

# Humanity, Virtue, Justice: A Framework for a Capability Approach

By

Benjamin James Bessey

A Thesis submitted to the University of Birmingham  
for the degree of DOCTOR OF PHILOSOPHY

Department of Philosophy  
School of Philosophy, Theology, and Religion  
College of Arts and Law  
University of Birmingham  
October 2015

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

# Abstract

This Thesis reconsiders the prospects for an approach to global justice centring on the proposal that every human being should possess a certain bundle of goods, which would include certain members of a distinctive category: the category of capabilities. My overall aim is to present a clarified and well-developed framework, within which such claims can be made. To do this, I visit a number of regions of normative and metanormative theorising. I begin by introducing the motivations for the capability approach, and clarifying some of its most distinctive features. Next, I focus on Martha Nussbaum's version of the approach, and identify several problems therein. The most important concerns epistemology, and especially the challenges that constructivist theories pose. The middle part of the Thesis presents an alternative, based on the work of John McDowell, which I argue has superior prospects. Then, I turn to two further problems: that of making sense of the universalistic aspirations of cosmopolitanism, and that of integrating the microscopic prescriptions of ethics with the macroscopic analyses of political philosophy. Using the Aristotelian interpretation of its core framework that I have developed, I conclude that the capability approach can provide compelling answers to important questions about global justice.

# Acknowledgements

Firstly, I would like to thank my two current supervisors, Heather Widdows and Iain Law, and my former supervisor Alex Miller. All three have contributed significantly through inspiration, feedback, and support. In particular, I'd like to thank Heather for providing a constant backbone to my project, and being tolerant (perhaps too tolerant!) of my periodic lapses, and Iain for being endlessly reassuring and such a fine armchair psychiatrist. I owe a debt to many of my fellow students at Birmingham for stimulating conversation and friendship over my time here, especially Herj Marway, Naomi Thompson Paul Broadbent, Khai Wager, Ben Matheson, and Natalie Ashton. Participants at IRSE in Caerleon in April of this year provided a morale-boost as well as nice feedback on parts of what has become Chapter 6. Kirk Surgener has pulled out all the stops in proofreading the whole Thesis over the past two weeks, and has provided me with helpful criticism thereon, as well as similarly constructive barbs throughout the project. Finally, Sarah Louise-Johnson has done a wonderful job of putting up with having two(!) Thesis-finishers in her house for the past month, and has been a source of huge friendship and support in all of my time at Birmingham.

Financial support for this research came from several sources at the University, for which I am very grateful.

I'd like to give thanks to my sisters, their partners, and Max (and, independently, to my Grandmother and uncle) for providing me with a refuge from this Thesis from time to time. My parents have also been a source of support, not least by refraining to ask if I'm finished yet.

If this Thesis had a dedication, it would be to Ruth; I wish she was still around to greet it with the loving incomprehension it would warrant.

# Table of Contents

Introduction	1
Chapter 1: Capability-Theoretic Justice in Context	4
• §1 Historical Context	4
• §2 Resourcism and Welfarism: Two Existing Approaches	5
• §3 Problems with Resourcism and Welfarism	10
• §4 Resources, Functionings, and Capabilities	14
Chapter 2: Nussbaum's Capability Approach: Problems and Lacunae	18
• §1 Justice and Flourishing	19
• §1.1 Justice and Sufficient Good	19
• §1.2 'Why Capability?'	23
• §1.2.1 'Why Capability?' in Nussbaum	25
• §1.2.2 Justifying the Capability Account	32
• §1.3 The Bounds of Justice	38
• §1.3.1 The Inner Bound: The Role of the State	40
• §1.3.2 The Inner Bound: The Personal as Political	44
• §1.3.3 Liberalism?: Individualism and the Privileging of Choice	48
• §2 Constructivism or Realism?	55
• §2.1 Introduction	55
• §2.2 Nussbaum's Methodology	60
• §2.2.1 The Early Approach: Human Nature and Evaluation	60
• §2.2.2 The Late Approach: Informed Desire and Substantive Good	65

• §2.3 Sketch of the Problem	70
• §2.3.1 Unclarities, Inconsistencies, and Lacunae	70
• §2.3.2 Plural and Local Specification: Bounded Constructivism? Or an Acknowledgement of Epistemic Limits?	71
• §2.3.3 What Explains the Convergence?	74
• §2.4 Jaggar's Critique	76
• §2.4.1 Jaggar's Epistemic Critique	76
• §2.4.2 Jaggar's Normative Critique	78
• §2.5 What is Required in Response?	80
Chapter 3: Justification, Objectivity, and the Human Good	83
• §1 Adaption: Realism or Constructivism?	85
• §1.1 Adaptive Preferences	85
• §1.2 Baber's Transworld Preferentism	91
• §2 Arguing About Constructivism	98
• §2.1 Privileging one's own Perspective – Authority	100
• §2.1.1 Paternalism and Alienation	100
• §2.1.2 Methodological Exclusion; Authority and Expertise	102
• §2.2 Privileging one's own Perspective – Bias	105
• §2.2.1 Individualism and Epistemic Path-Dependence	106
• §2.2.2 Marginalised Voices	108
• §2.3 Constructivism as Self-Effacing	111
• §2.4 Making Sense of Epistemological 'Naturalism'	116
• §2.5 Rawls' 'Political Constructivism'	123
• §3 The Locus of Epistemic Values; Contingent Virtue	141
Chapter 4: Shapeless Virtue	147
• §1 Secondary-Quality Realism	150

• §1.1 The General Picture: Realism, Idealism, Quietism	150
• §1.2 The General Picture: Rules and Rails	154
• §1.3 Worldly Facts; Human Responses	156
• §1.4 Sui-Generity and Circularity	162
• §1.5 Constructivism Redux?	167
• §2 Virtue as a Shapeless Unity	174
• §2.1 Particularism, Cognition, and Affect	175
• §2.2 Reasons Internalism as a Form of Proceduralism	183
• §2.2.1 Excursus: On the Barriers to Virtue	191
• §2.2.2 Defending the Divide: Alienation and Blame	194
• §Anthropocentricity and Universalism	201
 Chapter 5: A Circle: Humanity, Well-Being, Justice, Virtue	 204
• §1 Introducing Neo-Aristotelian Naturalism	205
• §1.1 Naturalism in Nussbaum	205
• §1.2 Sketching a Neo-Aristotelian Naturalist Capability Theory	207
• §2 Reductionism?	212
• §2.1 The Core View	212
• §2.1.1 Life-Form and Normativity	212
• §2.1.2 Foot's Gloss, and Hursthouse's	216
• §2.2 Anti-Reductionist Critique	219
• §2.3 Difficulties of Interpretation	222
• §2.3.1 Foot	223
• §2.3.2 Hursthouse	224
• §2.3.3 Thompson	228
• §3 An Alternative View	230
• §3.1 The Human Life-Form and the Virtuous Life	231
• §3.1.1 Life-Form and the Good Life	231
• §3.1.2 Cultural and non-Cultural Formation	236
• §3.2 The Right and the Good	239

Chapter 6: Humanity as the <i>Polis</i>	245
• §1 Justice as a Property of Individuals and Relations	247
• §1.1 What is Essential to an Aristotelian Approach?	247
• §1.2 Thompson's Dikaiological Matrix	251
• §1.3 Against Vogler's 'Promulgation Problem'	255
• §2 Personal and Structural Justice	259
• §2.1 Structural Irreducibility, and the Difference it Makes	261
• §2.2 Structural Injustice as a Barrier to Virtue	270
• §2.3 Two Alternatives: A Conditional Interpretation, and Exculpation-Maximalism	274
• §3 What are the Practical Consequences?	277
• §3.1 Regulating Regret	280
• §3.2 Motivating Imaginative Engagement	283
• §4 Capability-Theoretic Justice Revisited	286
• §4.1 The Wider Context: Cosmopolitan Universalism	287
• §4.2 Drawing the Threads Together	289
Appendix A: Nussbaum's List of Essential Capabilities	292
Appendix B: Neo-Aristotelian <i>Naturalism</i> ?	294
• §1 What Are the Meta-Level Claims?	295
• §2 Living Up To the Label	297
• §3 Anti-Reductionism vs. Pre-Theoretic Ubiquity	301
• §4 On Second Naturalism	303
• §5 Quietism and the Point of Metaethics Redux	308
Bibliography	310

# Introduction

One recently central topic of discussion in political philosophy has been concerned with justice at the global level, the level of all humanity; one increasingly popular – and now quite mainstream – approach to this topic has emphasised the concept of human capability. There are many questions that might be asked in a philosophy Thesis, against this backdrop: about the prescriptions the approach would generate when applied to real-world controversies; about its precise relationship to the social sciences; about precisely which capabilities are of normative significance, and for which purposes<sup>1</sup>. This Thesis, however, seeks to re-examine the concepts and proposals that are often in the background when academics or practitioners talk about the capabilities and their relevance to global justice, and so aims to set out a renewed framework within which more fine-grained questions might be pursued. As a result, while I shall occasionally touch directly on clearly first-order questions – as when I discuss the appropriate definition of capability in Chapter 2 – most of my work will relate to more abstract issues in methodology and epistemology, defending a capability approach indirectly by placing it as a coherent element within a broader philosophical context.

The overarching project of this Thesis aims to set out a normative approach to global justice, similar though not identical to that of Martha Nussbaum, which centres on a concept of 'capability'. Within this, however, a surprisingly diverse array of philosophical topics emerge as relevant, and much of the Thesis will be devoted to clarifying and arguing for responses to a number of these. The areas involved are vast, and so my coverage will often be limited,

---

<sup>1</sup> Robeyns 2006 provides a critical survey of practice-oriented work up to that time; since that time, work in different disciplines on capability-related issues has burgeoned further.

although I try to give a sense of how the different areas of thought fit together. Most importantly, almost nowhere will I present full arguments against the major competitors to the responses I endorse – the disputes I do discuss explicitly are largely<sup>2</sup> internecine. This general broadness of scope also requires me to give no more than sketches of answers to certain important and relevant questions<sup>3</sup>. Despite these obstacles, I make progress towards the broad goal above, in respect of two subsidiary aims. Firstly, to clarify the proper form of capability-theoretic justice, illuminating the role that the concept of capability plays, and motivating the resulting approach. Progress is made on this in Chapters 1, 2, 5, and 6. Secondly, to place the universalism and objectivism of the capability approach in a theoretical context, by linking them to wider philosophical questions. In turn this has two main parts: the connection of the subjects of capability-theoretic justice to a conception of *humanity* (Chapters 5 and 6), and the connection of objectivity of ethical judgement with metaethical notions of *truth*, *virtue*, and *reality* (Chapters 3 and 4). I hope, in the conclusions to these lines of inquiry, to bring out the deep congruence that I perceive between the capability approach and a broader sort of philosophical perspective, exemplified by John McDowell and other recent neo-Aristotelians. While this Aristotelian connection was a major emphasis of early work in the capability approach<sup>4</sup>, it has since been significantly eclipsed by a move towards more self-consciously pragmatic philosophical stances such as Rawlsian 'public reason', Deweyan ethical pragmatism<sup>5</sup>, or a conception of impartiality derived from Adam Smith<sup>6</sup>. Although I cannot argue against these developments here, describing the Aristotelian approach as a coherent

---

2 There are a few exceptions. For example, Baber's preferentist interpretation of adaptive preference, argued against in Ch.3, is not a remotely close relative of my own.

3 The most important example of this is my inability to make any progress towards a concrete answer to questions about the actual extent of the reasons of capability-theoretic justice that any contemporary individual has; I settle for the conclusion that everyone has *some* such reasons. See Ch.6, esp. §1.3.

4 See the Nussbaum references in Ch.2, §2.2; cf. Sen 1993.

5 Anderson (2014) provides the central statement of Deweyan methodology, and she applies this method throughout her work on the capability approach, e.g. Anderson 2003, and 2010.

6 Sen 2009 draws heavily on these ideas throughout.

whole will not only help to remove the motivation for them, but will also make both the first-order capability-centred account of justice, and the broader philosophical viewpoint in which I embed it, more attractive in their own right.

# Chapter 1:

## Capability-Theoretic Justice in Context

In this Chapter, I shall begin the overall project of the Thesis by introducing the contemporary family of normative views known as the capability approach, sketching both its distinctive theoretical contributions and their motivation. I shall begin by saying a little about the historical context within which the capability approach arose, and the issues with which it has been concerned. Next I shall introduce two dominant perspectives in contrast to which capability theorists have defined themselves, and the main criticisms that motivate the turn to a capability-centric account. This will be summed up in a set of desiderata for normative theory in the area under consideration. Finally, I'll briefly outline the central concepts of capability theory itself, and show how a theory using such concepts might begin to meet the desiderata.

### §1 Historical context

The capability approach emerged as a distinctive research program and group of related theoretical contributions in the late 1980's, primarily as a response to issues in development economics. In 1990 the UN Development Program (UNDP) released the first Human Development Report, containing the first instantiation of the Human Development Index. This index, the report in which it was embedded, and the new approach to global development that it represented, were inspired by work done at the World Institute for Development Economics Research (WIDER). Amartya Sen, who had been one of the central

researchers there, was also an architect of the index and report. The coalescence of the capability approach itself can be traced to a conference at WIDER, sponsored by the UNDP, in 1988. At this conference, and in exchanges and individual research that followed, a general dissatisfaction with the way that global development had been conceptualised and promoted was revealed, and a fruitful direction in which it might turn was identified. Complaints raised by this movement centred around the fact that development had generally been conceived of as if GDP growth alone mattered; and that existing theoretical alternatives did not seem to be adequate either<sup>1</sup>.

Despite this initial focus on the assessment of global development, the capability approach rapidly expanded to encompass responses to a wide range of theoretical issues. As well as the construction of indices to measure differential development in the global context, important subjects include political distribution and social justice; the kinds of freedom and equality that justice may require; and the proper account of human well-being. The work of Nussbaum, who has been a major representative of the capability approach since its origin, has been largely concerned with these latter, broadly philosophical questions. For my purposes here, the relevant question to which capability approaches provide an answer may be summed up as: 'What are the goods with whose provision and distribution social justice is concerned?' Because societies are, *prima facie* at least, collections of individual circumstances and relations, the assessment of individual well-being and the justice of individual circumstances have an obvious relevance to this question.

