefly look at an argument between Korsgaard herself and Bernard Williams. Their dispute is specifically about the role of reflective endorsement as a foundation for ethical knowledge, but it is illustrative of more general concerns relevant to normative judgment.

Though his journey is different, in his book *Ethics and the Limits of Philosophy* Williams reaches a point where he frames the problem facing morality very similarly to the puzzle that I have posed about normativity.<sup>59</sup> His concern is how to explain the use of what he calls thick ethical concepts – those like cowardice, brutality, gratitude (his examples) or promiscuity, gregariousness or being a Yank (we might add) that have both descriptive and connotative elements to them. In Williams' words, the application of these concepts is both “world-guided and action-guiding”<sup>60</sup> and that for an individual to properly use them requires an understanding cultivated from within a particular social world since they are a “cultural artifact they have come to inhabit”.<sup>61</sup> This can be read as suggesting that understanding the functions of normative discourse requires seeing them as growing out of the conditions of human social life. This is only the barest of sketches of Williams' argument. But his role as a prototypical Humean when it comes to normative reasons, means that his position on the social foundations of ethical knowledge is a useful demonstration of the way that theories of

---

<sup>58</sup> Korsgaard 1996: 135

<sup>59</sup> These points are mainly drawn from Chapter 8 of Williams 1985

<sup>60</sup> Ibid: 156

<sup>61</sup> Ibid: 163

normative judgment can be enriched by this perspective.

Though she rejects Williams' conclusions, Korsgaard also adopts a socially oriented theory of normative judgment. Her disagreement with Williams is about the source of ethical knowledge, as she argues that it comes not from our collective sentiments and dispositions but from the obligations that can be discovered as a consequence of our humanity as such.<sup>62</sup> Later she goes on to develop the idea that this process doesn't deliver us private reasons, but public ones. She wants to resist the notion that our social nature is too biological or contingent to play a role in arguments about the foundations of morality. Her suggestion, drawing on Wittgensteinian themes, is that the illusion of the privacy of our consciousness is what precludes normative reasons from being properly understood as communal.<sup>63</sup> Korsgaard's argument attaches a prominent role to sociality even as it retains the paradigmatically anti-Humean flavour of searching for standards that justify the truth of normative reason-claims.

What relevance does this exchange have for the account of the normative puzzle that I have provided in this chapter? The answer is that it shows how the social foundations of normative judgments can be incorporated into either of the Humean or anti-Humean theories that seek to solve the puzzle. And so though I haven't developed either Williams or Korsgaard's positions in much detail, they are instructive as they provide explanations which can change the emphasis of the solutions away from an individualistic basis. To incorporate this thought more generally, we can apply this lesson to the way that we frame the puzzle in the first place. We can do this by asking how the factualist and practicality platitudes themselves might be more thoroughly explained in social terms.

Start with the factualist platitude. The key question here is how shared frameworks of evaluation are

---

<sup>62</sup> Korsgaard 1996: 90-130

<sup>63</sup> Ibid: 135-142

constructed and perceived by individuals in a community. Importantly, normative discourse involves agents actively agreeing and disagreeing about normative reasons. But this practice implies that norms are themselves represented in a way that allows individuals to be agreeing or disagreeing about something. These representations need to be shared, as they would pick out the distinctive *communal content* which forms a part of human normative discourse. Discussing normative reasons involves individuals coming together to find answers to a set of questions, but the formulation of those questions needs to cohere at some basic level otherwise the enterprise would be fraught from the start. So what needs elucidation is just how individuals are able to come together to participate in the distinctive practice of normative discussion. The explanation here will need to account for the fact that these shared norms aren't so rigid that they never change, but neither are they so malleable that they are entirely determined by individual agents.

Now consider the practicality platitude. The key question here is why we are disposed to take norms seriously. Normative reasons seem to be able to intrude on our consciousness and engage the motivational apparatus of our minds. One of the reasons this seems to be the case is that normative content is imbued with a significance which enables it to be used as the basis for imperatives, either towards ourselves or others. We appear to be able to use norms to direct practical decisions. But this implies some degree of consistency in the way norms impact on human decisions, otherwise they wouldn't be useful in regulating acts, thoughts and feelings. The idea then is that there must be some *behavioural coordination* in the way that normative discourse typically functions. What requires further attention here is just how some sorts of human interaction could come to fulfill this function. We also need to be able to account for the fact that we normally can't help but react to normative imperatives, but we aren't normally overwhelmed by them in our thoughts.

This section demonstrates what can be gained from an appreciation of the social nature of normative discourse. The preceding two paragraphs only raise preliminary thoughts about the social

nature of normative discourse. But in relating them to the platitudes which I have used to frame the normative puzzle, the intention is that conceptual theories of normative judgment could take these modified platitudes to help solve the puzzle in a more satisfactory way. I myself won't explore *how* this would occur, but the point of this section is to support the contention that this *should* and *could* occur within both the Humean and anti-Humean traditions. In the next chapter, as part of a more empirical approach, I will explore other ways in which our social nature is important in understanding normative judgment. Before that though, I need to justify how this empirical approach is related to my arguments so far in this chapter. I do this in the next section.

## **2.6 – Naturalising Normativity**

There is a general objection to the line of reasoning I advanced in section 2.4. Perhaps proponents of Humean and anti-Humean theories only ever intended them to be theories that give us conceptual accounts of normative judgment. They might object that they never had any intention of trying to offer a descriptive theory of the kind I have suggested. If this is the case, then perhaps it is uncharitable of me to press them into service for a different task, and then criticise them for failing to execute this task. This is a fair response. To some extent I am shifting the dialectical goalposts. But this is deliberate. Though it may only be a matter of emphasis, it seems to me that the philosophical task we are engaged in would be well served by treating it as a natural phenomenon and using the methods and findings of empirical science to inform our understanding. This is what I mean by naturalising normativity. In the next chapter I will provide an example of this approach. I want to end this chapter by explaining and justifying why this is a relevant approach to take up.

This chapter has been primarily concerned with answering a key question: how does normative judgment work? I laid out a puzzle, and examined two prominent theories which aim to solve this puzzle. Both of these theories were primarily conceptual, in that they examine our thoughts about the phenomenon of normative judgment and tried to provide a coherent understanding of them. There is certainly nothing wrong with this generally; after all I did much the same thing in Chapter 1 in my analysis of the concept of norms themselves. But what I want to say here is that the philosophical task is itself a scientific one, and we can profit from decomposing our exploration of it into conceptual and empirical components. The reflexive nature of normative judgment means that thoughts about our thoughts will always need attention. But this introspection can be complemented by an empirical investigation which seeks to place some distance between the object and subject of our study, such as is possible.

To delve into this kind of investigation is to admit the relevance of a large body of scientific research to the understanding of normative judgment. But this empirical task needn't start from a blank slate. The various findings of the Humean and anti-Humean theories can guide the way we frame our understanding of the research as well as providing more specific hypotheses about the capacities we'd expect to find. So the traditional elements of philosophical psychology can almost be treated as a working scientific model. If it turns out to be entirely unsupported by the facts, then this would be a informative result. And just as concepts can guide empirics, the reverse is equally possible. So a thorough investigation of the psychology of normative judgment might excavate findings that shed some light on how rationality works, which might help us assess the plausibility of the deliberative or instrumental concepts discussed earlier. This probably wouldn't decide the debate between the Humean and anti-Humean, but could certainly be relevant to theory choice.

This is why taking up the empirical side of this connection can provide a contribution to our understanding of normative judgment. We don't know in advance what consequence this will have for our concepts, so we should just get on with it and see. So in the next chapter I will put this approach into action. As I proceed I will occasionally draw attention to the ways in which the empirical research resembles the various psychological theses that have been discussed throughout this chapter. The model of normative cognition that I will end up supporting has elements that cohere with some of the observations above. This isn't surprising, and is good indication that our conceptual platitudes aren't entirely off target. But it also provides a model that could be used to support the further development of the conceptual theories and I comment briefly on how this might occur in the conclusion of the thesis.

A potential ambiguity lurks here, and some clarification is in order. A 'naturalising' project in philosophy admits of multiple interpretations.<sup>64</sup> In one sense, it might mean giving an account of a

---

<sup>64</sup> Kitcher 1994 distinguishes four ways of "biologizing ethics" in his critique of early sociobiology, which I draw on



phenomenon that is scientifically minded. In our case, this would mean explaining the nature of normative judgment in a way that is consistent with a naturalistic world view. This descriptive task is certainly something that I will be doing in the next chapter, but it is fairly uncontroversial. The more contentious part of my aims is suggested by my comments above about the relationship between the empirical and conceptual aspects of a theory of normative judgment. The worry might be that I am claiming outsize ambitions for scientific research, that it is the key to meta-normative theory, and can settle traditional questions about the nature of normative judgment. This is related to a historically important argument known as the naturalistic fallacy attributed to G.E Moore.<sup>65</sup> Very broadly, the objection is that just as I criticised conceptual theories for trying to infer an 'is' from an 'ought' in the previous section, so too we should be wary of trying to infer 'ought' from 'is' when naturalising normativity. In a similar vein, Derek Parfit has recently attempted to bolster a non-naturalistic approach to normativity by painstakingly arguing that all available versions of naturalism are faulty.<sup>66</sup>

Without attempting to engage in a lengthy discussion of these types of arguments, I want to end this section by commenting on how they might apply to my naturalising project. I think that understanding normative judgment involves two parallel tasks, a conceptual one and an empirical one. This chapter has primarily been about the conceptual task and in this section I have argued that the empirical task is also worth pursuing, which is why I will pursue it in the next chapter. This much I think isn't vulnerable to any sort of naturalistic fallacy. I also suspect that this empirical task can inform the conceptual task, though it is hard to predict just how in advance. This may be vulnerable to some sort of objection, though the details would very much depend on the connection that one tried to draw. However, in this thesis I will not directly address the way that my empirical

---

here.