## §2 Resourcism and Welfarism: Two Existing Approaches

---

<sup>1</sup> See Nussbaum and Sen 1993 (1-6).

As I have said, capability theory as I am concerned with it is centrally concerned with two related topics: the assessment of the quality of human lives, and the standard of social justice, especially with reference to the kinds of freedom and equality that it may require. However, doing more than pointing out topics of interest for capability theory requires more work. Contemporary normative theory is extremely diverse, exhibiting a trend towards the complication and interpenetration of traditional normative perspectives such as Kantian deontology, derivatives of utilitarianism, and the revived, largely Aristotelian<sup>2</sup> programme of virtue ethics<sup>3</sup>. Moreover, this process is compounded, in the region of normative thought in question here, because such theorising typically incorporates not only traditional moral or ethical perspectives such as these, but also theses that have their origin in explicitly political or otherwise social thought. Indeed, it is typical of some prominent theories in this area (most notably the 'political liberalism' of John Rawls' later work) to argue that the move to reflection on life in contemporary (at least European and U.S.) society involves a categorical shift in the kind of theory required. The great complexity of each family of theories, and the degree to which individual theorists increasingly draw on insights from many supposedly distinct families at once, makes the task of providing an adequate analysis of the connected set of theories known as the 'capability approach' very difficult.

Given this, it is perhaps most helpful, in finding a starting point for an analysis of any given capability theory, to begin with what capability theories are self-consciously *not*. In this

---

2 The primary non-Aristotelian versions are the Nietzschean theory of Christine Swanton (2003), and the revived Humean sentimentalist theory of Michael Slote (2010). Throughout the Thesis, my focus is on ideas from the Aristotelian tradition; it is debatable whether the commonalities which warrant the label 'virtue ethic' would make these less traditional theories similarly well-coherent with a capability approach.

3 Perhaps the most extreme illustration of this tendency is Derek Parfit's 2011(a and b) magnum opus, which argues that three modern ethical traditions (rule consequentialism, Kantian deontology, and Scanlon-style contractualism) can be synthesised. Compare also Driver (2004, esp. Ch.4) on the amalgamation of virtue ethics and consequentialism, and Korsgaard (e.g. 1996) on a similar *rapprochement* between virtue ethics and Kantian thinking.

regard, the 'capability approach' arose as an alternative to Rawls' theory of primary goods as a metric for assessing the justice of societies. Rawls' view<sup>4</sup> involves the identification of a range of goods the possession of which is necessary for any individual to attain their good. Importantly, 'good' here refers not to a specific conception of the good life that might be held by the individual in question, or one that might be predominately held in the particular societal context in which they live, but rather a broad space of possibilities, representative (at least for the later Rawls) of the range of 'reasonable' comprehensive images of the good that are candidate conceptions at this time. This lack of awareness of a specific conception of the good life, such as will typically (if only vaguely and implicitly) be present in the worldview of each real-world agent, reflects the process of abstraction from particular, contingent circumstances that is Rawls' 'Original Position'. Rawls holds that the correct normative principles to govern life in the *polis* can be discovered and justified through a method of reflection on what abstract representatives of the individuals in a society would recommend, subject to a 'veil of ignorance'. This 'veil of ignorance' is supposed to avoid self-seeking choices by the representatives, by preventing them from knowing too much about the specific circumstances in which they would find themselves as individuals. One example of this comes in Rawls' reasoning that if people know what social position they will have, e.g. if they know that they will not be impoverished, they are likely to alter their proposed principles accordingly. Likewise, Rawls' narrative continues, if people know exactly which comprehensive normative worldview they will espouse, the resulting principles will reflect that perspective in an unjust way: it will no longer be neutral between what Rawls calls 'reasonable' viewpoints in the way that liberal-democratic principles such as tolerance,

---

4 My presentation is of the later Rawls' theory, as stated especially in *Political Liberalism* (2005). Although many of the criticisms that capability theorists level at Rawls apply equally to his earlier work as to the later (because what changes is largely a way of interpreting and justifying the basic normative framework, rather than the concepts themselves), it is his later work that provides a clear rationale for the restriction to general-purpose resources, and so is best juxtaposed against capability theory.

diversity, and respect for others require. In this way, the 'veil of ignorance' is supposed to model certain central intuitions about how people should be treated in society, i.e. as free and equal citizens in a system of fair social cooperation, under terms acceptable to all (to appropriate typical Rawlsian language). Rawls thinks that the core intuitions that the Original Position is supposed to embody are subject to an 'overlapping consensus', at least within contemporary liberal democracies, and at least provided that the citizens concerned have reflected adequately on their values, commitments, and so on (Rawls 2005, 3-46). Although the general tenor of Rawls' approach will be important in the following chapters, the essential thing to note for my current purposes is that Rawls' view is 'Resourcist': it only requires that each person be provided with a certain quantity of a number of highly generic, multi-purpose resources, such as an adequate education, opportunities for employment, and basic political liberties.

The second major group of relevant positions is one against which advocates of Resourcism and of the capabilities are united. This group consists of theories that are more or less straightforward applications to political questions of utilitarian or otherwise consequentialist approaches. Consequentialisms are distinguished from non-consequentialist theories by their possession of a few key features. Most importantly, consequentialisms aggregate goods. This can find expression in a variety of different ways. When assessing the rightness of an act based on its impacts on a number of people, a consequentialist may aggregate interpersonally, such that an act that makes most affected individuals considerably worse off may be correct even where an alternative act would make everybody better off (but would result in a lower total benefit). Similarly, a consequentialist may aggregate across qualitatively different goods, such that an act that affects one person's life negatively in many separate ways may be right

as long as it improves even one aspect of their life to a sufficient extent. Traditionally, consequentialisms have been 'Welfarist', in contrast to Resourcist or capability-theoretic: they have taken the view that the goods that are relevant to evaluation of individual well-being or societal health are states of individual welfare, such as happiness, the satisfaction of preferences, or gains in economic utility. Consequentialisms that are not Welfarist, or (potentially) Welfarisms that are not consequentialist, are not under consideration at this point.

It may be helpful to illustrate these positions with an example (which I shall continue to use throughout this chapter). Consider the case of the nutrition of a pregnant woman. For a Resourcist, what is necessary for this woman's situation to be sufficiently good in this respect is for her to have an amount of food sufficient for the needs of a typical individual. Due to the Rawlsian requirement that normative theories not presuppose comprehensive conceptions of the good, Resourcists will be restricted in the depth of detail they can incorporate in assessing situations such as these.<sup>5</sup> It might be that it would be problematically comprehensive to incorporate standards for who counts as a man or a woman, which would mean that the 'typical individual' here must be gender non-specific; or it might be that we should not privilege pregnancy by allocating extra resources to those who engage in it. Even if, as is likely, Resourcists do not wish to be *this* restrictive, it is clear that some restriction is implicit in the anti-comprehensive stance: that some features of individuals must be invisible to justice-related action that might otherwise be considered salient for evaluation of given cases.

For a Welfarist, what is necessary for the pregnant woman's situation with respect to nutrition

---

5 Insofar as Resourcists other than Rawls do not agree with his anti-comprehensive stance on uses of conceptions of the good, they will likely be able to incorporate much greater contextual specificity in their theories. To that extent, much of the criticism I introduce below may not apply to them. I am here taking Rawls as the paradigmatic Resourcist, as has often been standard in motivating the capability approach, which may do some injustice to Resourcism as a whole.

to be adequate will depend on the particular conception of 'welfare' involved. It might be that the state of *being nourished* is intrinsically welfare-enhancing, or it might be that it is so only instrumentally, to the extent that an individual prefers to be nourished, or is made happy thereby. All Welfarisms, however, are alike in that all they can use in evaluating the woman's situation are *actual results*; the etiology of a result is irrelevant. Finally, there will also be differences between Resourcism and Welfarism when we extrapolate out from the woman's situation to that of an entire polity. Here, where we may be concerned with the goodness of the polity *simpliciter*, or with its goodness specifically in terms of social justice, Resourcists like Rawls require that each individual have the relevant bundle of resources, while Welfarists may be more willing to trade off the welfare of individuals in order to ensure a greater good. While the Resourcist will be concerned with whether each individual has enough (relative to the generic individual) to eat, the Welfarist might be more concerned with whether the total amount of nourishment that occurs is as large as it might be, or with other variables such as GDP, or total levels of happiness.

### §3 Problems with Resourcism and Welfarism

Resourcism and Welfarism, as just described, form the theoretical background against which capability theory emerged. In this section, I shall bring out a number of criticisms that have been levelled at Resourcism and Welfarism by capability theorists, and which motivate the move towards a focus on capability.

To begin with, Resourcism has been criticised for being far too coarse-grained to really inform us about the justice-relevant goods at stake in individual situations and polities. By restricting attention to what *typical* individuals would need, they will fail to ensure that

*actual* people achieve, or have a realistic opportunity to achieve, states of value in their own lives. The argument here can take a number of forms. Firstly, Resourcist views do not allow for tailoring of the quantities of resources allocated to individuals to reflect the differential capacities of those individuals to convert resources into goods. A person with a disability, for example, will likely require more resources of many kinds to pursue the same goods as someone without; a pregnant woman requires more and different kinds of food to achieve the same level of nutrition as a non-pregnant woman (Nussbaum 1992, 233)<sup>6</sup>. Secondly, there may be external factors in the individual's context that affect their capacity to obtain value from a given bundle of resources. It may be that the pregnant woman has enough rice for her particular body to be nourished, but that cultural taboos prevent her from gaining access to a kitchen to cook it, forcing her to give it to a husband who will take some portion of what she needs for himself (233-4).<sup>7</sup> Finally, capability proponents note that it seems strange to include resources as goods at all when even Resourcists rarely consider many of them to have value in themselves. Income, wealth, employment (of a generic kind), even insurance contracts<sup>8</sup> are plausible candidate resources, but it is clear that each has little or no intrinsic value. It would be strange to conclude that purportedly value-laden concepts like freedom or equality are constituted solely by conditions that themselves involve no value. If an individual is not

---

6 This line of anti-resourcist argument remains at the forefront of these disputes, even while some of the others have arguably dwindled in significance, as resourcist positions have become clearer and more sophisticated. Given this increasing compatibility between capability approaches and resource-focused ones, some of the arguments here may not motivate capability approaches as powerfully as may first appear. Nevertheless, these criticisms are still important to note, even if only to motivate the desiderata identified below. The papers in Brighouse and Robeyns (eds.) 2010 provide a good introduction to the contemporary state of the resource/capability debate; see especially the papers by Pogge, Anderson, Arneson, Terzi, and Robeyns therein.

7 Putting it this way may suggest that I think that e.g. disability status is a fully intrinsic matter of the individual constitution, rather than something that arises because of the social signification given to attributes which are themselves well-being neutral, or because of the way that such intrinsic features *interact* with social structure. (Cf. Calder 2008; 2011). I do not need to take a position on these issues here, because I do not need to put argumentative weight on the distinction between 'internal' and 'external' factors, although I am inclined to think that a Calder-style interactionist ontology will be plausible for at least most states falling under the ordinary-language term 'disability'.

8 See Dworkin 2000.

inherently brought closer to the realisation of value by being made more free, or more equal with others, why is it that freedom or equality are of value themselves? (Sen 1993, 32-48; Nussbaum 1987, 9-10)<sup>9</sup>. Taken as a whole, the problem with Resourcism is that it isn't concerned with whether people can actually achieve goods with the resources they are allocated, not least because it (frequently, at least) is not oriented towards the goods themselves at all.

Welfarism, on the other hand, *is* directly concerned with the goods – or, at least, the alleged goods – that individuals actually realise. The central problem that capability theorists identify with it is that it ignores the importance of individual agency and other factors having to do with the etiology of states of welfare. To a large extent the arguments here overlap with the arguments for the importance of capability rather than functioning, which I shall address in the next Chapter (§1.2), so I shall leave it until then to draw those out. A *pro tanto* case, however, may be made by noting that what seems to be important for social justice is not only what people in fact do, but what they *could have done* or *can do*: opportunities to act are important as well as, if not rather than, exercises of agency or agency-independent outcomes. As Welfarisms are interested only in the extrinsic upshots of policies, institutions, and choices, and not in the constitutive processes that they involve, they cannot pay attention to the opportunities people have if those opportunities never bear fruit. If someone decides to fast, or to eat an unhealthy diet, this seems very different to a case of unchosen starvation; if the person *could have* been nourished had they chosen otherwise, this (at least) lessens the

---

9 It may seem, in the light of points that I emphasise later (Ch.4, esp. §§1.4-5), that endorsing this line of critique is hypocritical on my part, since it may seem to involve a demand that freedom or equality been shown to be valuable from without, whereas all explanations of a thing's value must reach explanatory bedrock at some point. However, any inconsistency is superficial: Sen and Nussbaum's point can be preserved as an insistence that freedom and equality must be located within a broader conception of ethics and the human good it accounts for, rather than appearing out of thin air; this Neurathian version of the claim is *highly* congruent with what I say later (cf. especially Chapter 5, §1.1)

badness and/or injustice of their situation (Sen 1993, 40). Welfarism also has problems at the societal level, in that it gives no inherent weight to considerations of distribution, abstracting away from the issue of whether each person is dealt with well. The problem with this cross-personal (and cross-good) aggregation may be expressed in a number of ways; as a failure to treat each person as an end, a failure to recognise the separateness of persons, or a blindness to the irreducible plurality of the good, and to the necessity of all its separate elements for a valuable life (Nussbaum 1992, 222; 2000b, 106)<sup>10</sup>. Given that social justice is centrally (if vaguely) concerned with people being treated as equals, this makes Welfarism ill-equipped to account for the justice of polities, which is (at least) a central element of their goodness. An injustice is present if the pregnant woman does not have enough to eat, even if others in her society get more (in total) as a result; likewise, something is deeply lacking in a life of (at least unchosen) malnutrition, even if that life is thereby made better overall, by improving other elements.

These arguments amount to a motivation for a number of desiderata for a theory of social justice, and of the goodness of societies. It will then be in successfully meeting these desiderata, without introducing further implausibilities, that capability theories will be successful over e.g. Resourcist and Welfarist rivals. Firstly, theories should accommodate the value of choice, and more generally be sensitive to the significance of etiology: there are many contexts in which it makes an evaluative difference what, how, and whether a person has chosen, as well as (potentially) other etiological factors. Secondly, theories should pay very close attention to whether each individual person is actually achieving states of value; it is not enough to provide for the needs of generic individuals, we need to provide for those of

---

<sup>10</sup> Sen arguably endorses interpersonal and cross-good aggregation (cf. 1982, 1983, 1993; all can be seen as advocating a more complex, but still fundamentally aggregative, consequentialism); at any rate, he does not seem to share Nussbaum's principled hostility to such.

*actual* individuals, and this is likely to require expansive examination of each particular circumstance. Thirdly, theories should seek to include as goods things that are plausibly of value in themselves, as opposed to things that are mere means; the most straightforward way to do this is to set up the goods that justice is concerned with as elements in human well-being. Finally, there is something deeply inadequate about theories which allow aggregation across persons or across goods to deprive any individual of essential elements of their good; doing so ought to be judged unjust.

## §4 Resources, Capabilities and Functionings

Capability theories build on the above criticisms, attempting to provide an account of what is valuable for justice-relevant purposes that avoids the flaws of both Welfarist and Resourcist approaches. Rather than focussing on the states of value that individuals actually realise, or on the quantity of general-purpose resources that individuals have available, capability theorists focus on the degree to which a given person is able to immediately realise relevant states of value. Firstly, some terminology: from this point, I shall frequently talk of 'functionings'. A 'functioning' is any way of being that an individual human may instantiate<sup>11</sup>, including relationships that they may have to external objects or individuals, or any activity that they perform. As David Crocker puts it, functionings are 'ways of doing and being' (1992, 590). Thus, *being well educated* is a functioning, as is *being respected by a friend*, as is *playing football*. To return to my earlier example, 'being nourished' would be a functioning that (roughly) consists in the body's having sufficient resources and the capacity to convert

---

<sup>11</sup> As I note below, there probably are many such functionings that an individual can only have in collaboration with others – or, more radically, if others realise them too. Thus, the only prospect that this view excludes would be factors contributing to an individual's good without that individual being in any way involved; it does not exclude the possibility that there might be functionings that can only be realised by individuals as parts of a group.

them into the chemicals it needs to perform all of its other functions adequately. For the normative purposes at hand in this Thesis, I am only interested in those functionings that are valuable for their bearer.

It is now possible to bring out the concept of 'capability' itself. My definition here is preliminary; the next Chapter will introduce some important complications. Capabilities are capacities, on the part of individuals, to function in some valuable way, such that the only condition that prevents the individual's actual functioning is the lack of a choice of that individual to do so. Put more technically, a capability is a disposition, of an individual being, to instantiate a functioning just in case they choose to; if an individual has the capability  $C_f$ , they will instantiate functioning  $f$  iff they choose to do so<sup>12</sup>. Coming to be in this state may involve many separate conditions. Firstly, an individual's constitution must be such that they can function in the given way. So, in a situation where a person will have to prepare food themselves, that person cannot be nourished by rice if they do not have the cooking skills to transform it into a digestible form. Such skills or technical knowledge, and physical characteristics of the body itself, such as physical strength, mobility, perceptual capacity, etc., are examples of potential 'internal' barriers to capability. Secondly, the individual's 'external'<sup>13</sup> situation must also be suitable for their functioning. The cultural taboo that enjoins, with a background threat of (e.g.) ostracism, the woman to give rice to her husband, who will appropriate some of it for himself, would be one such restriction. Another will be the necessary resource that is the rice itself, along with cooking equipment, energy sources, clean water, and so on. Notably, many of these resources may be collectively owned (or not owned by anyone), but that does not prevent their being resources for the individual in question as

---

12 This is a very strict definition – it is harder to possess a capability on this account than on at least most others. For some implications of this, and a limited defence of this strictness, see §1.2 in the next Chapter.

13 For the internal/external distinction, see Nussbaum 1987, 20-5.

long as the individual has permanent unconstrained access to them; this is particularly clear in the case of clean water, since that is almost always a common resource<sup>14</sup>. Overall, then, if there are neither internal nor external barriers to functioning, and if the only thing that prevents immediate functioning is the lack of a choice to do so, then the individual has the capability to function. Notably, all of the barriers to functioning may or may not be present depending on many contingent features of individuals and the environments in which they live. For the case of nutrition alone, relevant variables in the woman's case will include her pregnancy, characteristics deriving from her biological sex, her occupation and level of activity, local disease risk factors, etc., as well as the possible barriers already mentioned. What is required for one situated individual to function may be very different to that which is required for another to do so.

Capability theories claim that these kinds of states of individuals are highly relevant to social justice, and to the assessment of the goodness of individual lives and of societies. Capabilities and functionings thus take the place of bundles of generic resources for Resourcists, and states of actual welfare for Welfarists. From what has been said so far, it is clear that capability theory stands a good chance of being able to satisfy the first two desiderata that I have identified. That is, because what is relevant here is capability and not merely functioning alone, capability approaches can be etiology-sensitive, and will, at least to some extent, be capable of accounting for the value of choice. And, because of the specificity of barriers to functioning to individual circumstance, and the wide range of features of individual lives that are recognised to make a difference here, capability theories will likely be accurate in their assessment of actual human lives. However, it still needs to be shown that the final two

---

14 Sen's classic malaria-protection example emphasises this non-individualistic feature. See Sen 1992, 66-8.

desiderata, dealing with the identification of intrinsic goods, and a non-aggregative structure respectively, are met by capability theory. I discuss these within the context of Nussbaum's view of social justice in the next Chapter.

## Chapter 2:

# Nussbaum's Capability Approach: Problems and Lacunae

In the previous Chapter, I began the task of introducing and motivating a capability approach to global justice. In this Chapter, I continue towards that end, with a new critical focus on Martha Nussbaum's version of the capability approach. Much of the Chapter will be spent explaining the central features of her view, and identifying points at which greater detail is needed, or where there are tensions which require resolution. Although the bulk of my exposition here will be critical of Nussbaum's work, it should be clear that the reason for this critical focus is that I ultimately believe her contributions to be highly important and, in large part, correct. The Chapter is split into two sections, each devoted to a major area of theoretical controversy within Nussbaum's work, each of which in turn provides an impetus for several strands of the broader philosophical discussion that comprises the rest of the Thesis. The first section is devoted directly to questions about global justice and its relationship to human flourishing. I begin by bringing out the compelling connections that Nussbaum's approach incorporates between justice and the individual human good. Next, I concentrate specifically on the question of why *capability* is of importance in its own right, and argue for some minor revisions to Nussbaum's formulation as a result. Finally, I briefly discuss the universalist and (equivocally) liberal character of Nussbaum's approach to justice, setting the stage for more detailed development on my own behalf later in the Thesis. The longer second section turns to consider the ethical epistemology which Nussbaum employs,

exposing some nascent meta-level issues which will be especially important in what follows. Firstly, I introduce a background distinction between 'constructivism' and 'realism', and distinguish two aspects of the debate concerning these which are relevant to the justification of a capability approach. Then I identify Nussbaum's most significant proposals about methodology, making some critical points in the process. Finally, I describe an important criticism of Nussbaum by Alison Jaggar, and briefly evaluate its cogency. The problems I identify as arising from this critique will then form the motive force for Chapter 3.

## §1. Justice and Flourishing

### §1.1 Justice and Sufficient Good

I have provided – in the previous Chapter – a basic account of the background motivations for the capability approach, and of its central concepts. As each capability theorist uses these differently, however, it is necessary to isolate individual thinkers to get a clear picture of how the capabilities relate to social justice. Because Nussbaum's work will be most important in my project as a whole, I shall here provide an outline of the way that she does this. I shall begin by sketching the structure of her account. Then I shall highlight two methodological considerations that are at work therein. Overall, I shall conclude that an account which preserves key features of Nussbaum's view has good prospects to meet the final desideratum I identified in the last Chapter: incorporating an element which exposes the ethical limits of cross-personal and cross-good aggregation.

Central to Nussbaum's approach is the idea that each being who is a subject of justice must have a certain set of capabilities, if the world is not to be unjust. This set will include some degree of each valuable capability. This degree constitutes a threshold below which no-one

should fall. An individual situation, or an entire polity, will not be just (and so cannot be good) if any person in it is below the threshold on any capability (2000a, 81-6). This provides a clear framework that relates a view of the minimum capability necessary for an adequate life (the threshold), to a view of social justice at the global level (*inter alia*).

Importantly, justice here is an absolute standard. Although a case might be more or less just, there is some absolute point at which it will tip over from minimal justice to injustice. This means that the approach maintains a non-aggregative structure in both interpersonal and cross-good respects. For Nussbaum, a policy (or individual action) that brings or leaves anyone below the threshold for any kind of good will be unjust, no matter how many others it might bring or leave above it<sup>1</sup>. This is not to say that we might not be forced to engage in such choices, but if we must do so our choice will be tragic, and some irreducible wronging will remain, no matter how good the aggregated results (2000a, 81). Nussbaum's view, then, has a strong chance to accommodate all four desiderata, and so seems to have clear advantages against the Welfarist and Resourcist views that I began with.

Nussbaum provides two ways of thinking about the human good, and so derivatively of the capability threshold that is relevant to social justice. The earlier method involves asking which features of our lives function as criteria for humanity itself; what in life is so essential that a being that lacked it could no longer be fully regarded as a human *at all* (1990a, 218-9; 1992, 214-16; 1993, 260-7; cf. Also 2006a, 181-2). The later method instead asks which features of our lives are necessary for human *dignity* (2000a, 72-4). It is not entirely clear

---

<sup>1</sup> Notably, many political questions involve aggregations that are plausibly 'above the threshold' – involving balancing and weighing considerations, but not in a way which places anyone's sufficient good in the balance. Nussbaum's view can wholeheartedly endorse such aggregations. (Cf. 2000a, 211-12, on inter-good balancing in the context of 'religious freedom' jurisprudence).

what 'dignity' means here; Nussbaum doesn't seem to think that it introduces any substantial content of itself<sup>2</sup>. It seems fair to assume, then, that a life with dignity is a sufficiently good human life, in some sense, although it is clear from the demandingness of the standard of dignity that Nussbaum sets that the absence of dignity doesn't necessarily make a life so bad as to not be worth living. On the other hand, it is not obvious that the sorts of functioning which Nussbaum identifies as constituents of dignity are sufficient for a person to live a 'good life' in the ordinary sense of that phrase<sup>3</sup>. Nor is it intuitive that an injustice is *necessarily* done to someone if they do not flourish, as such: perhaps it is their fault that they do not flourish; perhaps their flourishing is something that humanity could not ever ensure, even if socially-contingent circumstances were very different. So, the standard of sufficiency involved in dignity is likely to be a quite specific one: not identical with, but lying in between, the standard of the life worth living and that of the good life (*simpliciter*)<sup>4</sup>. What is more, both the humanity-theoretic and the human-dignity-theoretic approaches explicitly involve reflection on our common species, even if that reflection must be overtly related to the concept of well-being. This has the merit of enabling a clear explanation of why the class of beings for whom justice is relevant is co-extensive with the class of humans<sup>5</sup> – if humanity is implicated by the notion of dignity itself, it will make no sense to regard justice as having any other scope<sup>6</sup>. However, I need not consider the possible differences between these two

---

2 For example, there is no reason to think that dignity in this sense involves conventionally 'dignified' conduct.

3 See the list given in Appendix A.

4 Vallentyne 2005 (361-3) argues that approaches which use a notion such as necessity for human flourishing to restrict the space of justice-relevant functionings (or capabilities) are inadequate because in principle any way of functioning (whether good or bad for human beings in itself) can make a difference to whether or not an individual has a decent life. This is not well-targeted against Nussbaum, at least as I am developing her thought, because although one could use the language of *necessity* for humanity (/human dignity) here, the thought is actually that the relevant bundle of functionings are necessary *and sufficient* for the relevant status as one who is living a good life.

5 All of Nussbaum's early papers, 1987-1995, contain some notable development on or around this view; this link to a norm of humanity will be particularly important in my project as a whole.

6 However, cf. Ch.5 fn.12 on this point.

approaches further here<sup>7</sup>.

A second methodological feature involves two respects in which judgements about the goodness of candidate capabilities are constrained. Firstly, because in holding that a given capability is essential we hold that people have a claim on it (which is presumably matched by corresponding obligations on the part of institutions and agents), we have to match our assessment of these values and their demandingness to what is realistically possible in the contemporary context. One motivation for this is that it makes little sense to think that humans require things for a dignified life that few (or no) humans could ever actually achieve. We are reflecting on what is essential for the good of the kind of being we are now, not what would be so for some other kind of thing (1995, 105-9, 120). A second is that we believe that obligations are limited to what people can reasonably be expected to do, and if valued capabilities are unrealistic or thresholds out of reach, this criterion will not be met (2006a, 280-1)<sup>8</sup>. Secondly, because *each* of the capabilities is required, and we cannot justly trade-off one against another, we need to consider what effect particular judgements about a threshold level, or about a candidate capability, will have on others. If one threshold were set too high, it would make it unrealistic that others could be met (1987, 37). There is a potentially devastating tension in Nussbaum's methodology here. On one hand, Nussbaum's notion of tragic choice concedes that sometimes it may be necessary to act in a way that is inadequate or unsatisfactory when it comes to justice, because it brings someone to, or leaves someone with a life of insufficient dignity, but is also the least of the unjust acts available. On the other, we are given reason to weaken our normative claims whenever they outstrip the

---

7 I use recent neo-Aristotelian work to support a somewhat similar methodology for thinking about dignity in Chapter 5.

8 The interpretation I eventually develop in Ch.6 may appear to violate this criterion; in fact it does not do so, because the norms that extend beyond the current capacity of individuals to live by them do not generate deontic burdens: people's 'obligations' still do not outstrip what can reasonably be expected of them.

possibilities that actual, contemporary situations afford. This threatens either to re-introduce aggregative reasoning – because the considerations which justify the 'tragic' option will be aggregative – thereby violating a key desideratum, or to introduce a distance between our evaluations of the justice of events and our judgements about what people ought to do, justice-wise. The difficulties thrown up by this tension will be a principal topic of Chapter 6, where I argue for a resolution by locating my Nussbaum-style sufficientarian account of global justice as an ideal theory, aimed at guiding our practical imaginations and our general evaluative responses to events, rather than our moment-to-moment decisions.