<sup>65</sup> Moore 1903. The historical use, and abuse, of this argument in relation to moral naturalism is discussed in chapter five of Joyce 2006.

<sup>66</sup> Parfit 2011

model might (or might not) fold back onto the conceptual task. So unless some sort of argument can be found that resists any form of naturalism *tout court*, I can proceed in the next chapter immune from this sort of criticism as long as it is remembered that what I am doing there is providing an alternative, but purely *descriptive* account of normative judgment.

## **2.7 – Conclusion**

In this chapter, and particularly the last two sections, I have made arguments that are at the heart of my thesis. I began by outlining a puzzle that arises from the clash between our intuitions about the the practical and factualist aspects of normative discourse. I used this to set-up the key problem that we would want any account of normative judgment to answer. I then presented and critiqued both Humean and anti-Humean theories of normative reasons, which I called non-revisionary since they attempt to squarely solve the normative puzzle. I argued that both of these types of theories were deficient in the same way due to their tendency to attempt to explain a descriptive phenomenon by positing an unexplained conceptual ability. Then, in the last two sections I suggested two ways in which non-revisionary theories might improve. In the next chapter I will attempt to take up both of these suggestions. I will survey empirical research into normative judgment that is an example of what I called naturalising normativity. In addition, this investigation will be shaped by my contention that the social foundations of norms are an important part of what needs to be explained.

---

## 3

# The Roots of Normative Cognition

---

This chapter advances the agenda introduced at the end of Chapter 2 by turning to my attempt to develop an explanatory account of normative judgment. Specifically, I will survey some evidence that supports a model of human normative cognition that emphasises its social elements. The particular empirical hypothesis I will develop in this chapter isn't contingent on the truth of my earlier conclusions and neither is it a necessary consequence of those conclusions. Understood one way, this chapter can be read simply as an attempt to understand how our capacity for normative cognition works. But more broadly, my project here is a consequence of the argument of my thesis. Up till now, I have looked to develop a closer connection between a conceptual theory of normativity and the actual psychological operation of our normative thought. So this chapter can be understood as an attempt to vindicate the merits of that applied approach.

My plan of attack is as follows. In the first section, I will start by canvassing some of the diverse empirical work on norms and drawing out a general working definition of the concept, and the psychological capacity it implies. I will argue that despite some research which attempts to show that norms in different domains are processed differently, these distinctions don't undermine the prospects of such a general definition. In the next two sections I will present and critique two types of theories which try to explain our normative cognition. The first argues that important elements of normative cognition are innate. The second argues that normative cognition evolved out of an ability to follow commands backed by the threat of social punishment. My position is that neither of

these theories is in a position to successfully explain the whole of the operation of normative cognition, primarily because they are inattentive to the shared social foundations of normative judgment. This will lead me to my last two sections, where I will present and defend my own model that provides a better framework for normative cognition. In section five I present the model and show how it can focus attention on issues that extant theories under-emphasise. Then lastly I will sketch some evidence which supports this framework, both in terms of how it operates and how it might have evolved.

### **3.1 – Norms as a Social-Scientific Concept**

There has been a burgeoning social-scientific interest in investigating the psychology of norms in recent years. By necessity, this investigation has been piecemeal and tentative, as we'd expect when the target phenomenon is of such complexity. Researchers have been interested in asking two separate but related sets of questions. The first are questions about how normative psychology operates and develops in typical humans today – issues of ontogeny. The second are questions about how we evolved to have these capacities – issues of phylogeny. These two issues are distinct, but often difficult to prise apart at such a nascent stage in the explanation. Understanding how a mental capacity evolved often involves speculating about its function and form, which isn't entirely independent from the question of how it might have evolved in the first place. So sensibly, researchers haven't tried to impose an artificial firewall between these two parallel lines of inquiry. I shall follow this lead throughout the chapter, though I shall try and be clear about the distinction when it is possible.

Furthermore, many of those who investigate normative psychology do so through the lens of moral psychology. Again this is sensible, since morality is probably the normative domain which is the most prominent in our everyday lives and poses interesting questions of its own. But as I draw on much of this literature I shall liberally, though charitably, interpret it as applicable to broader issues about normative cognition. I have given conceptual reasons why I favour this approach earlier in my thesis, and below I shall go on to argue why such an approach makes sense even on an internal reading of the evidence as well. I mention these as brief methodological preliminaries which inform the approach I am taking.

So what are 'norms'? In a typical social-scientific spirit, no conclusive definition exists, but nor

should it be demanded since the nature of norms themselves is the target of empirical scrutiny.<sup>67</sup>

However, various informal accounts have been offered which tend to emphasise the same features.

At the broadest level, there seems to be a tentative acceptance of the claim that whatever norms are, in some rough sense they link attitudes and behaviours with judgments of permissibility. Here is a sprinkling of quotations which illustrates this:

“As we use the term, a norm is a rule that specifies actions that are required, permissible, or forbidden independently of any legal or social institution.”<sup>68</sup>

“Roughly speaking, by norms this literature refers to informal social regularities that individuals feel obligated to follow”<sup>69</sup>

“As we shall understand them, norms are attitudes toward types of actions, emotions, thoughts, or other traits....Such norms can prescribe or forbid a thought, behaviour, or any other characteristic”<sup>70</sup>

“Norms are an important species of social institution...[they] reinforce certain patterns of behaviour, but they do so in their own way, by representing those patterns as peculiarly desirable or obligatory.”<sup>71</sup>

A persistent observation is that norms operate by way of rules which prescribe the conditions or circumstances when the norm has been violated. The suggestion seems to be that norms centrally play a deontic function, which sees them as associated with a cluster of notions about what we must, should, ought, or are permitted to do. This doesn't seem to elicit much disagreement at the descriptive level. Moreover, there is a general consensus that these rules possess “independent

---

<sup>67</sup> Though in Chapter 1 I gave my own conceptual definition of norms, here I am addressing the different issue of definitions that emerge from within the social-scientific literature on norms.

<sup>68</sup> Sripada & Stich 2007: 281

<sup>69</sup> McAdams 1997: 340

<sup>70</sup> Machery & Mallon 2010: 12

<sup>71</sup> Pettit 1990: 725

normativity” which means that they can regulate social behaviour on their own and don't need any the backing of other sorts of institutions like a judicial process or formal social sanctions. Norms get their authority on their own and don't require any formal social sanctions to support their operation, even though they might be reinforced in this way. This empirical observation has been inferred from the evidence gathered in a variety of different disciplines from anthropology and sociology through economics and legal studies. We might sum it up by saying that *norms are rule-bound*. This isn't a conceptual analysis of norms. But the fact that we have the ability to understand norms does imply a capacity to be able to reason about a certain suite of concepts associated with normativity.

Though we might identify other generalities about norms, the fact that they are rule-bound is an important feature. It resonates with my with my conceptual analysis of norms, which analysed them as involving a relation between a correctness value, a psychological response and a social framework. It picks out a foundational element which we could treat as a part of all sorts of normative discourse from the moral and prudential to the conventional or epistemic domains. It would also seem to vindicate the idea that we can think of a capacity for normative cognition which encompasses these domains. However, the idea that we process norms as rule-bound in such a general way encounters some resistance. The worry is put clearly and forcefully by Leda Cosmides & John Tooby:

“Social interaction across many domains involves moves & countermoves, expectations, obligations, prohibitions, and entitlements. But the nature of these can be very different... [and] may employ some *version* of a particular deontic concept. But each version may differ from the other by virtue of the unique inferential role it plays...If this picture is even remotely correct, then the project of creating a deontic logic that is both general yet empirically descriptive may be doomed.”<sup>72</sup>

---

<sup>72</sup> Cosmides & Tooby 2008: 58



This is a legitimate concern. If it could be demonstrated that humans have separate inferential systems for distinctive domains of deontic reasoning then this could undermine the claim that there is a domain-general capacity for normative cognition. This would mean that it would be better to adopt a more splintered research programme that looked to describe more specialised psychological capacities. And it is certainly true that there are some domain-specific characteristics which accompany different sorts of normative rules, not least of the sort explored by Cosmides & Tooby themselves in their studies of the way that humans understand if-then conditionals in different contexts. However, to demonstrate such a fissure it isn't sufficient to just show some contextual differences between the way deontic concepts are understood. There would need to be a demonstration that the uses of terms like *should* and *ought* in different normative domains involve fundamentally incommensurable patterns of thought. In the absence of such a decisive empirical result, there is utility in continuing to assume that there is some mileage to be made out of a domain-general understanding of norms as rule-bound.<sup>73</sup>