These complications aside, engaging in one or the other of the two methods above will result in a list of capabilities that are valuable, and a judgement about just how much of each capability is necessary in each life for a person to be suitably enabled to flourish.<sup>9</sup> These lists, in congruence with the fact that they originate in reflection about human life itself, are supposed to be universal, applying to all human<sup>10</sup> lives in the contemporary world, across boundaries of nationality, ethnicity, cultural identity, gender, and so on. This raises important meta-level questions about justification, which I begin to address later in this Chapter (§2.2 onwards). The compelling core of Nussbaum's view remains, however. As she argues, justice requires<sup>11</sup> that every human being have that bundle of capabilities and functionings, suitably distributed over their lives, that constitutes a life of dignity.

## §1.2 Why Capability?

---

<sup>9</sup> See Appendix A for an example list, from Nussbaum 2006a, 76-8.

<sup>10</sup> Precisely what is involved in these kinds of references to 'humanity' is a central topic of Ch.5.

<sup>11</sup> Put this way, it will appear that I am proposing a supposedly complete account of social justice. However, this is not a necessary consequence (and I shall suggest later – in Ch.6 – that it is not very plausible). The claim, rather, is really just that this is *part* of what justice requires. Nussbaum usually inserts similar caveats (e.g. at 2000a, 75; 2006a, 274).

As the preceding section has exemplified, capability theorists maintain that facts about people's capabilities are essential to evaluation of the justice of individual situations and of polities. Thus far, however, I have not given a characterisation of exactly why it is plausible to hold that *capabilities* are so important, as it is *functionings* that might more naturally be seen as components of human well-being<sup>12</sup>. To see this, note firstly that capabilities are always capabilities for something: for some functioning. What is more, capability advocates do not hold that just *any* functioning will be of importance here. It is hardly important, *prima facie*, to enable people to efficiently count all the blades of grass in a field<sup>13</sup> – and this appears to be because that functioning would be of no value to its potential bearers. This might suggest that the value of a capability is at best derivative of the value of the functioning with reference to which it is defined. However, capability advocates (or, at least, Sen and Nussbaum) are committed to resisting this thought, and arguing instead for a more independent good-making role for capabilities. In this section, then, I turn to consider the justification for this focus on capability. I do this in two subsections. In §1.2.1, I examine the justifications that Nussbaum has considered for a focus on capability, together with the question of how exclusive this focus is within her own work. Then, in §1.2.2, I build on this discussion to sketch a model capability theory that I hold to be well-supported by some Nussbaum-style justifications, and which will inform my arguments concerning the often more abstract or higher-order topics that form the subject-matter of most of the rest of the Thesis. In this subsection I also provide a brief defence of the strictness of my definition of capability.

---

12 See e.g. Arneson 2010, for one argument that functionings are of great importance. As I have already noted, capability theorists typically integrate considerations about well-being very thoroughly into their views about justice.

13 The example is, famously, used by John Rawls (1999, 379); cf. Foot 2002, 107).

Before I move on to that, however, it is worth re-emphasising some of the points I made in the last Chapter. There, I presented four desiderata for an account of global justice, of which the first three are of importance here. Firstly, I claimed that accounts of justice need to be sensitive to choice and other agential etiological factors – this will be of obvious importance in what follows, since the very concept of a capability implicates choice, on my strong definition. Secondly, I noted that theories need to be able to tailor their evaluations of individuals to those individuals' circumstances to a high degree – this desideratum will be in play insofar as different ways of understanding the capabilities focus might affect the amount, and kind, of detailed information that is identified as salient. Thirdly, I held that theories should place value on intrinsic goods as far as possible, in order to render as clear as possible the value of justice for the individual humans that it affects.

### **§1.2.1 'Why Capability?' in Nussbaum**

In this subsection, then, I shall concentrate on Nussbaum's work, and the role of different justifications within that, in order to provide a relatively complete picture of one way of thinking about the importance of capabilities in comparison to functionings. There are two major kinds of argument for this conclusion. The first places the connection that well-being is supposed to have to politically-relevant conceptions of justice in the foreground. The second emphasises the essential role of practical reasoning and choice in human well-being itself. As well as the difference between these two kinds of argument, there is a difference in what is being argued for: sometimes it seems that Nussbaum in particular holds that *only* capability is relevant to social justice and well-being; sometimes it seems that capabilities are only *one vital ingredient* in complete conceptions of such. I will expand on this later in the subsection, but it is important to note from the outset that all that is actually *necessary* for capability

theory to meet the third desideratum is that capabilities be essential to adequate thought about the human good; it is not necessary that nothing else be too. What is more, if, as I claim, the proper role of a capability approach involves providing characterisations of both the justice-relevant capabilities and *other* justice-relevant goods (other functionings), the fact that the capabilities aren't of exclusive significance is even less threatening.

One instance of the first kind of answer to 'Why Capability?' would trade on the fact that the capability approach aims to provide an account that links well-being with values such as justice, freedom, and equality. This has potential benefits in both directions: by explaining how well-being relates to judgements about the correct form of political arrangements it should make it less mysterious how politics relates to ethics, but, equally, linking conceptions of freedom and equality to well-being, such that being made more free or equal always clearly contributes to one's good, helps to clarify how it is that those things are valuable. Consider 'freedom': 'negative' conceptions of freedom, as the absence of coercion, have frequently been considered to be the most plausible analyses of everyday concepts named 'freedom', but this common-sense plausibility is marred by the obscurity of any grounds to think that purely negative freedom could be of intrinsic value. What we need is an analysis of something that can genuinely be called 'freedom', but which also exhibits a strong, clear, and ideally intrinsic, normative pull. The approach that capability theory proposes, which preserves most importantly the common-sense judgement that freedom, as a distinctive value, involves opportunities to act rather than (or, at least, as well as) exercises of agency<sup>14</sup>, has a good chance of satisfying the former criterion<sup>15</sup>. So long as the link to well-being is plausible,

---

14 See Taylor 1991, 143-4, for a general distinction between 'opportunity' and 'exercise' concepts of liberty. Whether capability-theoretic freedom can really be regarded as exclusively an opportunity concept is debatable, but it is not important for my purposes to address this here; it is at least clear that any capability approach worthy of the name will involve opportunity-side considerations to some extent.

15 It is, of course, impossible to do justice to the disputes surrounding the analysis of concepts like freedom

it will satisfy the latter too. Turning to equality, it seems implausible that any kind of equality that is mandated by justice could be independent of the choices that agents make in their lives. One ubiquitous objection, associated most prominently with 'luck egalitarian' views, to the Thesis that justice requires us to equalise the amount of good that people experience over their lives<sup>16</sup> is that individuals should not be responsible for the irresponsible actions of others. For example, if someone who has had available a genuine opportunity for post-secondary education wastes this, there may not be any injustice involved in their not having that opportunity again at a later stage of their life, even if the lack of such education would normally be an injustice.<sup>17</sup> This kind of intuition supports opportunity-involving conceptions of equality, too: surely justice does not require us to continually allocate resources to people who, despite having sufficient opportunities in the past, have consistently chosen<sup>18</sup> to waste them, and have not thus attained equal functioning with others? From this perspective, then, the fact that a capability-centric approach takes as relevant values opportunities, rather than just choice-insensitive outcomes, counts in its favour for the purposes at hand.<sup>19</sup>

---

and equality, as they relate to the concept of capability, here. See e.g. Cohen 1993, and Sen 1993, for some important arguments in this area.

- 16 As should already be clear, the capability-theoretic view of justice does not advocate a straightforward equalisation of goods in any case, but it does involve equal attainment of a threshold level of good, which could be susceptible to similar objections.
- 17 There is room for considerable nuance here: for many capabilities, perhaps including the opportunity for rewarding higher education, I would probably hold that people *should* have this opportunity at multiple points in their lives, even if they have previously chosen not to function. But it is clear that this only complicates the picture, rather than changing its deeper structure: it remains the case that at various points in space and time no injustice results when someone chooses not to function well, whereas a failure to function with any other etiology would constitute an injustice.
- 18 This sort of claim need not (and should not!) operate with a simplistic conception of the requisite conditions for ethically-relevant (luck-excluding) choice. As I flesh out a little below, the references to 'choice' in this section – and, arguably, throughout Nussbaum's work too – should be taken as potentially including a very wide range of constraints on the sorts of choice that are genuinely of significance here. Cf. fn.30 below.
- 19 Nussbaum, at least, does not explicitly present full arguments of this form. However, it is clear that Nussbaum holds that the capabilities approach is designed to be linked to conceptions of political justice, and (at least) of equality (e.g. at 2006a, 75-6) although she does not put emphasis on it as an analysis of freedom in the way that Sen does (cf. 1993, esp. 33-5). It is also clear that, quite generally, she does not see our extant judgements about social justice or equality as inert in the generation of a normative theory that relates to them (cf. e.g. 2006a, 171-179).

A second instance of this politics-centric kind of answer, which is emphasised in Nussbaum's more recent work (1999, 40-7; 2000a, 65-105; 2000b, 129-30; 2006a), justifies the concentration on capabilities as necessary for the relevant normative theory to fall under an 'overlapping consensus', in the sense proposed by Rawls in his *Political Liberalism*. I cannot explain Nussbaum's use of this Rawlsian context in detail here. However, her reasoning is straightforward enough not to require much background information. Nussbaum gives several rationales for a capabilities focus that amount to appeals to its necessity for consensus to be possible. The first appeals to particular cases in which it is clear that people whose normative perspectives Nussbaum thinks 'reasonable' – and which must therefore be accommodated within the 'overlapping consensus' – would wish to forgo functioning themselves, but might nevertheless support bringing themselves and others to the point of choice about whether to do so or not. The principal example here<sup>20</sup> is of religious fasting: the ascetic, for example, wishes to fast, thereby depriving themselves of the functioning of nutrition; but they need not oppose themselves or others being given the effective choice of fasting or not (Nussbaum 2000a, 87; cf. 2006a, 79-80). The second, rather than appealing to particular cases, argues that the kind of general independence from particular 'comprehensive' conceptions of the good that political liberalism requires is violated if we require people to function in particular ways (Nussbaum 2000a, 87, 95-6).

The second kind of justification for the focus on the capabilities, that which aims to do work within the conception of the human good itself, also comes in two forms. I'll deal with these quite briefly now, since there is considerable overlap with the argument I endorse below. One form, which is most central in Nussbaum's early work, stresses the central, 'architectonic' role

---

<sup>20</sup> The example also occurs at Sen 1993, 40.

of individual practical reason in the good life. Nussbaum claims that practical reason – the ability to assess and select individual actions and ways of living among plausible alternatives – is such an important element in the human good that it infuses all the rest. Even functionings such as eating and drinking seem less than fully human if they occur in the total absence of reflection and deliberation (1987, 44-50; 2000a, 82, 85 and 92). The other form argues that many particular activities could not even minimally be regarded as valuable human functionings in the absence of reflection, deliberation, and effective choice. Among the examples given here are love, which is not genuinely present unless it originates in the self, rather than being imposed from outside; and play, which is not present if activity is coerced (2000a, 88).

Now that the grounds which Nussbaum gives for concentration on the capabilities have been sketched, the remainder of this section will be taken up with an assessment of their coherence. Firstly, it is worth noting that she seems to hold that some justifications of the first – political – kinds will be necessary for a defence of any rigid focus on capability rather than functioning. This is because it is clear that she believes, as an ethical matter, that some functioning of at least some kinds is necessary for a genuinely good human life. Very plausibly, capabilities themselves are not *enough* to ensure flourishing. This shows that Nussbaum does not think that one can rely exclusively on appeals to this second kind of grounding for the capabilities-focus, if that focus is to be exclusive: if the relevant conception of the human good requires even *some* actual functioning, it will not be sufficient, to ensure a minimally good life, to merely guarantee the capabilities<sup>21</sup>. The existence of the second kind

---

21 Unless, that is, the actual functionings that are required are necessary prerequisites of more complex and difficult-to-instantiate capabilities. For example, nutritional functioning throughout childhood would obviously be a necessary prerequisite for a productive and autonomous role in the adult workforce. I take it for granted, however, that not all of the functionings that are plausibly required for a flourishing life are going to be components of complex capabilities in this way.

of well-being-based argumentation is probably not enough to prevent this problem, because it is not plausible, at least *prima facie*, that very many of the individual functionings must (by definition) flow from capability in the way that play or love must. Plausibly, the sort of functionings that seem absolutely necessary for human dignity or human life itself will tend to be of a very simple sort – such as the consumption of adequate food and water, or enough of (e.g.) the social bases of self-respect<sup>22</sup> to ward off crippling mental disorders – and it is just such functionings that are least likely to display the requisite conceptual dependence on practical-rational choice.

Secondly, there are reasons to doubt that Nussbaum's explicit endorsement of the capabilities-only view (e.g. at 2000a, 87) is consistently upheld. The key example here is what she sometimes calls 'dignity', which seems to stand in for a large bundle of functionings centred on the development and maintenance of flourishing relationships with others. Firstly, Nussbaum claims that the failure to have dignity in relationships will damage other capabilities, and thus that mandating dignity may be justified on these instrumental grounds (2000a, 91-5)<sup>23</sup>. However, she also holds that coercive promotion of dignity functionings is warranted regardless of such means-end reasons:

Suppose a state were to say “We give you the option of being treated with dignity. Here is a penny. If you give it back to us, we will treat you respectfully, but...you may keep the penny, and we will humiliate you.” This would be a bizarre and unfortunate nation, hardly compatible, it seems, with basic justice. We want political principles that offer respect to all citizens, and, in this one instance, the principles should give them no choice in the matter. (2006a, 171-2).

The existence of this exception suggests that the ultimate rationale for the focus on capability is of a flexible, non-absolutist sort. Importantly, this is a flexibility that Nussbaum is much more reticent to countenance in the case of direct trade-offs *between* distinct capabilities

---

22 See Rawls 1999, 54.

23 See also 2000b, 128-32; where these points are made again in response to Arneson.

themselves (2000a, 81); there, she holds that losses to one capability cannot be straightforwardly made up by gains to another. The elasticity of the normative conception that must lie behind such reasoning seems ill-suited to representation in terms of the kind of deontological, individual autonomy-based rationale that Nussbaum appropriates from Rawls. If a substantial amount of justificatory work were being done by the general injunction against forcing people to act in ways contrary to their 'reasonable' comprehensive conceptions of the good, we would expect to see a clean, absolute prohibition on the direct promotion of functioning in cases where a 'reasonable' conception of the good is plausibly at stake. This is especially clear given that Nussbaum misrepresents the situation in the quote above by focusing on government actions; the interventions which governments should take to prevent violations of dignity in non-governmental contexts such as family life are also, for the capability theorist, to the point<sup>24</sup>. It is far less clear that the exemption for 'dignity' can be squared with the Rawlsian prohibition on promoting comprehensive conceptions when the interventions in question might involve (e.g.) shutting down a ubiquitous network of private members clubs where gendered humiliation is practised for money, etc. If Nussbaum thinks that 'reasonable' conceptions of the good are not at risk of contravention in such 'dignity' cases, we should expect to see argument to that effect.

However, an argument for a general capabilities focus which works within the conception of well-being, or which appeals to links to conceptions of freedom, justice, or equality, has much better prospects to support the necessary flexibility. Cases of choice-incompatible functioning may be justified on the grounds that the value for well-being of practical-rational planning and choice-making is reduced or absent for exceptional cases like that of dignity

---

<sup>24</sup> I stress this in elaborating on Nussbaum's general rejection of public/private distinctions below (§1.3)

(perhaps because e.g. a choice to degradingly submit to the dominance of others is destructive of practical reason itself). Alternately, perhaps the cogency of the capability-theoretic conceptions of freedom or equality will not be significantly damaged by relaxing the prohibition on the promotion of functioning in these limited cases.

Taken together, these points show that it is obscure what essential work the appeal to violations of people's moral autonomy in forcing them to function actually does. On the one hand, it does not seem plausible that considerations of well-being alone can justify an exclusive restriction to the promotion of capability; this might suggest that some Rawls-style liberal rationalisation must fill the gap. On the other, Nussbaum appears not, in fact, to think that such an *exclusive* restriction is appropriate. Ultimately, then, most of the normative work here seems to be being done by conceptions of well-being, rather than by restrictions originating in thoroughly independent political thought. Given that in early versions of her approach this was the only justification given, this also suggests that not much of substance in her account has changed with respect to this issue, despite the large superficial change that has resulted from her endorsement of Rawlsian political liberalism<sup>25</sup>.

### **§1.2.2 Justifying the Capability Account**

As I made clear in the previous Chapter (§4), the primary notion of capability that is intended to do work in the version of the capability approach that I want to defend is extremely strict, in one respect. This is because someone only counts as having a fully-fledged capability, on this view, if they are in a state such that a properly-constituted (uncoerced, etc.) act of their will cannot fail to make a difference, in determining whether or not they instantiate a valuable

---

<sup>25</sup> This does not become fully explicit until 2000a, but Nussbaum's constructive engagement with Rawlsian liberalism begins with her 1997 and (especially) 1998.

functioning. If they choose to function, they will do so; if they do not so choose, they will not. This is a much more rigorous and precise specification than that which generally plays a role in Nussbaum's work (or that of Sen). It is, correspondingly, also a state that will be comparatively rare. On the other hand, as the previous section suggests, I am very sceptical that capability – on even a much weaker construal than my own – could provide anything like a *complete* basis for an account of justice, or (even more clearly) an account of the human good. Firstly, as critics such as Arneson have repeatedly stressed, sometimes (at least) justice requires that people function, rather than *just* having capability, and in some further cases it is obscure why capability is of great value, over and above the value of its related functioning.<sup>26</sup> Secondly, there is a potential tension between the strictness of a conception of capability, and the tightness of the focus upon capability so conceived. The more strictly capability is defined – the harder it is to possess, relative to its functioning, and thus the 'further away' it is from its functioning<sup>27</sup> – the more difficult it will become to justify an exclusive or near-exclusive focus on it. Thus, if a fairly strict conception of capability is warranted, as I shall argue is the case, only a relatively loose focus on capabilities as components of the good life or of the life-relative-to-which-justice-has-been-done (what I am calling 'dignity') will be warranted in turn.

As things stand, I have given some reason to hold that we should not pursue an exclusive focus on capability, and I have been working with a fairly strict biconditional definition for

---

26 It may well even be that there are cases where causing people to have effective opportunities to function rather than causing them to function is actually a very *bad* thing, rather than just being of negligible value. This will be the case to the extent that giving people a choice about doing things always leaves open the possibility that they won't do them; less obviously, it might also be the case that some things that are of value in even adult lives are best when they *aren't* planned, deliberated about, and chosen from a broad space of open alternatives. But I won't pursue this thought further here.

27 As is often the case, this notion of distance can be identified with some function that determines a closeness ordering of possible worlds, but I won't attempt to give one here; I think that the intuitive idea will suffice.

capability itself. This opens the way for a defence of capability, strictly conceived, as something that is of intrinsic value in human lives, and something that, in many instances, human lives need to contain in order for justice to be done with respect to them. In what follows, I will pursue both ends (the defence of the strict definition, and the answering of the 'why capability?' question) simultaneously. To begin with, schematically, the case for capability as a distinctive component of good lives, and one with which justice should be particularly concerned, is as follows. Capability is essential because it guarantees that people are able to actively choose the good, in those cases where the good is not exhausted by choice-involving functionings directly<sup>28</sup>. The thought is then – again, schematically – that a chosen good is, *eo ipso*, better than an unchosen good.<sup>29</sup>

However, 'choice' is hardly a transparent concept here, and must be unpacked in several ways<sup>30</sup>. One way appeals to the diverse array of variants of a functioning that (plausibly, at least) can only arise when the creativity of agents with open possibilities is in play. This is probably part of what Nussbaum has in mind when she talks about the value of being e.g. able to choose what and how to eat, over and above the nutritional or gustatory value of what is actually consumed. Functionings that can be instantiated in more or less *creative* ways – which is, arguably, most of them – are, then, particularly apt for a capability-theoretic

---

28 As Nussbaum says (see above), there are probably some good functionings that *constitutively involve* practical rationality, and the permanent possibility of effective choice, to such an extent that anyone who instantiates them *must* possess capability in even my very strict sense. Clearly, for these cases no additional argument for the value of capability is required, over and above an argument for the value of the functioning.

29 A different, more theory-laden way of putting this would be to note that the *virtuous* choose the good, whenever the good is apt for choice; virtue is at least usually a species of autonomy. Anyone who is heteronomously pushed into instantiating a justice-related functioning will thereby be prevented from living the life of virtue; at least in many cases, this will be deeply unjust. Cf. Kymlicka 1988, 183, on the necessity of a life 'lived from the inside'. See also the next paragraph.

30 Arguably, there might be as many qualitatively distinct but similarly good-making features of the processes that generate a functioning as there are distinct valuable functionings to be generated (maybe more, indeed). 'Choice' would then merely be a theoretical place-holder term for this otherwise-undistinctive mass of features. I am particularist enough not to be worried if this turns out to be the case. Cf. Ch.4, §2.1.

treatment. Alternatively, it can be valuable to *plan* certain activities with the veracious awareness that one's plans will make a difference to how things go. Similarly, *deliberation* about whether or how to go about something, alone or with others, can be a valuable process, and can be more valuable, again, when it will be effective in changing the course of events. Finally<sup>31</sup>, and most importantly, these processes are generally, if not necessarily, involved in coming to recognise the value of functionings for oneself. It is often only through deliberation, discussion, creative exploration, and so on, that the true value of activities, relationships, etc., becomes clear. This is especially important within the context of the Aristotelian tradition that will become increasingly prominent as this Thesis progresses, since it is almost a commonplace within that tradition that the virtuous – those who lead good lives – do not merely do the right thing (neutrally described), they do it for the right *reason*; this requires that a person can respond appropriately to the right features of things in functioning, and agential etiology is clearly important for this to occur. But even apart from a commitment to Aristotelian thinking, it seems as plausible as anything – ethically speaking – that doing something in recognition of its value, and (thus) because of that value, is generally better than doing it incidentally.

This provides a solid motivation for some degree of focus on capability as well as functioning, but I should still spell out why it is particularly relevant to the more rigorous definition of capability that I have proposed. Each of the aspects of agential involvement in

---

31 The etiology of a functioning can also make a difference in ways that do not obviously have to do with agentic capacities. For example, it might be better to instantiate some (or all!) functionings when they are performed against a background of social equality, respect, etc. Or, perhaps, it might be best if they occur in the context of widespread societal consensus, perhaps as a result of democratic deliberation. Sen distinguishes between 'process' and 'opportunity' aspects of freedom, somewhat along the same lines (e.g. at 2009, 228-31). I won't, however, explore these etiological factors further, since – however important they might be to a truly comprehensive overview of the components of a good or dignified life – they are less distinctive of the capability approach, and that is my target view.

functioning that I have highlighted are connected, in more or less direct ways, to the possession of effective choice. The virtue of the strict definition, then, is that it guarantees that effective choice will be present in a way that will actually make a difference to the way the chooser's life goes. For contrast, consider a slightly weaker variant of the strict definition. This would drop the biconditional clause, thus: if an individual has the capability  $C_f$ , they will instantiate functioning  $f$  if they choose to do so; not, that is, *only* if they choose to do so. Here, the capable agent does have an effective choice, but it is left open that their choice may not make any difference to the way things go for them, since it may be that they would have instantiated the functioning regardless<sup>32</sup>.

I do not want to suggest that a person who has the weak capability for some valuable functioning is in a bad situation – they do have a genuine opportunity to realise something that will make their life go well. However, some of the features of the agentic capacities that make capability so significant are not necessarily fully present in this case. Take Sen's classic malaria example<sup>33</sup> – he holds that someone who has a public health protection from malaria thereby gains a valuable capability. On the weaker definition, this is plausible – they do gain an effective choice to be malaria-free (and malaria-freedom is a very valuable functioning). Nevertheless, on the stronger definition, this is not a complete capability, since the choice to be malaria-free can make no difference; in the counterfactual case where the choice is not made, the agent instantiates the functioning regardless. However, consider the valuable processes involved in choice, outlined above, at this point. Firstly, is it plausible that the (*ex hypothesi*) capable individuals would gain a valuable sort of creativity from malaria protection? No. People *might* collectively plan an anti-malaria policy, and engage in

---

32 Compare Vallentyne 2005, esp. 363-6, on the distinction between 'control' and 'effective' freedoms, and his argument for the only partial importance of the former.

33 Sen 1993, 54-5. Cf. Cohen 1993, 33.

deliberation to that end, but this could only be an extremely limited process, since once effective protection was in place, there would no longer be any need to engage in such activity. Deliberation and planning would become pointless for at least almost all of the individuals involved – apart, that is, from the technical (i.e. medical; environmental) professionals who would be implicated by the continuing process of securing the capability through time<sup>34</sup>. It appears, then, that the distinctive agential processes that pick out capability as ethically significant are not significantly present in this sort of case<sup>35</sup>. They are only likely to be fully present where strong capability is realised. This gives us excellent reason to conclude that it is only if we include some elements of strong capability into our approach that we can really remain true to the distinctive motivations for the capability approach *per se*.

To conclude, what does the view of the life of dignity look like, once this is taken into account? What does justice require that every human life contain? My answer is that lives must contain many different functionings – some of them (like adequate nutrition, adequate health, and the social bases of self-respect) must be constantly present, while others (like fulfilling employment and leisure time) are temporary, and must be distributed appropriately across life. In many cases, it may also be of some value for people to have relatively weak forms of capability for these functionings, even where strong capability is impossible. But human lives, in order for the value of practical rationality, creativity and collective deliberation to be adequately represented in them, also need, at certain relatively rare points, and with reference to certain functionings only, to contain capability in the strictest sense.

---

34 Clearly, any creativity involved for these people will constitute an aspect of their valuable labour, not an aspect of their malaria-freedom itself.

35 Arguably, what this suggests is that protection from malaria is not plausibly a valuable *capability* after all – having the opportunity to be malaria-free isn't significantly more valuable than just (securely) being so.

People need many things to flourish; one of them is for their agency to make a real and positive difference to the world.

### §1.3 The Bounds of Justice

Conceptions of justice can helpfully be distinguished from one another in terms of the thought that each involves substantive outer and inner bounds. The outer bound concerns the maximal set of patients of justice – those with reference to whom justice should make claims; the inner bound concerns the relative size of the 'private' and 'public' spheres – the regions of human relations which form the proper ambit of the norms of justice. In this section, I set the stage for the discussion of universalism and the proper role of theories of global justice in the Thesis by briefly summarising, and then noting some important problems with, Nussbaum's central claims about the location of these bounds.