But this is a little too hasty. The critique in the previous paragraph can be put in a less ambitious but still dangerous way. It isn't necessary to show that there are as many different deontic inferential systems as there are normative domains. It is sufficient to argue that there is simply more than one in order to pose trouble for a domain-general understanding of normative rules. A prominent research tradition hones in on just such an attempt, and centers on what is known as the Moral/Conventional (M/C) distinction. It argues that there is a fundamental psychological difference between the nature of rules in the moral domain and those in other domains. Drawing on a wide array of developmental research<sup>74</sup>, the M/C model presents a host of evidence that from an early age, humans treat transgressions of moral and conventional rules very differently. Moral rules are general in scope, authority-independent, appeal to notions of harm, justice and rights, and

---

<sup>73</sup> Mallon 2008

<sup>74</sup> Turiel 1983, Smetana 1993, Nucci 2001 all provide good summaries of the research.

violations of these rules are judged to be particularly serious. Conventional rules are essentially defined as rules which lack these distinctive features. The ability to understand this distinction emerges early and it is cross-culturally robust. These facts are taken to be suggestive of a deep psychological difference in the way that the human mind processes norms. Though both moral and conventional rules pervade the social order, they belong to separate conceptual domains and this explains why we have different patterns of response to them. This in turn suggests that functionally distinct components of the mind deal with each domain. If this hypothesis is borne out, then it would be a reason to suppose that there isn't a single reasoning ability that underpins our understanding of rules across different normative domains.

Though this research programme is wide-ranging and has generated robust and replicable observations, there is some reason to doubt its conclusions. Daniel Kelly and Stephen Stich note that the signature moral and conventional response patterns that the M/C model reports aren't always as tightly bound to violations within their respective domains.<sup>75</sup> One example of contradictory evidence they offer is based on studies that show that violations of etiquette rules – like spitting in one's water before drinking it – tend to be judged as authority-independent even if they have nothing to do with the distinctive moral notions of harm or justice.<sup>76</sup> Since the M/C model claims that there is a sharp qualitative difference between moral and conventional domains these sorts of results are anomalous and count against it. If the psychological wires can be crossed so easily then it suggests that perhaps we don't have dedicated mechanisms for moral rules. We can concede that the M/C model has probably identified an interesting natural sub-category of normative rules while resisting the claim that this sub-category carves out its own mental fiefdom.

Time will tell whether similar disconfirming empirical results accumulate, but for present purposes

---

<sup>75</sup> Kelly & Stich 2008: 360-366

<sup>76</sup> Nichols 2002

we will proceed on the assumption that the evidence for a sharp distinction between different kinds of normative rules is unconvincing. So I will assume from here that norms are rule-bound in a domain-general sense. This understanding of norms implicates a general psychological capacity to reason about them. This is the capacity for normative cognition which will be the focus of the remainder of the chapter. Just what sort of cognitive architecture would subserve the capacity to reason about norms? The question that I will address is one of design rather than implementation. We aren't looking for an elaboration of the neurological mechanisms which are at play when we reason about norms. Instead we are looking for an account of the psychological structures which would need to be in place to explain normative cognition. In exploring this issue, I will place some emphasis on the public functions of normative judgment that I discussed in the previous chapter. To recap, these were the communal content and the behavioural coordination that normative discourse seems to require. I will treat these features as desiderata of theories of normative cognition.

### **3.2 – Normative Nativism**

The first account of normative cognition to consider is one that we can call 'normative nativism'. Essentially, it advances that our competence in normative thinking is attributable to an evolved psychological capacity that is part of our common biological heritage, much like our capacity for vision or language. There should be little resistance to the general idea that our evolutionary history has something to do with the way we are built. We got here and we used to be there so presumably there is some historical story about how we made the journey. Nativism also make the further claim that our capacity for normative cognition was adaptive in that it conferred a fitness advantage on our ancestors. This is a more substantive assertion about the way this capacity may have evolved but also one that I will not dispute. The issue which will be discussed at greater length in this section are the claims made by nativism about the *innateness* of normative cognition.<sup>77</sup> Roughly speaking, this is the idea that our normative capacities are 'instinctual' which is a hypothesis about the structure of our minds. The claim that I will focus on here is the hypothesis that our normative thinking involves dedicated psychological systems that help us conceptually navigate normative discourse. So this is a ontogenic thesis, about the operation and development of our normative psychology. In this section my main contention will be that nativism isn't a successful theory of normative cognition because its claims about innate understanding don't mesh well with the reflective social elements of normative judgment. In the parlance of the previous chapter, it doesn't account for the communal content that is at the heart of the factualist platitude.

There has been a burgeoning interest in defending normative nativist views in recent years. Much of this interest can be traced, directly or indirectly, to the Chomsky-inspired advances in linguistics which rigourously argued that much of our ability with language is based on innate cognitive

---

<sup>77</sup> Joyce 2013 points out that the ways that the literature about moral nativism is plagued by multiple ambiguities.

structures.<sup>78</sup> This inspiration has been transported over to the normative realm in a number of ways. Predictably, moral psychology has provided extremely fertile ground for these ideas to take root. A number of thinkers such as John Mikhail<sup>79</sup>, Susan Dwyer<sup>80</sup>, Marc Hauser<sup>81</sup> and Gilbert Harman<sup>82</sup> have elaborated various views about the degree to which we have innate moral knowledge, some of which we will look at in more detail shortly. But morality isn't the only normative domain where innateness has found some favour. For example, a nascent but growing research programme known as Literary Darwinism argues that our understanding of the narrative form is a result of psychological capacities that enable us to understand fiction.<sup>83</sup> Though only in its early stages, this is often interpreted as the claim that features of our aesthetic psychology are innate.

The most thoroughly worked out versions of normative nativism brandish their Chomskian credentials proudly. To make the case that normative cognition is innate they typically rely on a version of a 'Poverty of the Stimulus' argument which at its core is quite simple. The idea is that our competence in understanding normative rules is much too high to be the result of general learning given the paucity of the data we are exposed to in our environment. As Dwyer notes, there are two challenges that we face in learning to reason about norms.<sup>84</sup> The first is that there are untold regularities that we are exposed to from an early age, some of which are as mundane as learning which side of the pantry the cereal is stored. How do we learn to distinguish these accidental regularities from those which are more avowedly normative, like learning that about the rules which govern knife and fork use? Secondly, even if we could make this distinction, how do we learn to distinguish between different sorts of normative rules, like making the crucial distinction between

---

<sup>78</sup> Chomsky 1980

<sup>79</sup> Mikhail 2007

<sup>80</sup> Dwyer 1999; Dwyer 2007

<sup>81</sup> Hauser 2006

<sup>82</sup> Harman 2000

<sup>83</sup> For instance, Boyd 2009 argues the case for fictional narrative and Dutton 2009 applies a similar perspective to art more generally.

<sup>84</sup> Dwyer 1999: 172

rules of etiquette and rules of morality? The typical Poverty of the Stimulus argument is that these tasks are far too complex to be the result of simple observational learning unless we have a richly pre-specified understanding that guides our exploration of the normative domain. So it is posited that we are born equipped with a normative reasoning capacity.

The structure of this psychological faculty is sometimes explained via a 'principles and parameters' approach, which plumbs the linguistic analogy fairly deeply. Nativists don't deny that the environment can influence the specific norms that we acquire, just as it would be foolish to deny that where we grow up influences which specific language ends up being our mother tongue. But this experience merely 'sets the dials' on our normative repertoire, which is argued to be invariant in significant ways. For example, Hauser presents evidence to suggest that complicated moral principles like the Doctrine of Double Effect are implicitly understood and applied by individuals all over the world, even those who have no training in ethics.<sup>85</sup> Research of this sort is intended to show that there are universal features of the way we reason about norms which are evidence of an innate 'normative grammar'. This grammar involves some general high level principles which provides us with a body of implicit knowledge about norms that is a necessary component of the psychological structure that underpins our capacity for normative cognition.

With this nativist account on the table, we can turn to assessing its plausibility. There are a couple of reasons to doubt the view that the capacity for normative cognition is founded on something like a normative grammar. To begin with, the universality of certain patterns of normative judgments don't necessarily count in favour of an innate capacity. Even if we grant the empirical assumption that many individuals across the world share characteristic intuitions about some sets of normative questions, it doesn't follow that these intuitions are produced by an invariant normative grammar. It

---

<sup>85</sup> Hauser 2006:177-262 provides a comprehensive presentation of this sort of evidence, particularly in reference to the famous 'Trolley Problem' which has enjoyed extensive airtime in this debate.

may be the case that some fairly sophisticated set of principles and parameters would consistently explain this pattern of judgments. But it is quite another thing to use this as evidence for the existence of psychological mechanisms that are causally responsible for producing these judgments. When the nativist claims that implicit grammatical knowledge structures our normative thinking they imply the strong claim that our judgments are likely to be caused by this particular cognitive mechanism. But pointing to the universality of judgments doesn't count in favour of this hypothesis, since it doesn't exclude alternative mechanisms which may have produced the very same judgments. And some of these other mechanisms seem like they don't require anything resembling a body of innate knowledge to produce similar patterns of empirical results. For example, Josh Greene & Jonathan Haidt have produced research<sup>86</sup> which emphasises the extent to which emotion and affective systems in the brain are a significant driving force in the process of moral judgment, systems that are conspicuously different from those invoked by the nativist. This criticism is that the role of normative grammar is too prominent in the nativist account of normative cognition.

Another criticism is that the nativist account isn't capable of explaining the role of reflection and discussion that our normative thinking involves.<sup>87</sup> This objection targets the core of the linguistic analogy used to support the nativist case. Our linguistic competence may well be unexplainable without a body of innate syntactic knowledge which helps structure the way we learn languages. But this knowledge is inaccessible to most competent language users. Absent formal linguistic training, people can't really explain why they judge that certain grammatical sentences are correct and others are not. But no such barrier applies when it comes to normative thinking. People can and do articulate the process that leads them to form particular moral, aesthetic or epistemic judgments, even if this might fall short of individuals possessing a full and coherent set of explicit normative principles. This is a problem for a nativist view because it is unclear just how ordinary normative

---

<sup>86</sup> Greene & Haidt 2002 summarises some of this work.

<sup>87</sup> Sterelny 2010

reflection interacts with the innate parameters that come pre-installed. If these parameters can be dramatically altered by explicit discussion and thinking of the sort that seems commonplace, then the hypothesis that normative cognition is structured similarly to linguistic cognition looks shaky. The normative grammar approach doesn't account for the reflective nature of normative thought.

Both these objections to the nativist account contend that normative cognition is unlikely to be explained by positing an innate and specific psychological structure. These objections shouldn't be interpreted too broadly though. Nichols makes the useful distinction between the positive and negative conclusions that Poverty of the Stimulus arguments are deployed to establish.<sup>88</sup> The negative conclusion is that a particular competence can't be acquired through empiricist learning and the positive conclusion is that, as a consequence, domain-specific psychological mechanisms are required to explain this competence. The arguments presented in this section do little to weaken the negative conclusion. It is likely that the capacity for normative cognition isn't grafted onto an entirely blank slate; understanding and processing rules is probably too complex a task for that. But they do provide reason for rejecting the positive claim because they cast doubt on the specific mechanism posited by nativism. In short, we can retain the idea that there is an innate capacity of some sort while resisting the idea that this is best explained via the structure of a grammar. So we need to look elsewhere for an account of normative thinking.

---

<sup>88</sup> Nichols 2005



### **3.3 – Socially Embedded Normative Guidance**

The previous section examined a proposal that explained normative cognition in terms of how it operates in humans today. In this section, we will examine a proposal that looks to explain the capacity for normative cognition in terms of its evolutionary function. This view is the one outlined and defended by Philip Kitcher recently in his book *The Ethical Project*.<sup>89</sup> Kitcher espouses what he calls a pragmatic naturalism that attempts to conceive our ethical practices today as gradually evolving over the course of tens of thousands of years. He wants to argue that the original function of ethical discussion usefully illuminates an ongoing project we are still engaged in. To that end, he offers a history of what he calls “the capacity for normative guidance”. He argues that the capacity for normative guidance began as a response to altruism failures. To overcome these, our ancestors acquired the ability to obey commands based on the threat of punishment. This led to the development of a conscience. In conditions of social equality, the subsequent elaboration and evolution of a system of rules formed the functional basis of the normative discussions we continue to have to this day. This is a short and rough precis of Kitcher's analytical history, which I will return to very shortly. My main contention in this section will be that this picture isn't sufficiently rich to fully explain the nature of normative cognition.

Before laying out this picture in slightly more detail, it is worth pausing to consider why I have picked out this particular view to discuss, rather than trying to reconstruct a more general view as I did in the previous section. To begin with, though Kitcher is ultimately occupied with trying to ground a particular ethical stance, his approach emphasises the continuity between morality and other normative domains that has been a characteristic theme of my argument. This is strongly suggested by his analysis in terms of the notion of normative guidance, which is congenial to the approach I have taken in trying to explore the nature of normative cognition generally. In addition,

---

<sup>89</sup> Kitcher 2011

Kitcher's specific attempt to explain the genealogy of normative guidance places a great deal of emphasis on its place as a form of social technology that has a public function. This too displays coherence, at least at a broad level, with my aim of trying to find an account which incorporates the social foundations of normative judgment which I consider to be crucial. For these two reasons – in addition to its recency, scope and rigour – Kitcher's view is one that provides an important foil to my own. A careful discussion of his account is more useful than a general treatment of the surfeit of evolutionary views that have been proposed in recent years.

Kitcher's starting point is early hominids who are capable of psychological altruism, which is the ability to form intentions which directly promote the wishes or interests of other people. This is carefully distinguished from biological altruism – acting in ways that promote the reproductive success of another organism – and behavioural altruism – acting in ways that benefit others because this is judged to be in the agent's extended self-interest. The notion of psychological altruism is meant to cover those cases where an agent is capable of forming truly other-directed desires and is a minimal requirement for creatures that participated in an extensive cooperative social life. Kitcher recognises that this is a "preethical state" and very far from a fully normative life but is a necessary precondition for its emergence. The characterisation he offers is one where "the limitations of their psychological altruism cause the tensions of their social lives and prevent them from....participating in more complex cooperative projects".<sup>90</sup> This is a result of altruism failures where the range or scope of other-directed desires is limited in crucial ways which prevent deeper social coordination. In an environment like this, there is strong adaptive pressure to find a mechanism which can strengthen altruistic behaviour in order to reap the benefits of social living.

The solution to this predicament was the acquisition of the ability to follow orders backed by the threat of social sanction. Normative guidance emerges here: "conceived in terms of the difference

---

<sup>90</sup> Ibid: 67-68

made to one's action guiding preferences by the recognition of commands".<sup>91</sup> The outcome of a social order governed by rules in this sense is to supplement our natural but limited inclinations to play nice with an external bulwark against bad behaviour. This produces altruistic behaviour in more contexts and strengthens those impulses in existing contexts. Obviously, a mechanism like this can't have got going without some sort of enforcement mechanism and here Kitcher relies on social punishment. Accounts of this sort traditionally face problems, since a stable system of punishment is thought to require a system of rules to be already in place, creating an unwanted regress. Who will punish those who fail to respect the punishment norms and let the initial norm-violators get off lightly – as the objection goes. In response, Kitcher draws heavily on work by Boyd & Richerson<sup>92</sup> which provides game-theoretic models of how retributive punishment can work to stabilise punishment. This is deployed as a proof that punishment could possibly have evolved, which is deemed sufficient for purposes in his account.<sup>93</sup>

The last crucial step in the analytical history is to explain how this rudimentary normative guidance paves the way for the fuller conception of the ethical project. Here Kitcher stresses two important points. The first is to articulate a liberal conception of how normative guidance operates, which rejects the notion that only a faculty of cool reason is invoked in overcoming our dispositions. In doing so he welcomes the idea that a conscience can form as a result of a "hodgepodge of considerations and feelings, and it is foolish and unnecessary to limit the full range of psychological possibilities".<sup>94</sup> To put it in terms of normative theory, we might say that the internalisation of commands is likely to involve both desires and beliefs. But secondly, normative guidance is socially embedded, meaning that its development in individuals was always guided by the groups they were a part of. The reason for this is that the ethical project first arose in the unusual circumstances of

---

<sup>91</sup> Ibid: 75-76

<sup>92</sup> Boyd & Richerson 1992

<sup>93</sup> Kitcher 2011: 89-90

<sup>94</sup> Ibid: 82

relative equality amongst small bands of hominids who faced shared survival challenges, and “the codes thus devised are *social* products: they represent a *joint* reaction to the altruism failures previously afflicting the group”.<sup>95</sup> This communal picture figures centrally for Kitcher as it serves as the source of authority for our capacity for normative guidance.

This evolutionary account certainly has much to recommend it. Kitcher's emphasis on the complexity of normative psychology and its social dimensions are welcome and share affinities with the approach I endorsed in the previous chapter. In particular the connection of normative guidance with social punishment provides an idea of how a focus on the behavioural coordination function of normative discourse might be explained, which is important because it gets to the heart of the practicality platitude. However, there are a couple of points of caution I want to explore. These involve questions first about the explanatory resources of the model and secondly about the target of his analysis.

The first questions are about the analytical history. Is the path from altruism failure to normative guidance really convincing? Here it is important to note that Kitcher's ambitions are perhaps more modest and in the spirit of his pragmatic naturalism he claims to be giving a mere 'how possibly' explanation which is compatible with the best evidence on hand. But if this hypothetical account is to serve as a serious foundation for an account of normative cognition then it is appropriate to scrutinise it a little bit. We don't have to hold it to a 'how actually' standard of proof, but a 'how probably' standard seems fair. The most vulnerable aspect of the history is the contention that normative guidance became socially embedded when our distant ancestors were able to sit down together and articulate ethical codes in the cool hour around the campfire. This picture seems to assume too much about the social order that Kitcher is trying to explain. There are serious questions about whether social settings of equality could stabilise given what we know about punishment in

---

<sup>95</sup> Ibid: 98

other hominid lineages, where violence is often used as tool to entrench hierarchy.<sup>96</sup> Even if this weren't a problem, a picture of deliberators who engage in normative discussion about ethics seems to presuppose an already complex repertoire of normative concepts to structure that discussion, like norms about how to decide what counts as a rule and how to change rules.<sup>97</sup> This critique is that Kitcher's account of the crucial transition from a pre-normative condition is underexplained but it might be easily met by simply enriching the descriptive account, while leaving the basic structure of the explanation intact.