Nussbaum's approach to the former question is relatively easy to summarise: Nussbaum's theory has always had a global scope, and rests on a cosmopolitan perspective, at least to the extent that it applies to all human beings as individuals, regardless of nationality. She notes that it does not follow from this that any implementation of the approach must occur at the level to which it most fundamentally applies (2000a, 101-105; 2006a, 306-324), but consistently demonstrates a willingness to advocate pragmatic interference with nation-state practice to protect capabilities entitlements, at least where the most serious injustices are involved. I do not disagree significantly with this aspect of Nussbaum's thinking. I accept both the fundamental cosmopolitan answer to the question about application, and the recognition that this has no immediate consequences of itself for what justice would prescribe. Rather than constituting (i.e.) a directly applicable decision procedure for NGO,

global-governance, or nation-state policy-making, the capabilities approach provides a more-or-less general ethical orientation, which Nussbaum sometimes (e.g. at 2000a, 102-6) compares to that contained within the necessarily-vague promulgations of constitutional statute<sup>36</sup>. My only significant argument in the vicinity is that, until the ethical role of humanity is completely clear, the justification for a universalist cosmopolitanism will not be clear either. Nussbaum's adoption of a Rawlsian 'political-liberal' ban on the involvement of 'comprehensive' or 'metaphysical' doctrines would preclude the kind of Aristotelian story of which she was once a proponent<sup>37</sup>; clearly, there is no consensus that the nature of the human species generates norms of itself<sup>38</sup>! It is not clear what could take its place: if the suggestion is that it simply follows from contemporary global political culture that all and only individual humans are patients of justice, this will – at best – simply beg the question, since (however contingently) all of the participants in that culture are human.

There are two important potentially problematic aspects of Nussbaum's thinking about the domain of justice and its practical role. Firstly, some of her language links capability-theoretic justice too cavalierly with the purview of state institutions. This limits the applicability of the capability approach, as well as building in dubious statist assumptions without proper argument. Secondly, her commitment to Rawlsian public reason threatens to implicate a public/private contrast which would be inconsistent with her feminist commitments. Reflecting on her attempts to avoid such implications helps to clarify the extent and nature of her liberal assumptions about justice, and demonstrates the need for

---

36 It is also reasonable to suppose that Nussbaum agrees with this thought since it is close to Rawls' view of the status of principles of justice. See esp. Rawls 1997, 766-68.

37 See Chapter 5, §1.2 for a brief account of Nussbaum's original Aristotelianism about humanity, and Chapter 5 more generally for my appropriation of Aristotelian ideas in giving a similar story.

38 'Of itself' is important here – many proponents of 'human rights' rest their case on the (putatively complete) overlap between personhood and humanity, for example – and in this account personhood is doing the normative work, not humanity itself.

alternative claims at certain points. In the first two subsequent subsections, I focus on these different questions about the inner bound. In the third, I take a broader view, evaluating the liberal bent of Nussbaum's feminism, and arguing that, so long as negative-libertarian values are not privileged over others, and once the role of individualism has been clarified, her liberalism – if it is such – is acceptable.

### **§1.3.1 The Inner Bound: the Role of the State**

The first problem concerns the relation between the requirements that capability theory is supposed to generate, and the proper domain of state – or other large-scale institutional – action. This is a salient issue for my purposes because Nussbaum thinks of the standard of capability-theoretic justice as generating prescriptions primarily for governmental or other macroscopic institutional action: typical wording, from *Women and Human Development*, speaks of 'the structure of social and political institutions' (2000a, 75) as the normative target of the capabilities approach<sup>39</sup>. In this section, I shall argue that this is problematic, firstly because it implies that state and other macroscopic institutions must have a large amount of power over very many domains of life, which is both dubious and unnecessary, and secondly because it would limit the conception of dignity to elements which large-scale, powerful, and typically coercive institutions can actively support.

The first aspect of this problem concerns the coercive and intrusive (distinctively *power-laden*) tendencies of state action. Nussbaum exposes herself to attack on these grounds in much of her work. Most explicitly, she says that 'the central capabilities always supply a

---

<sup>39</sup> This is not unambiguous across Nussbaum's corpus as a whole; in 'Aristotelian Social Democracy' (1990a), by contrast, she begins by setting herself 'the task of political arrangement' (208-9), which seems neutral between institution-focussed and more basically relational strategies, although she soon glosses this as a question about 'the job of government' (214).

compelling interest for purposes of government action' (2000a, 280). This may not entail that failures of dignity will always require *actual* government action – the government could have a prerogative that it (contingently) need not make use of, because it happens that the relevant needs are already met – but it does put the emphasis strongly on state-based solutions. Several thinkers have criticised this aspect of Nussbaum's work, usually with reference to a sub-set of her central capabilities for which it seems especially likely that state action is unlikely to be appropriate. Eric Nelson (2008, esp. 98-103)<sup>40</sup>, Henry Richardson (2000, 314-15), and Thom Brooks (2014, 6-7) all raise versions of this criticism. The point is clearest in Brooks<sup>41</sup>:

Consider the capability to bodily integrity. For Nussbaum, this capability includes 'having opportunities for sexual satisfaction'. What would it mean for a community to secure the capability of its members to have such opportunities? The community cannot require any fellow citizen to love you, nor desire sexual satisfaction with you. Neither may the community enforce existing couples to desire sexual satisfaction with their partners. Either scenario would be highly intrusive and unacceptable. Moreover, it is far from clear how a community might act through public policy to ensure the capabilities threshold is reached. Would we expect to find government-funded matchmaking websites or speed dating clubs? (6-7)

One option, in response to criticism of this kind, would be to abandon the claim that the capability for sexual satisfaction is an aspect of dignity. Consider also, against this background, what the state might have to do to ensure that individuals can '[enter] into meaningful relationships of mutual recognition with other workers', or '[be] able to have attachments to things and people outside ourselves'<sup>42</sup>. Thus, there are likely to be many

---

40 Nelson's primary concern is actually with *neutrality*, and whether the role of the state in Nussbaum is compatible with the neutralist (Rawlsian) liberalism she endorses. I agree that she can be attacked on those grounds, but I do not ultimately endorse neutrality anyway, so I need not press this point here.

41 An important aspect of Richardson's and Brooks' criticisms concerns whether the state could coherently promote only capability in many instances, because there is no real gap in practice between functioning and capability in some cases, because the functionings in question are intrinsically other-involving. The capability for sexual satisfaction is like this; Richardson also suggests that various political rights are too (2000, 319-26). For example, a person cannot be capable of living in mutual respect with others unless others actively respect them (which is the relevant functioning).

42 See Appendix A.

capabilities for which similar difficulties arise, at least once the massively power-involving nature of governmental involvement is made clear, so the same is likely to go for many of the most intimate capabilities on Nussbaum's list. However, a capability approach limited in this way to 'bread and butter' capabilities (Nelson 2008, 102) will clearly fail to satisfy the motivations which, I have noted, gave rise to it. In Nussbaum's conception, at least, the capability approach is supposed to be motivated by the thought that justice requires everyone to have an equal basis for a minimally good life. If 'minimally good life' here merely referred to (e.g.) the failure to starve to death, this revision would be compatible with that. Clearly, however, this is not what any capability theorist has had in mind – what we want is not merely for people to be enabled to continue to (just barely) *live*; what we want is for them to be enabled to *flourish* (at least, in some substantial sense).

What is more, it is not really clear from the outset why a restrictive emphasis on governmental action should be warranted. If the motivation for the capability approach is to enable people to live with one another in a way compatible with their human dignity, it is far from obvious that the government always need be implicated even as a safety net. An alternative option is to begin by thinking more holistically about how people need to be relating to one another (including through the institutions that affect them) in order for their dignity to be secure. On this approach, government action only needs to be implicated when A) it is actually needed, and B) when it is genuinely the best relational means available. For cases like the capability for sexual satisfaction, at least, these conditions will rarely be met<sup>43</sup>, but this does not necessitate a watering-down of our conception of dignity.

---

43 I say rarely, not never: sexual/relationship education policy should probably play a role here, as well as the prohibition of genital mutilation that is Nussbaum's focus (2000a, 94).

The second aspect of the problem, then, concerns whether governmental action will always (or even usually) be the *best means* of securing capability, if I am correct that, fundamentally, the motivation for the capabilities approach runs broader than the governmental sphere<sup>44</sup>. One lacuna that arises is Nussbaum's approach concerns a lack of attention to non-governmental action aimed at (most obviously) the removal of societal (but not paradigmatically institutional) barriers to equal capability, such as sexism and racism. These structural injustices affect capability both on a broad scale – because they generate inequalities of treatment across most of the capabilities on the list – and on the narrower scale – because they intrinsically violate certain capabilities, such as the capability to 'be treated as a dignified being whose worth is equal to that of others'<sup>45</sup>. The clear relevance of non-governmental action, e.g. that occurring within social movements such as feminism, suggests that this lacuna is also motivated by Nussbaum's emphasis on the institutional sphere. As I have already noted, her claim may not be that the *entirety* of capability-theoretic normative thought is restricted to what falls within paradigmatically institutional (usually, state-level) frames. If not, however, it is unclear why Nussbaum never discusses how capability-theoretic justice could govern the actions of 'private' individuals and groups, by – for example – motivating campaigns against sexist or racist advertising, or anti-bystander initiatives to reduce the incidence of microaggressions. After all, these largely non-institutional forms of praxis have often been more prominent in action for the effective promotion of social justice over recent decades than legal or regulatory changes.

In conclusion, there appear to be good reasons not to adopt a restrictive emphasis on the

---

44 Of course, it will often be the case, empirically speaking, that legal guarantees and mandates for state action *are* the only way to reliably ensure that the capability threshold is attained by all. If so, state action will usually be required. But this would still not follow from the basic reasoning that motivates the capability approach itself.

45 Again, see Appendix A.

space of (especially governmental) institutions when thinking about the practical and evaluative targets of the capability approach. In what follows, I shall operate on the understanding that – in principle – any way in which individuals relate to one another might be a target for change, with state action providing only one of the means to ensure the equal dignity of all.

### **§1.3.2 The Inner Bound: the Personal as Political**

The second of the problematic regions of Nussbaum's thinking about the domain of the political concerns her response to an important line of feminist and other radical critique, which corresponds loosely to the slogan 'the personal is political!' Generally speaking, this critique generates suspicion of attempts to restrict the space of the political to traditionally 'public' institutions and the spaces they affect, arguing instead that justice and other political norms must be applied unrestrictedly across society. Many feminist thinkers have noted ways in which the traditional liberal presumption that certain regions of relational life are beyond politics, and cannot be evaluated or interacted with on the basis of justice-involving critique, has functioned to protect the unequal standing of powerful or otherwise privileged individuals. I cannot provide an expansive overview of this work here. However, Jenny Saul provides a nice encapsulation of the orientation:

Many feminists...maintain that there is much more to power relations between the sexes than the presence or absence of laws against sex discrimination. In particular, they argue for the significance of matters traditionally considered private and personal, rather than political – the structure of families, sexual relationships, appearance-related concerns, or the words we use. (2003, 3).<sup>46</sup>

Nussbaum has attempted to accommodate this species of anti-privacy move within the

---

46 Cf. esp. Benhabib 1993; Jaggar 1991; Young 2009; as well as many of the essays within the Saul volume itself, for several useful variations on the theme.

confines of political liberalism, relying on the work of Okin (1989) in doing so (e.g. at 1999a, 66-7). She argues that social institutions such as the family must be regarded as parts of the Rawlsian basic structure – the 'public' side of the public/private divide (2000, 241-297, esp. 252-283). However, even if this were sufficient to assuage Okin-style doubts about privacy, Nussbaum's acceptance of a Rawlsian framing – with its constitutive constraints on the sorts of claim that can be made within the 'political' domain – generates *of itself* a public/private contrast. This is necessarily the case, because there must be some non-empty set of 'non-political' normative considerations, and parts of life to which they pertain, if political liberalism is not to amount to a ban on most<sup>47</sup> ethical disagreement *tout court*<sup>48</sup>. For example, she has explicitly claimed, following Rawls, that 'churches and universities' (2000a, 272-3) are not part of the basic structure, and so are not fully subject to the general requirements of justice. The explicit reason she gives is that these entities 'have a life outside the state...; [while] families do not'. The test is, then, whether or not an aspect of human interrelation 'is a social and legal construct in a... fundamental and thoroughgoing way' (2000a, 275).

However, it is unclear both why churches and universities are not recognised as social, if not legal constructs, and why this is a relevant criterion at all. In the absence of a strict restriction of capability-theoretic attention to the domain of direct state action, which I have already argued would be untenable, there is no obvious reason why the existence of a 'life outside the state' should have this deep significance. Perhaps, if the rectification of injustice were always

---

47 It is necessary to say 'most', because the later Rawls eventually concedes the existence of a plurality of 'reasonable' liberal conceptions of justice, in addition to his own conception involving the two principles: by Rawls 1997, the public reason methodology is described as generating a 'family of liberal political conceptions of justice' (e.g. 767, and throughout). Thus, there is some reasonable disagreement even within the political domain of basic justice.

48 This criticism will only apply to 'consensus'-based versions of political (or 'justificatory' (Gaus 1999)) liberalism. 'Convergence' models, such as those of Gaus (2012) and Vallier (2011) would not have this implication, even if they were extended to cover a more inclusive conception of the political domain.

a matter of state action (especially coercive state action), it would make sense to restrict the application of justice-concepts solely to areas of life with which the state was already deeply involved (if that is not, in any case, an all-encompassing set). But I have argued that injustice does not always consist in, and is not always best rectified by, state action. A plausible alternative explanation is that Nussbaum's real reason for adopting a restriction to the 'basic structure' is a classically liberal concern with negative freedom, reflecting her evaluative rejection of a world in which religion- or education-involving choices are constrained by (even non-state?) others in the name of justice. It is compatible with everything I have and will say for this to be a sound evaluative judgement within the capability approach. But, again, there is no need to invoke a principled public/private boundary in order to do so; if a certain action – even one motivated by concern for a core capability – would significantly<sup>49</sup> violate a dignity-involving capability (which at least some non-state interventions within academic or 'religious' life plausibly would), that is sufficient to regard it as unjust. The appeal to the 'basic structure' in bounding the *polis* is at best a theoretical third wheel, and at worst an illegitimate means of smuggling in negative-libertarian values without argument.

Once the anti-privacy motivation is on the table, there is no reason to preserve any public/private distinction in principle; instead, the contours of political critique of human relations are justified in directly political terms<sup>50</sup>. We are always justified in deep hermeneutic suspicion (Ricoeur 1970, 32)<sup>51</sup> of proposals to rigidly cordon off areas of human life on grounds of privacy, given the effectively anti-egalitarian history of such proposals. It may be

---

49 Enough, that is, to bring or leave someone below the relevant threshold.

50 I am assuming throughout that, since justice is an ethical concept, there is at least a sense in which politics is continuous with ethics. When my discussion turns to meta-level matters, I often refer to ethics as my subject-matter; the 'ethics' I have in mind throughout includes the concept of justice, and is, to that extent at least, also political. It may well be the case, as a result, that 'politics-first' modes of political philosophy such as political realism are beyond the dialectical scope of this Thesis.