This leads to the second set of questions we might pose of the analytical history. Simply put, is normative guidance the same thing as normative cognition? There is reason to think that there is still a bridge waiting to be crossed. Kitcher claims that the evolutionary function of normative guidance was to remedy altruism failures. But normative cognition has a broader function than that – to enable us to process rule-bound norms in a very general sense as has been the target of this chapter. Buchanan puts this point as a complaint that Kitcher fails to explain open-ended normativity and views it arising from an equivocation in the function of normative guidance. Sometimes it is framed as a response to altruism failures only insofar as they cause social costs, and at other times it is framed as a response to altruism failures simpliciter.<sup>98</sup>

Granted, there are other functions to which morality could be put. It isn't an objection to just show that normative cognition in current humans is used in contexts outside of those involving social costs. The problem for Kitcher though is that he explicitly wants to ground the historical progression of the ethical project in this specific account of its origins. There needs to be an explanation of how the many other domains of human normative discourse *could possibly* have arisen from this picture of campfire moralising. This is what would allow us to argue that socially-

---

<sup>96</sup> A review of this evidence is provided by Sterelny 2012a.

<sup>97</sup> Baumann 2012

<sup>98</sup> Buchanan 2012

embedded normative guidance might lead to generalised normative cognition, even if its roots are in ethics. But this isn't the sort of general reflective setting suggested by the model. Kitcher doesn't explain how our lineage started to engage in discussion on a host of issues that have less to do with social disorder but still count as normative – such as those discussed in the prudential or epistemic domains which are an important part of what we are trying to explain.

Kitcher himself might object to the idea that there really is such an independent capacity for reflection implicated by our normative discussion. He is certainly very emphatic when it comes to rejecting broadly rationalist views within moral philosophy, saying that “the entire conception of the ethical point of view is a psychological myth devised by philosophers”.<sup>99</sup> Even if we want to grant that point (which would be hasty) a broader skepticism about the concept of the normative point of view is harder to stomach. This analytical history might be a good way to explain the emergence of the ethical project, but it falls short of doing the same for the normative project more generally.

---

<sup>99</sup> Kitcher 2011: 81

### **3.4 – The Modified Sripada & Stich Model**

In the last two sections I have considered two different accounts of the capacity for normative cognition, and concluded that neither is entirely satisfactory. In this section I turn positive, and outline a framework that is stronger and better accounts for the phenomenon. Recently, Chandra Sripada and Stephen Stich have proposed an account of the process which underpins normative psychology.<sup>100</sup> This theoretical model is intended to provide a framework which incorporates existing research and can guide future investigation in the area. In what follows I will defend a modified account of the model, explaining why the basic Sripada & Stich (henceforth S & S) Model provides a useful structure of normative cognition by carefully addressing the functions it serves. Then, in context of the discussion of the previous two sections, I will argue that this model should be slightly modified in order to better account for the social foundations of normative judgment. My main contention is that the resulting Modified S & S Model provides the most coherent general account of the capacity for normative cognition.

At its core, the S & S model posits two linked psychological devices: a Norm Acquisition Mechanism and a Norm Implementation Mechanism. These two mechanisms serve separate but interconnected functions in normative psychology. The acquisition mechanism is mainly responsible for detecting norms that prevail in the cultural environment and inferring the content of those norms. It plays the role of explaining how we can reliably latch on to the norms of the world around us. The implementation mechanism is mainly responsible for storing norms and then executing behaviour that those norms prescribe, both by motivating individuals to comply and motivating them to punish those who violate norms. It plays the role of explaining why these norms have some motivational force for individuals. Both these mechanisms are meant to be innate. Below is a diagram which illustrates the essentials of the S & S Model.

---

<sup>100</sup> Sripada & Stich 2007

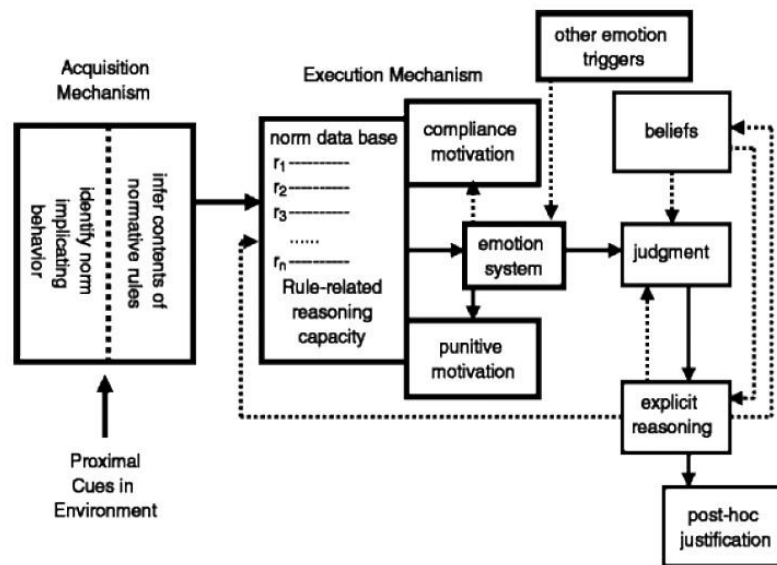
Figure 1: A Basic Sketch of the S & S Model<sup>101</sup>

This model provides a good functional specification of the kind of psychological architecture that could subserve the capacity for normative cognition. Though it is very simple, by breaking down the capacity to simpler processing tasks it provides a nice illustration of the structure required. The innateness claims it makes are minimal as well, since it only proposes a basic structural competence rather than any rich knowledge about particular normative domains. As Sripada & Stich point out, this initial model raises more question than it answers, but does so in a way that provides a framework for addressing those questions. They go on to explore some of these questions, most notably the role that emotion and explicit reasoning play in this framework. Citing an enormous body of research, they conclude that emotional systems most likely mediate both punishment and compliance motivations and also that emotions play a role in moral judgment. In addition they think that explicit reasoning about our judgments can be both a cause and effect of these emotional processes since they are influenced by emotions but can also generate new rules and norms themselves. The details need not detain us here, but below is a diagram that illustrates how they view a host of supplementary evidence as figuring into the model.

---

<sup>101</sup> Ibid: 290



Figure 2: A sketch of the elaborated S & S Model<sup>102</sup>

The first issue with this more elaborate model is that in trying to incorporate more data it loses much of the functional clarity that made it useful in the first place. This is partly due to an overemphasis on incorporating results strictly from moral psychology into a model designed to explain normative psychology more broadly. For example, the cluster of boxes pertaining to explicit reasoning are added as a response to research on moral dumbfounding experiments that show that emotional responses trump explicit reasoned judgments in certain circumstances. But it is not clear whether these results imply a peripheral function for explicit normative reasoning in the way that the model implies. In addition, though it is highly probable that emotional systems are indeed involved in the implementation of norms, the model makes no suggestion about the extent to which these same systems might play a role in the acquisition of norms. This is despite Sripada and Stich's informed discussion of some of the very evidence that is suggestive of constraints and biases in the way humans learn and process rules towards the end of their article.<sup>103</sup> In general, the elaborated model seems to give a too specific account of the implementation mechanism while being too bare

<sup>102</sup> Ibid: 297<sup>103</sup> Ibid: 298-301

when it comes to the acquisition mechanism.

The second issue with the elaborated model is that it is inattentive to the interaction between the individual and the cultural environment they are a part of. The only inputs into the system are proximal cues from the environment into the acquisition mechanism and little is done to explain how this process might work, though a range of options are possible from explicit reasoning, instruction, or individual inference. Also nowhere in the model is there any suggestion that there might be any input from the environment to the implementation mechanism, an omission which seems odd given the plausible role that social punishment can play in supporting intrinsic motivation to comply with norms. Moreover, there is no real consideration of the role that an individual's normative psychology might play in producing output that shapes the cultural environment. On the implementation side this seem like it should be a natural part of the model to make sense of the role of punitive motivations, and on the acquisition side the oversight is the role that discussion about social rules might play in influencing the norms that prevail.

This second criticism gets at a more fundamental issue. In general the S & S Model under-emphasises the social foundations of the capacity for normative cognition that I have argued serve an important role. In the last chapter I argued that both communal content and behavioural coordination are important elements of normative discourse on a conceptual level. Even though this psychological model is empirical, we should still expect that it coheres with these elements, particularly given the way that these emerge from an internal reading of the evidence as well. So the earlier discussion of nativism illustrated that the ability to understand normative rules requires making sense of the interaction between individuals and their social environment, in light of the factualist platitude. And Kitcher's account of the evolution of normative guidance gave special prominence to the importance of social punishment in stabilising a system of ethical norms. Though I may have raised issues about the details of both of these accounts, I found their general approach

to be sound. So the fact that the S & S Model does so little to incorporate these aspects of normative cognition is a significant weakness.

However, this doesn't require a wholesale rejection of the model. By altering the framework slightly, it can be strengthened to provide a better account of normative cognition. Below is a diagram of a Modified S & S Model that both consolidates some of its excesses and extends it to better account for the interaction between individual and the cultural environment.

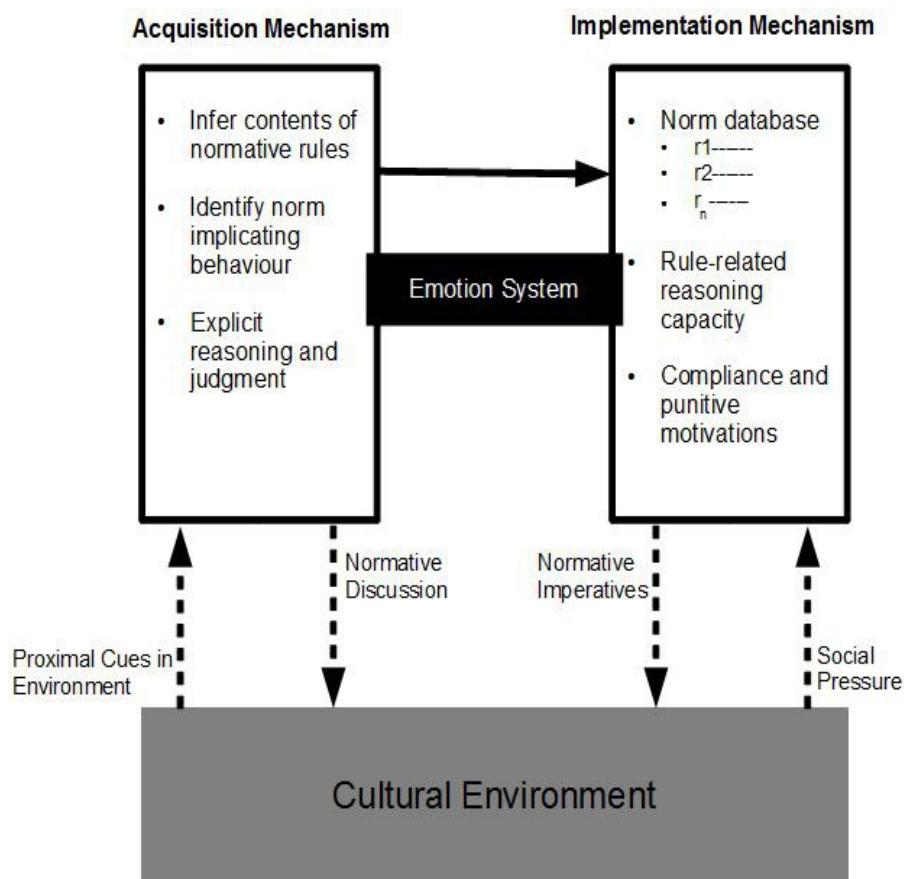


Figure 3: A sketch of a Modified S & S Model to explain normative cognition

This modified model has a number of features worth noting. The first is that it strips away much of the more complex additions that surrounded the implementation mechanism in Figure 2. Most of this didn't helpfully clarify the basic structure of normative psychology and some of the

elaborations are incorporated into the diagram in different ways. For instance, the role of explicit reasoning has its place as a direct part of both the acquisition and implementation mechanisms. This befits its importance without committing the model to over-specific hypotheses about the role it plays in normative psychology. Similarly, and secondly, the modified model views the Emotion System as independent and able to influence both mechanisms, not just the implementation of norms. This is visually represented by its position overlapping both mechanisms in the diagram. These two features don't involve any additions to the model, but rather capture the framework that Sripada and Stich seem to have in mind in a simpler way. The lists of functions it attributes to each sub-system are the provisional hypotheses it is committed to. The task of testing how these fare can be roughly thought of as the task of identifying the process by which these functions operate in the brain.

Where I have made a substantial addition to the model is with the grey box representing the Cultural Environment. The dotted arrows between it and the two mechanisms represent my provisional attempt to model the interactions between the individual and their surroundings that are important to understanding fully-fledged normative cognition. On the acquisition side, the input from the environment retains the cues that are used to infer the prevailing norms, much like the original model. But the output to the environment is labeled 'Normative Discussion' to represent the role that individuals have in not just accepting but playing some role in shaping the prevailing norms, by contributing to the social project that is involved in constructing a public body of rules. On the implementation side, the input from the environment is 'Social Pressure' – that process by which influence from those around us can supply motivations to comply with the normative rules we have internalised. But we aren't just recipients of this pressure, and the output arrow is labelled 'Normative Imperatives' to represent the role that individuals can play in creating just this same sort of pressure on those around them.

These additions serve as a preliminary sketch of the ways in which the social foundations of normative discourse might be incorporated into the basic psychological framework of the S & S Model. It makes particular commitments about the structure by which this operates. And these hypotheses are guided by the conceptual discussions of the previous chapter. The Proximal Cues/Normative Discussion pair on the acquisition side is an empirical hypothesis about the operation of the factualist platitude. The Social Pressure/Normative Imperatives pair on the implementation side is an empirical hypothesis about the operation of the practicality platitude. By delineating the 'internal' mechanisms of individual thought from the 'external' mechanisms of social thought, this model provides an account of normative cognition that is consistent with what we know, while focusing attention on the communal functions that are important to explain. But is there any evidence that supports this modified model and can start to put some flesh on this skeleton? I conclude this chapter by presenting some suggestive evidence in favour of this model.

### **3.5 – Some Evidence for the Model**

In this section I will, very briefly, present some recent evidence which supports the model that I offered in the previous section. I will use this to illustrate the plausibility of the hypotheses I ended the last section with. I will look at a research programme led by Tomasello that stresses the importance of shared intentionality and use this to support the importance of output from individuals to their environment. Then I will look at an apprentice learning model of human cognition presented by Sterelny and use this to support the importance of the input from the environment to individuals. This completes a sketch of an approach which vindicates the merits of a focus on the social foundations of normative thought.

Michael Tomasello and his colleagues have produced a body of research<sup>104</sup> that aims to illustrate that shared intentionality is a key species-unique form of cognition. It is essentially the ability to participate with others in collaborative activities that have shared goals, which requires not only the ability to understand others' intentions but motivations to share psychological states with others. This is because participation in shared activities involves public representations of the goals of these endeavours. There needs to be mutual knowledge that parties are collaborating and a coordination of their plans in order to achieve joint goals. They argue that these capacities have analogues in other primates who can participate in rudimentary social learning. But the crucial competence they lack is the ability and motivation to share aspects of their inner mental lives with those around them. By contrast, infants from a very early age demonstrate this ability through behaviour like gaze-following and gestures like pointing in the direction of objects they want to draw attention to. In addition to marshalling a host of evidence for this developmental claim, Tomasello et al also propose an evolutionary path that explains why this can serve as the basis for the development of a

---

<sup>104</sup> Tomasello 2009; Tomasello et al 2005; Tomasello & Carpenter 2007

normative community.<sup>105</sup> A first step would have been the context of collaborative foraging that selected for the skills of shared intentionality. But as these initial societies became larger, groups would face competition from other groups. This meant that the initial skills of shared intentionality were 'scaled-up' to help foster a sense of group mindedness that involved the creation of collective social institutions bound together by explicit norms.

This is only the barest possible sketch of an expansive range of research. But it is enough to trace why this account helps support my model of normative cognition. Notably, it provides evidence for the two output arrows in Figure 3 that represent the influence that individuals have on the cultural environment. Normative discussion emerges as part of what is required for individuals to form public representations of joint plans. Coordinating social behaviour through collaboration becomes an integral part of the way that humans acquire norms. But this doesn't just occur through discussion but also by sharing one's emotions and using these to direct the motivations of those around us as well. This can be seen as an illustration of the normative imperatives which links the implementation system with the environment and another way that individuals can influence the prevailing norms around them. Tomasello's account of shared intentionality gives us a good explanation of not just the basic structure of normative cognition but also the way that actively engaging in the social world is a part of this capacity.

By contrast, Kim Sterelny presents an evolutionary model of cultural transmission that helps see how the individual is shaped by their social environment.<sup>106</sup> The key idea is that the human lineage possesses a non-genetic inheritance mechanism by which skills and information can be passed down the generations. The environment that individuals are exposed to is seeded with information that means they don't have to learn new skills from scratch. This can begin as a side-effect of adult

---

<sup>105</sup> Tomasello et al 2012

<sup>106</sup> This is drawn from Sterelny 2012b, particularly Sections 2.3 & 3.3, and Chapter 7.

behaviour and leads to a process of apprentice learning which allows the accumulation of cultural knowledge across generations. Though the apprentice model is primarily intended to explain the evolution and retention of technical expertise, it is extended to account for the growth of norms in human evolution. Once a robust system of cultural inheritance is in place, selective pressure can lead to adaptations which help engineer capacities that make it easier to engage in normative thought. For instance, adults can engage in explicit instruction and involve children in the evaluative life of the community, and children who are sensitive to the emotional rhythms of the individuals that surround them will find it easier to participate in these societies. In essence, apprentice learning can lead to minds that are biologically prepared for normative cognition. Norms are a product of a synergistic interplay between biological and cultural mechanisms of inheritance.