51 Cf. Leiter 2004, which provides a fairly expansive development of the concept.

that there are some aspects of our lives together which are genuinely private, in the sense that they rarely (or never) affect human dignity, and so will be of negligible (or no) relevance to considerations of justice. But, if so, this is a normative thesis in its own right, and the reasons for hermetic encapsulation of such parts of life can be expressed within the space of political (justice-related) norms using existing concepts. For example, rather than stipulating that the bedroom, or child-rearing, or academic mentorship, or comedic speech (or speech *per se*) exist beyond the properly political realm, one can simply argue concretely that (e.g.) the putatively valuable capability (that is, freedom) to engage in sexist joke-making is precluded by (e.g.) public expressions of rejection, non-state direct action, and/or state coercion<sup>52</sup>. But this argument must be made, clearly, and on the merits, rather than surreptitiously relying on the liberal assumption of privacy that has done so much to protect injustice in the past.

For this reason, as well as those surveyed in the previous subsection, in the remainder of the Thesis, I shall operate (*pace* much of the liberal tradition) with the presumption that the *polis* – the domain to which norms of justice apply<sup>53</sup> – has no inner bound. This will be important

---

52 Clearly, the additional claim – about vocal rejection or other non-state action – will have to be made in many cases, since I do not closely link the rectification of injustice or preservation of justice to state action alone. Many political actions which might offend against the putative value of (some conception of) privacy, then, would not do so apart from the deployment of state coercion.

53 There is a risk of straw-manning the (e.g.) liberal opponent here – i.e. the Rawlsian strands in Nussbaum – by overstating the extent to which principles of justice do not extend to elements of relationship outwith the basic structure. Rawls' late position is that although principles of justice do not apply fully and directly to the 'internal life' of non-basic institutions (e.g. at 1997, 789), they do constrain the form of such institutions from the outside, guaranteeing most pressingly an effective capacity to exit the institution (789-90). The upshot of this, Rawls claims, is that 'the spheres of the political and the public, of the nonpublic and the private, fall out from the content and application of the conception of justice and its principles. If the so-called private sphere is alleged to be a space exempt from justice, then there is no such thing' (1997, 791). It is not the case, then, that Rawlsian justice incorporates a private sphere whose boundaries are not determined by the conception of justice (and other political values). Rather, liberal justice is supposed to be self-limiting, restricting its central principles to a sub-set of relationships on grounds of freedom-promotion, or non-applicability, or both; the general liberal rationale suggests the former, some of Rawls' comments, e.g. about non-political conceptions of justice within the family (*ibid.*), suggest the latter. I do not think it obvious that a vague sufficientarian conception such as that which I am developing here would have to incorporate such theory-internal restrictions, even if it were proposed as a complete theory of justice; the additional flexibility of the capability approach should help to avoid non-applicability worries, at least.

in what follows insofar as I eventually endorse – in Chapter 6 – a fundamentally relational Aristotelian conception of justice which represents traditionally 'public' (governmental, civil-societal, etc.) modes of interrelation as comprising only part of the *polis*.

### **§1.3.3 Liberalism?: Individualism and the Privileging of Choice**

The two strands that I have discussed, taken together, motivate a turn away from some of Nussbaum's restrictive emphases in developing her capability approach. In Chapter 6, I will make use of both elements in articulating a somewhat different interpretation of the proper role of capability theory in guiding political practice. However, another controversial aspect of Nussbaum's work on the political role of the capabilities needs to be discussed; here, too, my project will need a solid basis for further development. This aspect concerns the liberal nature of Nussbaum's proposals. I discuss two forms that Nussbaum's liberalism takes within her deployment of the capability approach, and arrive, after clarifications, at the conclusion that both her reformed individualism and her incorporation of the value of free choice are sustainable in principle.

Firstly, I need to consider the difficult issue of whether or not the liberal nature of Nussbaum's political framework builds in a problematic metaphysics of the individual self, assuming that it does (or should, and therefore can) stand apart from, and be in rational control of, every element of its interactions with others. However, there are related areas of recent dispute with which I am not concerned. Firstly, I am not interested here in any accusation that liberalism builds in a suspect picture of the self as 'unencumbered' by (irrevocable) 'constitutive ends' (Sandel 1984, 86), or that it deprives people of the security of socially given (and thus involuntary) roles (MacIntyre 2007, esp. 241-256). Apart from the

question of whether such attacks were ever well-targeted against Rawls (cf. Bell 2012; Rawls 2005, 22-8), it is not clear that Nussbaum would be vulnerable to them in any minimally plausible form. Her account does not imply that people are or should be detached from others (Nussbaum 1999a, 59-62). It probably does imply that we cannot give significant *weight* to all those conceptions of social roles, projects, or relationships of which it is true that a person 'could not understand [themselves] without it' (Sandel 1984, 86), but this is for good reason, since many such conceptions are incompatible with dignity, for the person in question or for those they interact with. It is also likely that a plausible moral epistemology would violate any such restriction in another way, by requiring that people be able to be at least minimally critical about any element of their lives, even (e.g.) relationships of motherhood (Nussbaum 1999a, 71-8; Kymlicka 1988, 190) or a sense of religious calling. However, it is hard to see how any plausible and widely-applicable ethical perspective could fail to violate such a restriction. What is more, a Nussbaum-style capability approach, because it explicitly takes a stand on (at least some) questions about the good, rather than hiding beneath a veneer of superficial neutrality, is well-placed to take the bite out of many such communitarian criticisms. In particular, it is possible within the capability approach to address concerns about damage to relational or communal goods at face value; if it is genuinely the case that a justice-relevant good is at stake – and my approach has been extremely permissive about what could count as a justice-relevant good – the capability approach should be able to take note of this fact. Secondly, I am also not concerned with criticisms of the idea that individuals are the fundamental locus of politically-relevant goods, such as the good of dignity, at least once that idea is distinguished from the claim that individuals cannot be benefited or harmed in virtue of their relation to groups (Nussbaum 1999a, 62-3).

Instead, I am interested in whether Nussbaum has a plausible enough social ontology to accommodate the strong role that group membership and interrelation plays, both in the good of individuals and in ethical epistemology<sup>54</sup>. My primary worry is that she has never discussed the role that membership in oppressed social groups can play in precluding dignity for a person in adequate detail. Her *oeuvre* does include helpful materials for this purpose. For example, her influential paper on objectification and pornography contains discussion of the effects of objectification on women *as women* rather than merely as (non-gendered) individuals (1999b, 213-240). What is more, she has consistently emphasised the relevance of culturally-ingrained norms (e.g. of gender) to whether or not an individual can function<sup>55</sup>. However, a satisfactory framework for global justice requires explicit acknowledgement that these sorts of thing are linked: that membership in a socially-salient group can *eo ipso* damage dignity, and thus that the status of such groups is of substantial relevance to capability-theoretic justice.

Since Nussbaum has not provided such an account within her capability approach, the question is then whether one can be incorporated which will not conflict with her liberal outlook. In particular, it may seem that the liberal commitment to taking the individual as the central concern of political norms is likely to prevent recognition of these important social factors. However, as Charles Mills has recently argued (2012), in a survey of various attempted justifications for a divergence of radical critical philosophy from liberalism, it is not obviously essential to liberalism to build in a problematic concentration on the individual (methodological individualism), beyond the locus-of-value claim that I have already accepted ('normative individualism'). In particular, a view which accepts that individuals are the locus

---

54 Insights from feminist social epistemology are central in the next Chapter.

55 The relevance of gender norms concerning 'modesty' to acquittals for rape in India is noted at (2000a, 30), for example. There are many other instances in her corpus of work.

of political value, and accords a significant role for individual freedom of choice, does not have to suppose that individuals have a significant degree of control over their group memberships, the constitution of the groups of which they are members (i.e. whether they are oppressed or otherwise hierarchically positioned), or the effects that their group memberships have on their well-being. We need an account of group social ontology and the group-individual relation which will not presuppose that individuals have this implausible degree of control. Chapter 2 of Ann Cudd's 2006 work, *Analysing Oppression* (29-54), presents a view of the group-individual relation which satisfies this desideratum<sup>56</sup>. I need not discuss most of the details of her account here – since for current purposes all that is required is a way of making sense of the irreducible contribution of social groups to justice-relevant states of affairs – but the upshot can be seen in the following definition:

A social group is a collection of persons who share (or would share under similar circumstances) a set of social constraints on action. (44)

As Cudd makes clear (42), 'constraints' are facts which 'frame' or 'guide' action; they can be incentives as well as (e.g.) threats or preventions. Examples of such social constraints include racial stereotypes and their attitudinal consequences (43), and facts about differential likelihoods for significant consequences, such as the increased likelihood of sexual assault when walking alone (42)<sup>57</sup>. Importantly, on this account many groups – including many of those with the largest actual-world effects on dignity – are 'non-voluntary', in the sense that they 'are formed not by the intentions of the individuals in them... but by the actions, beliefs, and attitudes of others, both in the group and out, that constrain their lives in patterned and

---

56 My endorsement here is solely of Cudd's analysis of the supervenience of social groups on the individuals that they comprise, and her basic thought that damage accrues solely to individuals, even when it accrues in virtue of group membership, and where the group is explanatorily irreducible to individuals and their non-relational properties. I do not need to (and do not) accept her analysis of oppression itself, or the rational choice psychology she draws upon in giving her individual-level explanations, or many of the other details and applications of her account.

57 Cudd does not specifically mention implicit biases or their effects, but these would readily fit within her account.

socially significant ways' (44). What is more, as well as its being the case that the group's existence, constitution, and particular membership, are not subject to individual volition, it is also possible for individuals to be dramatically *mistaken* about which groups they are members of and/or the nature of those groups (45). This account is suitable to my purposes, because it recognises the limits of individual choice in securing dignity while remaining compatible with the normative individualism that, as Mills argues, lies at the heart of the liberal tradition.

The second aspect of Nussbaum's liberalism which requires clarification and evaluation concerns the role of liberty itself in her approach. I have already noted some ways in which freedom functions as a value within the capability approach, including within the justification for a focus on capability itself. What is more, several traditional liberties are included on the list itself – freedom of association, freedom of speech, freedom of conscience<sup>58</sup>; this makes these freedoms 'central and nonnegotiable' (2006a, 80), at least without tragic costs. However, one might be suspicious, given Nussbaum's more recent placement of the capability approach within the Rawlsian tradition of liberalism, with its lexical prioritising of liberty over e.g. redistributional change (Rawls 1999, 214-220), that she builds in an implausibly strong role for freedom of choice. Nussbaum has explicitly and repeatedly disavowed the idea that 'negative' or 'neoliberal' freedom is of intrinsic or all-purpose importance (e.g. 2003, 43-6; 2006a, 216-17). Nevertheless, there are two areas of remaining concern.

Firstly, the Rawlsian framework builds in a large role for choice through the methodological restriction to elements which can be part of an 'overlapping consensus' between 'reasonable'

---

<sup>58</sup> See Appendix A.

ethical conceptions. This is potentially an incorporation of space for choice of vast scope, depending on how many, and which, conceptions are judged 'reasonable', and how much overlap there then turns out to be amongst them. I suspect – although I cannot substantiate this here – that a view such as Nussbaum's could not possibly be the subject of such consensus, without an unusually restrictive interpretation of 'reasonable'. If this is right, her acceptance of Rawlsian assumptions would mandate a space for choice that could not be interfered with on grounds of justice of a width that would be implausible by her own lights; she would have to give up much of what she wants from a theory of global justice as a result<sup>59</sup>. Since I cannot substantiate this here – both because of space constraints and because it is in part a difficult empirical (sociological) question – I shall leave this concern to one side, although the alternative I'll present through the rest of the Thesis will not be vulnerable to this type of problem, because I drop the Rawlsian requirement entirely.

Secondly, Nussbaum appeals to a presumption against paternalistic interference (esp. 2000, 51-5), and plainly regards this as a powerful principle against which normative accounts should be judged. This is problematic not as a normative judgement *per se* – although I am suspicious of whether Mill's 'harm principle', or other ideas which rely on a distinction between self- and other-affecting conduct, can make much substantive difference in practice. Rather, it is problematic because it appears to be espoused as a judgement outside the capability approach itself, which will then set limits on what forms that approach could take. This is potentially very damaging to the approach, because it clearly needs to make large numbers of judgements about individuals' goods in order to be practically workable<sup>60</sup>; if all such judgements have to coincide with the self-conceptions of individuals at all times, most

---

<sup>59</sup> Nelson 2008 is particularly trenchant on this point.

<sup>60</sup> See also the brief discussion in the next Chapter (§2.1.1), for a little more on the alleged wrongs of paternalism.

such political action will become unjustifiable. As a result, it had better not be the case that paternalistic interference with freedom of choice is ruled out in principle on Nussbaum's view.

Fortunately, the most that can be generated by the capability approach she proposes, at its most basic level, is the conclusion that since various freedoms (whether incorporated via a restriction to capability-promotion over functioning-promotion, via a more pluralistic conception of each valuable capability/functioning, or otherwise) are entitlements of dignity, violating them in the name of dignity itself will be – at best – tragic. Nussbaum's approach does not give us reason to privilege freedom of choice in general, or even core freedoms such as freedom of speech or conscience, over other capabilities. This puts accusations of paternalism in a new light. Firstly, some forms of paternalistic interference may not be problematic *at all* – as Nussbaum appears to acknowledge (2000a, 53). Others might violate dignity, but be (tragically) necessary to secure the most dignity possible. Finally, some actions might be made wrong by being paternalistic which would otherwise be acceptable, because the paternalism in these case precludes dignity while not simultaneously preserving/promoting it to a greater degree. Only in the final set of cases will anti-paternalist considerations have the sort of immediate deontic consequences that liberals have often proposed, but which threaten to severely restrict the scope for political action.