Again, this is an extremely brief precis of an account which integrates research from a wide array of anthropological and ethnographic sources. But it provides us with some evidence that supports the model of normative cognition proposed in the previous section. The apprentice learning approach is consistent with a picture that sees the cultural environment having a role in feeding input to both the norm acquisition and implementation mechanisms. On the acquisition side, the role of explicit instruction and the scaffolding of an individual's learning environment is an example of a process that can be thought to make it easier to learn the norms that prevail around us. On the implementation side, being attuned to the emotional life of the worlds we are a part of can be a way that we are structured to be susceptible to social pressure that helps support our compliance with the norms we have internalised. So Sterelny's picture is consistent with the input side of the model, that predicts that our cultural environment will have a significant influence on the operation of normative cognition.

This breezy sketch of research is obviously far from a decisive vindication of the Modified S & S Model. But this evidence is consistent with the model, and can be interpreted to support the key



additions I made about the relationship between the individual and the cultural environment in the operation of normative cognition. So it is sufficient to demonstrate that there is a potentially fruitful line of empirical inquiry that can be guided by the framework that I offered.

### **3.6 – Conclusion**

This chapter defended a model of normative cognition that was informed by my conclusions about norms and normative judgments in the previous two chapters. I started by outlining a basic social-scientific definition of norms. This treated them as being linked to rules in a way that cuts across different normative domains, an empirical treatment that resonates with my conceptual analysis of norms in section 1.1. I then examined two attempts to explain the capacity for normative cognition. The first was a nativist account that made claims about the developmental basis of the capacity and the second was an evolutionary account that made claims about the origin of the capacity. Both these accounts were critiqued, although they each illustrated some important features of an adequate psychological model of normative cognition. I used these to inform the positive account of normative cognition that I presented in the final two sections. This model outlined a plausible psychological structure that could aid the understanding of normative cognition that places emphasis on the social elements that underpin it.

---

## Conclusion

---

At the start of this thesis I staked my ground in the everyday, quotidian exchanges where norms are important in our social discourse. The task I set myself was to explain how these social exchanges provided glimmers of insight into the very operation of normative judgment. In this conclusion, I will explain how the claims I have defended come together to constitute a sustained argument in favour of this central claim. Though the specific points of emphasis of my three chapters were disparate to a certain extent, together they built support for the main purpose of this thesis, which was to demonstrate that an understanding of normative judgment requires a direct understanding of its social functions. The way they demonstrated this is best understood in light of the two broad themes that I outlined in the introduction. The first of these was the insistence that normativity is best understood as a general conceptual phenomenon, rather than being balkanised into separate sub-domains. The second of these was the insistence that empirical research is a necessary supplement to our understanding of normativity.

In Chapter 1, I provided a conceptual account of norms and normative judgment that primarily bolstered the first theme of the thesis. There were two arguments in it that were important. The first argument was the analysis of norms, which held that they were a structural relation that delivers a verdict about the correctness of a response relative to an evaluative framework. This general analysis was supported by observations about the unity of different normative domains in our everyday discourse. The second argument was that Normative Judgment Internalism was the best account of the role that norms play in human thought. This established that sincerely accepting a

normative statement necessarily implies an endorsement of a normative reason. This was defended on the basis that when we try to predict when the responses of other intentional agents 'make sense', this necessarily involves a normative appraisal. These two arguments were important in setting the stage for Chapter 2. But additionally, they demonstrated the utility of thinking about normativity as a general phenomenon, rather than as loosely connected sub-domains of morality, prudence, aesthetics and so on. This was the important contribution of Chapter 1 to the overall aim of the thesis.

In Chapter 2, I explored accounts of how normative judgment works, in the context of a puzzle that arises from the clash between our intuitions about the factual and practical platitudes about normative discourse. The issues I explored here stood at the nexus of the two themes of the thesis. I made one important negative argument, and then offered two positive arguments. The negative argument was that after tracing both Humean and anti-Humean attempts to resolve the normative puzzle, it was clear that both of these orthodox approaches converged on a solution that failed, since it tried to explain a descriptive phenomenon by positing an unexplained conceptual ability. As a result, I offered two suggestions for reform. The first positive argument was that conceptual theories of normative judgment would be enriched by incorporating the social functions of normativity, by explaining the factualist and practicality platitudes in more detail. The second positive argument was that naturalising the study of normativity would provide another useful avenue to explain the unusual conceptual ability humans seem to possess. The arguments in Chapter 2 were very important to the overall purpose of the thesis. They continued on from the issues presented in Chapter 1, and so were a further illustration of the way investigating the nature of normative judgment in a general fashion was a worthwhile enterprise – the first theme of the thesis. But in addition they demonstrated why an empirical understanding of normative judgment was so important, by showing that even a general conceptual explanation of the phenomenon reaches its

limits – so began an exploration of the second theme of the thesis. The connection between the two themes that was evident in Chapter 2 is important, and one that I will return to shortly.

In Chapter 3, I explored empirical research on the nature of normative cognition and defended a specific model of its operation: the Modified Sripada and Stich Model. This put the conclusions of Chapter 2 into action in service of the second theme of the thesis. The S & S Model was an empirical model of the normative cognition that emphasised the interaction between an individual and their cultural environment. I presented it as an upgrade of two theories of the capacity for normative cognition, one that was developmental and treated it as innate and another that was evolutionary and linked it with institutions of social punishment. I then marshalled some evidence in favour of the model to demonstrate its plausibility. As I noted, the model I defended was a working hypothesis about the psychological structures that underpin normative cognition, one that emphasised the social elements of normativity. As more evidence accumulates, some details of the model will surely need to change, and other details may remain intact. But the importance of the findings of Chapter 3 for the overall argument of the thesis was to demonstrate the way that empirical research can contribute to an understanding of the nature of normative judgment – a vindication of the second main theme of the thesis.

The preceding three paragraphs trace the ways that the arguments of the thesis come together to assemble the case that understanding normative judgment requires an understanding of its social foundations. What is the importance of this claim? We can begin to tease this out by observing a couple of potential connections that emerge when the findings of Chapter 3 are brought to bear on the arguments of Chapter 1. The capacity for normative cognition is explained by the S & S Model as emerging from the two way interaction between an individual and their environment, both in development and in evolutionary history. Importantly, this capacity is a general one, not restricted to any particular normative domain. If an empirical model like this was developed further, it could

help to explain why we observe the similarities that were the basis of my argument in favour of a unified conceptual analysis of normative domains. Additionally, a crucial argument in Chapter 1 was the interconnection between the intentional and normative perspectives of interpretation. An empirical model of the kind I offered, if developed in much more detail, might be able to explain how this is the case by providing a path for exploring the emergence of normative cognition in evolutionary or developmental terms. Of course, neither of these speculative points is established by this thesis. But they are potentially fruitful lines of inquiry, and they would be areas that would be ripe for investigation if the approach of this thesis was carried further.

Norms are both a singularly important and incredibly puzzling part of our lives and societies. This thesis has attempted to contribute to understanding how these norms work. By blending our best conceptual accounts with our most detailed empirical observations the nature of our engagement with norms comes into better focus. Normative judgment both enables, and is enabled by, the sociality that is at the heart of the human condition.

---

## References

---

- Altham, J.E.J (1986). "The Legacy of Emotivism" in G. Macdonald & C. Wright (eds.), *Fact, Science and Morality: Essays on A.J Ayer's Language, Truth and Logic*, Oxford: Blackwell.
- Anscombe, G.E.M (1957). *Intention*, Oxford: Blackwell.
- Ayer, A.J. (1936). *Language, Truth and Logic*, London: Gollancz.
- Baurmann, Michael (2012). "The Golden Age of the Campfire: Should We Take Our Ancestors Seriously", *Analyse & Kritik*, 34: 39-50.
- Blackburn, Simon (1993). *Essays in Quasi-Realism*, Oxford: Oxford University Press.
- Blackburn, Simon (1998). *Ruling Passions: A Theory of Practical Reasoning*, Oxford: Clarendon Press.
- Boyd, Brian (2009). *On the Origin of Stories: Evolution, Cognition and Fiction*, Cambridge: Belknap Press.
- Boyd, Robert & Richerson, Peter (1992). "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups", *Ethology and Sociobiology*, 13: 171-195.
- Brink, David (1986). "Externalist Moral Realism", *Southern Journal of Philosophy*, Supplement: 23-42.
- Brink, David (1989). *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.
- Brink, David (1997). "Moral Motivation", *Ethics*, 108(1): 4-32.

- Buchanan, Allen (2012). "The Open-Ended Normativity of the Ethical", *Analyse & Kritik*, 34: 81-94.
- Chomsky, Noam (1980). *Rules and Representations*, Oxford: Blackwell.
- Cosmides, Leda & Tooby, John (2008). "Can a General Deontic Logic Capture the Facts of Human Moral Reasoning? How the Mind Interprets Social Exchange Rules and Detects Cheaters." in W. Sinnott-Armstrong (ed.), *Moral Psychology: Volume 1: The Evolution of Morality*, Cambridge, MA: MIT Press, 53-120.
- Cuneo, Terence (2007). *The Normative Web: An Argument for Moral Realism*, Oxford: Oxford University Press.
- Darwall, Stephen (1983). *Impartial Reason*, Ithaca, NY: Cornell University Press.
- Davidson, Donald (1963). "Actions, Reasons, and Causes", *The Journal of Philosophy*, 60(23): 685-700.
- Dennett, Daniel (1987). *The Intentional Stance*, Cambridge, MA: MIT Press.
- Dutton, Dennis (2009). *The Art Instinct*, New York: Bloomsbury Press.
- Dwyer, Susan (1999). "Moral Competence" in K. Murasugi & R. Stainton (eds.), *Philosophy and Linguistics*, Westview Press, 169-190.
- Dwyer, Susan (2007). "How Good is the Linguistic Analogy?" in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind: Volume 2: Culture and Cognition*, Oxford: Oxford University Press, 237-256.
- Falk, W.D. (1947). "'Ought' and Motivation", *Proceedings of the Aristotelian Society*, 48: 111-138.
- Finlay, Stephen (2007). "Responding to Normativity" in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics: Volume 2*, Oxford: Oxford University Press: 220-239.
- Finlay, Stephen (2009). "Oughts and Ends", *Philosophical Studies*, 143: 315-340.
- Finlay, Stephen (2010). "Recent Work on Normativity", *Analysis*, 70(2): 331-346.
- Foot, Phillipa (1972). "Morality as a System of Hypothetical Imperatives", *Philosophical Review*,



81(3): 305-316.

Frankena, William (1958). "Obligation and Motivation in Recent Moral Philosophy" in A.I. Melden (ed.), *Essays in Moral Philosophy*, Seattle: University of Washington Press, 40-81.

Gert, Joshua (2001). "Skepticism about Practical Reasons Internalism", *The Southern Journal of Philosophy*, 34: 59-77.

Gibbard, Allan (1990). *Wise Choice, Apt Feelings: A Theory of Normative Judgment*, Oxford: Clarendon Press.

Gibbard, Allan (2003). *Thinking How to Live*, Cambridge, MA: Harvard University Press.

Greene, Joshua & Haidt, Jonathan (2002). "How (and where) does Moral Judgment Work?", *Trends in Cognitive Science*, 6(12): 517-523.

Hampton, Jean (1998). *The Authority of Reason*, Cambridge: Cambridge University Press.

Harman, Gilbert (1975). "Moral Relativism Defended", *Philosophical Review*, 84(1): 3-22.

Harman, Gilbert (2000). *Explaining Value and Other Essays in Moral Philosophy*, Oxford: Oxford University Press.

Hauser, Marc (2006). *Moral Minds: How Nature Designed our Universal Sense of Right and Wrong*, London: Abacus.

Hume, David (1740). *A Treatise of Human Nature*.

Hume, David (1751). *An Enquiry Concerning the Principles of Morals*.

Joyce, Richard (2001). *The Myth of Morality*, Cambridge: Cambridge University Press.

Joyce, Richard (2006). *The Evolution of Morality*, Cambridge, MA: MIT Press.

Joyce, Richard (2013). "The Many Moral Nativisms" in K. Sterelny, R. Joyce, B. Calcott & B. Fraser (eds.), *Cooperation and its Evolution*, Cambridge, MA: MIT Press.

Kant, Immanuel (1785). *Fundamental Principles of the Metaphysic of Morals*, trans. T.K. Abbott.

Kelly, Daniel & Stich, Stephen (2008). "Two Theories about the Cognitive Architecture Underlying

- Morality” in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind: Volume 3: Foundations and the Future*, Oxford: Oxford University Press, 348-366.
- Kitcher, Philip (1994). “Four Ways of “Biologizing” Ethics” in E. Sober (ed.), *Conceptual Issues in Evolutionary Ethics*, Cambridge, MA: MIT Press, 575-586.
- Kitcher, Philip (2011). *The Ethical Project*, Cambridge, MA: Harvard University Press.
- Korsgaard, Christine (1986). “Skepticism about Practical Reason”, *The Journal of Philosophy*, 83(1): 5-25.
- Korsgaard, Christine (1996). *The Sources of Normativity*, Cambridge: Cambridge University Press.
- Lewis, David (1989). “Dispositional Theories of Value”, *Proceedings of the Aristotelian Society*, Supplementary Volumes, 63: 113-137.
- McAdams, Richard (1997). “The Origin, Development, and Regulation of Norms”, *Michigan Law Review*, 96(2): 338-433.
- McDowell, J. (1978). “Are Moral Requirements Hypothetical Requirements”, *Proceedings of the Aristotelian Society*, Supplementary Volumes, 52: 13-29.
- Machery, Edouard & Mallon, Ron (2010). “Evolution of Morality” in J. Doris et al (eds.), *The Moral Psychology Handbook*, Oxford: Oxford University Press, 3-46.
- Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong*, New York: Penguin.
- Mallon, Ron (2008). “Ought we to Abandon a Domain-General Treatment of “Ought”” in W. Sinnott-Armstrong (ed.), *Moral Psychology: Volume 1: The Evolution of Morality*, Cambridge, MA: MIT Press, 121-130.
- Mikhail, John (2007). “Universal Moral Grammar: Theory Evidence and the Future”, *Trends in Cognitive Sciences*, 11(4): 143-152.
- Moore, G.E. (1903). *Principia Ethica*, Cambridge: Cambridge University Press.
- Nagel, Thomas (1970). *The Possibility of Altruism*, Princeton, NJ: Princeton University Press.
- Nichols, Shaun (2002). “Norms with Feeling: Towards a Psychological Account of Moral

- Judgment”, *Cognition*, 84: 221-236.
- Nichols, Shaun (2005). “Innateness and Moral Psychology” in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind: Volume 1: Structure and Contents*, Oxford: Oxford University Press, 353-430.
- Nucci, Larry (2001). *Education in the Moral Domain*, Cambridge: Cambridge University Press.
- Parfit, Derek (2011). *On What Matters: Volume 2*, Oxford: Oxford University Press.
- Pettit, Philip (1990). “Virtus Normativa: Rational Choice Perspectives”, *Ethics*, 100(4): 725-755.
- Phillipson, Nicholas (2010). *Adam Smith: An Enlightened Life*, New Haven, CT: Yale University Press.
- Railton, Peter (1986). “Moral Realism”, *Philosophical Review*, 95(2): 163-207.
- Scanlon, T.M. (1998). *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
- Schroeder, Mark (2007). *Slaves of the Passions*, Oxford: Oxford University Press.
- Slote, Michael (1971). “The Rationality of Aesthetic Value Judgments”, *The Journal of Philosophy*, 68 (22): 821-839.
- Slote, Michael (2012). *Moral Sentimentalism*, Oxford: Oxford University Press.
- Smetana, Judith (1993). “Understanding of Social Rules” in M. Bennett (ed.), *The Development of Social Cognition: The Child as Psychologist*, New York, NY: Guilford Press, 111-141.
- Smith, Adam (1759). *The Theory of Moral Sentiments*.
- Smith, Michael (1994). *The Moral Problem*, Oxford: Blackwell.
- Sripada, Chandra & Stich, Stephen (2007). “A Framework for the Psychology of Norms” in P. Carruthers, S. Laurence & S. Stich (eds.), *The Innate Mind: Volume 2: Culture and Cognition*, Oxford: Oxford University Press, 319-343.
- Sterelny, Kim (2010). “Moral Nativism: A Sceptical Response”, *Mind & Language*, 25(3): 279-297.
- Sterelny, Kim (2012a). “Morality's Dark Past”, *Analyse & Kritik*, 34: 95-115.

- Sterelny, Kim (2012b). *The Evolved Apprentice: How Evolution Made Humans Unique*, Cambridge, MA: Bradford.
- Stocker, Michael (1979). "Desiring the Bad: An Essay in Moral Psychology", *The Journal of Philosophy*, 76(12): 738-753.
- Sturgeon, Nicholas (1985). "Moral Explanations" in D. Copp & D. Zimmerman (eds.), *Morality, Reason and Truth*, Rowman & Littlefield, 49-78.
- Svavarsdottir, Sigrun (1999). "Moral Cognitivism and Motivation", *Philosophical Review*, 108(2): 161-219.
- Tomasello, Michael et al. (2005). "Understanding and Sharing Intentions: The Origins of Cultural Cognition", *Behavioral and Brain Sciences*, 28:675-735.
- Tomasello, Michael & Carpenter, Malinda (2007). "Shared Intentionality", *Developmental Science*, 10(1): 121-125.
- Tomasello, Michael (2009). *Why we Cooperate*, Cambridge, MA: MIT Press.
- Tomasello, Michael et al. (2012). "Two Key Steps in the Evolution of Human Cooperation: The Interdependence Hypothesis", *Current Anthropology*, 53(6): 673-692.
- Turiel, Elliot (1983). *The Development of Social Knowledge: Morality and Convention*, New York: Cambridge University Press.
- Wedgwood, Ralph (2007). *The Nature of Normativity*, Oxford: Clarendon Press.
- Williams, Bernard (1981). "Internal and External Reasons" in *Moral Luck*, Cambridge: Cambridge University Press.
- Williams, Bernard (1985). *Ethics and the Limits of Philosophy*, London: Routledge.
- Zangwill, Nick (1998). "Direction of Fit and Normative Functionalism", *Philosophical Studies*, 91(2): 173-203.
- Zangwill, Nick (2003). "Externalist Moral Motivation", *American Philosophical Quarterly*, 40(2): 143-154.