With these clarifications in place, it appears that the sort of liberalism involved in Nussbaum's overall approach is largely benign. However, I have noted several areas where a better basic account is needed, especially concerning the scope of the *polis*, the domain of justice as such. These modifications will have effects at various points in the Thesis, but especially in Chapter

6, where I draw my claims about justice together, and illuminate the resultant whole.

## §2. Constructivism or Realism?

### §2.1 Introduction

As I have presented it, any capability theory will require a theory of the good (or goods) to fill out its content. Firstly, it needs an account of which *functionings* are goods – as without such it will not be capable of identifying the capabilities associated with those functionings as things with which justice is concerned<sup>61</sup>. Secondly, it needs to be able to claim that the *capabilities themselves* are goods, as otherwise capability theory will be vulnerable to one of the same criticisms that it generally appeals to in motivating itself as an approach – it, like resourcist theories, will be focusing on mere *means*, rather than on something that is of ultimate significance for people's lives.

But in addition, this theory of the good must be *objective*, in the sense of being epistemically secure in the face of persistent disagreement, because it needs to be capable of critiquing contemporary social realities and people's ideas about them. If the normative premises used in critique have no better epistemic standing than those being criticised, such critique will fail to be rationally cogent. The central issue with which I shall be concerned in this second part of the Chapter is that of the correct sense in which the claims about the human good at the

---

<sup>61</sup> Here, as elsewhere, I presuppose a certain conception of the role of capability theory. Some may claim that capability theory may be justified as an essential part of a theory of justice *without* involving any theory of the good. For example, Robeyns (2003; cf. also 2005) claims that we can select a list of relevant capabilities as an input into a theory of justice without presupposing that they are goods in any deep sense. In her view, the fact that there are consistent inequalities of capability and/or functioning (which we cannot suppose are ultimately down to unconditioned choices) *of itself* justifies policy to eliminate them; this argument does not rely on the claim that the things that are unequally distributed are goods. I cannot argue against this adequately here, but I think it is highly dubious to suppose that the mere fact of different treatment could be of normative import if no-one could be said to be harmed by it. This does not require that the harm be direct; perhaps the harm accrues e.g. through a violation of respect or substantive inequality? But then respectful relation to others, or valuable equality with others, would be figuring as an element of the good.

centre of capability theory *are* objective. This comes against a backdrop of disagreement between two (very) broad groups of thinkers in the meta-ethical literature. One group, 'constructivists', identifies the objectivity that (i.e.) capability theory might attain as consisting in agreement, under procedurally idealised conditions<sup>62</sup>. To call an idealisation 'procedural' here is to claim that it does not introduce substantive normative content into the standards which constrain the idealised situation, such that this methodology becomes immediately circular (even in part). This requirement is necessary because even the strongest anti-constructivists may be able to agree that for *some* group of agents objectivity could consist in agreement: so long as relativism is false, the set of agents *who know the facts* will trivially agree in their beliefs<sup>63</sup>, and their beliefs will trivially be true *ex hypothesi*. The second group, 'realists', identifies objectivity as consisting in a relationship to a set of facts that exist independently of any merely procedurally constrained agreement between human agents. A second relevant – if less well-explored – dichotomy among conceptions of objectivity, which differentiates two possible camps of realist views, lies between perspectives according to which objectivity can be 'culture-laden', and those according to which it cannot. 'Culture-laden' perspectives are those that do not begin by including the positions of all existing agents<sup>64</sup>, identifying the substantive constraints on ethical epistemology in a way which is incompatible with some extant enculturated viewpoints, and thus beginning instead with

---

62 Because I am only interested in approaches that are compatible with universalism here, I will not be discussing versions of constructivism which are overwhelmingly likely to generate non-universalist conclusions when applied to questions of justice at the global level. Thus, the 'Humean Constructivism' of Street (2010 and 2012), and Velleman (2013) will not be under detailed consideration here, although Enoch's argument (Ch.4, §1.5) still applies to these versions.

63 This is important: the McDowellian version of culture-laden realism that I endorse in Chapter 4 is surprisingly close to constructivism in some respects, and (as I shall show there) its refusal to endorse any form of proceduralism is one thing that preserves the difference.

64 Here, and throughout the Thesis, I assume that normative epistemology/methodology must resemble the method of reflective equilibrium at least to some extent: such methodology must begin with existing responses, and build from there. If there are methodologies that genuinely radically differ from this (which I doubt), my central distinction in epistemology between constructivists and culture-laden realists will not capture them, and they will be outside the scope of my argument. Compare Setiya 2010, 215-18.

some non-maximal set of perspectives that may, if appropriately developed, admit of objectivity. This distinction will be of central importance in Chapter 3 and 4, and so I shall leave a more detailed account of it until then (beginning in Chapter 3 §2).

Issues about objectivity have especial importance for capability theories because these (as a kind) give a certain significance to the individual will and individual choice. Insofar as they can genuinely be described as justifying a concentration on capabilities as well as functionings, they must include an account of the importance of choice and its underlying agential processes that does not entirely reduce to the importance of the things chosen: if they do not do this, it will not be clear why we should not just promote functionings themselves, without placing importance on capability. It may seem, then, that a perspective on objectivity which puts the substantively-unconstrained will front and centre, as constructivists' focus on what people can/do/will agree upon evidently must do, will be more congruent with the impetus of capability theory as a whole. Whether or not this rationale is ultimately convincing (and my position will be that it is not), it provides some motive to examine whether realism is even *compatible* with the capability approach.

This issue is also of significance because if constructivist or non-culture-laden realist views of objectivity are correct, this will have (potentially very radical) consequences for the way that we must engage in normative thought, and (in particular), for the attempt, typical of Nussbaum's capability approach, to draw substantive conclusions about the way the world should be in the absence of any obvious consensus. Constructivists take objectivity to be a function of agreement among all competent agents, where the relevant conception of competence cannot – at least directly – bring in substantive normative content. However,

there is very little agreement about – for example – whether the capability to achieve sexual satisfaction is of universal value, unless the relevant set of judges is substantively constrained<sup>65</sup>. They are, then, not merely committed to the view that *we do not know* whether a person who judges something like this to be the case is objectively correct; they are committed to saying that she *is not* objectively correct (at least, not yet). This problematic is particularly pressing because many proponents of constructivism in the arena of political philosophy advocate radical actual-world versions of it, involving proposals of deliberative democracy that are in most cases very distant from the contemporary status quo<sup>66</sup>. Advocating such methodologies threatens to put on hold all prescriptions and actions that are not aimed at bringing such a deliberative-democratic situation into being, because until such a procedure is realised we can have no objective guidance as to what to do.<sup>67</sup> Other versions of constructivism settle for a merely *hypothetical* consensus, usually including a very large amount of procedural idealisation, which might reduce the extent to which there is a conflict in practice between advocates of tradition-laden realism and these opponents<sup>68</sup>. This division among constructivist positions complicates matters because one preliminary reason to embrace culture-laden realism – the threat of a paralyzing sceptical *aporia* – is not obviously

---

65 Tamir 1996 provides a classic and vivid example of such dissensus.

66 Both the Habermasian and (more arguably) late-Rawlsian wings of contemporary contractarian thought are examples here. Gaus' (2012) more explicitly actual-world-based development of the public reason tradition is an even clearer example. Similarly, this is one motivation behind the turn to deliberative democracy: the thought is often not merely that such democracy will make *for* a good society, but that the relative presence of democratic consensus would *constitute* the worth of a social order.

67 That this is the case is perhaps responsible for the dominance of the rhetoric of democratic authority and distinctively *political* equality in much of today's radical egalitarian discourse, although few take this line quite as strongly and explicitly as Elizabeth Anderson, who appeals to democratic equality *alone* in (at least some of) her work on the capabilities, e.g. her 2003. I suggest that the potentially paralysing *aporia* of actual-consensus constructivism may be *bypassed* by presenting prescriptions as justified by pro-democratic considerations alone (if such justifications are successful), or *disguised* by such presentation (if they are not).

68 The classic recent example is Michael Smith (1994). Rawls' own approach in *A Theory of Justice* (1999), if not in his later work, is also of this strongly hypothetical kind, although parts of what I say immediately below – about the difficulties in assessing the consequences of strongly idealised theories do not apply to his view, because at least some of its *negative* consequences – in terms of what it *precludes* people from saying about justice – are comparatively clear. Compare §2.5 in the next Chapter here.

relevant to these constructivisms. This is because when constructivists endorse criteria for the constitution of objectivity that are wildly counterfactual, it becomes very hard to assess whether or not they will conflict in practice with culture-laden realism in the way that actual-consensus views must do. Regardless, I would not have space to argue meticulously that particular hypothetical-constructivist approaches would have this paralytic effect here.

However, the arguments that I will give – which make up most of the next Chapter – will be a mixture between direct reflection on how we should respond to situations of substantive disagreement, and meta-level argumentation<sup>69</sup> about the underlying basis for different methodological approaches. This cocktail is intended to remove much of the motivation for *any* constructivist approach in this area, and in this way, more forcefully motivate the alternative I offer. The following two chapters will continue this admixture of first-order and meta-level critique, although first-order considerations will be prevalent in Chapter 3, and meta-level ones dominant in Chapter 4. To clarify, the thought will ultimately be that no constructivist view is more promising than the culture-laden realist alternative I begin to present in Chapter 4, for a number of reasons set out in the next two Chapters. Firstly, relatively simplistic proceduralist constructivisms are very likely to be extensionally inadequate, because they face a general 'problem of adaption'. This argument is made throughout Chapter 3, §1, with reference to H.E. Baber's sophisticated form of preferentism.

The problem with more complicated, and more clearly value-laden, actual-process

---

<sup>69</sup> I generally present the disputes at issue here as resting, at least in large part, on a disagreement about *objectivity itself*, rather than just on less radical disagreement about how many (and which) people we ought to count as peers of our epistemic community – as those whose judgements can be expected to possess such objectivity, such that they count as evidence for us, or more broadly as those whose testimonial relationship to us is otherwise constitutive of our objectivity. Even if one accepts that not all culturally-located points of view on normative issues need to be given epistemic weight – that is, even if one is open to a culture-laden methodology – it will still be an open question which points of view have such epistemic standing and which do not. Disputes of this latter kind, then, cannot be fully resolved by meta-level argument, and I do not provide any detailed thoughts about this in the Thesis, although my conclusion (in Ch.4) is that thought about this first-order matter should link peerage to virtue.

constructivist methodologies is that they cannot make proper sense of the feedback between theses about methodology and those expressing evaluations/prescriptions without *ad hoc* gerrymandering or extensional inadequacy. This argument is made primarily in Chapter 3, §§2.3-4, but, as with the other criticisms made, its most basic form is given by David Enoch's 'Why Idealise?' challenge. Because of the vastly different implementations which are compatible with the basic strategy of hypothetical-process constructivism, it is relatively difficult to identify an ideally general flaw. In the absence of that, in Chapter 3 §2.5, I discuss an extremely influential example – Rawls' 'political liberal' constructivism – in some depth, and give reasons to think that it, in particular, does not provide a promising option for my purposes. Finally, in Chapter 4 §§1.4-5, I endorse Enoch's fully general, and deep problem for constructivist views, as I am construing them here, and use this diagnosis to finally clarify the ultimate difference between constructivism and the culture-laden realism that I prefer.

The structure of the rest of this section will be as follows. In §2.2, I provide an overall sketch of Nussbaum's position on normative methodology and the justification of capability approaches. In §2.3, I outline the problems for Nussbaum in this area. In §2.4 and §2.5, I sharpen this critique using two strands of criticism taken from a paper by Alison Jaggar (2006) as paradigms. Finally, in §2.6, I arrive at an overall assessment of the respects in which Nussbaum's view is implausible, incomplete, or incoherent here, and point the way to a solution, a more detailed construction of which will be the topic of Chapters 3 and 4.

## §2.2 Nussbaum's Methodology

### §2.2.1 The Early Approach: Human Nature and Evaluation

Nussbaum's early work<sup>70</sup> on the capabilities usually involves the exegesis, adaptation, and promotion of ideas from Aristotle, and her approach to normative methodology is no exception. In this subsection I shall describe her early view about how the capabilities approach can be justified (as always) for a global context. The core resource here is her 'Non-Relative Virtues' (1993), but it will be necessary to place this paper in a broader context.

I'll begin, then, by reconstructing the central argument of 'Non-Relative Virtues'. Officially, Nussbaum's target in this text is cultural relativism, the view that normative perspectives, and in particular conceptions of the good, can only ever have validity relative to some cultural context. However, in arguing against that view, she provides the most detailed (on the meta-level at least) account of a justification for capability theory in her early work. She begins with the claim that, for Aristotle:

[Some particular account of the human good] is supposed to be objective in the sense that it is justifiable by reference to reasons that do not derive merely from local traditions or practices, but rather from features of humanness that lie beneath all local traditions and are there to be seen whether or not they are in fact recognised in local traditions. (243)

This is a clearly realist and universalist conception. She next introduces the notion of 'spheres' within human experience – a 'sphere' being an area of an experienced life that can be seen in all lives that can count as 'human'. Then she offers the concept of a 'thin' virtue, this being “whatever being stably disposed to act appropriately in [a given sphere] consists in”. Each cultural grouping will have (probably numerous) different 'thick' conceptions of virtue, which fill out this formula with substantive content. These will compete with one another, and only one can be correct (245).

---

70 Nussbaum (2006b) presents a time line of her work, divided into four stages. Roughly, the first phase is between 1987, when her first work on the capabilities approach appears, and 1997, when she begins to introduce specifically Rawlsian 'political liberal' ideas into her work. (1316). When I refer to her early work, this is the period I have in mind; the other three phases all count as 'late'.

Given the universality of the 'spheres', it then looks like we can isolate an equally universal cross-cultural debate about the virtues that takes each single sphere as its subject, even if the participants in this debate have no idea that there are other participants who disagree with them, and even if the thick conceptions that the participants endorse are radically divergent. This position is made available by externalism about the semantic values of certain terms, which Nussbaum endorses here, claiming, indeed, that Aristotle was an externalist too (247). Because “[t]he reference of the virtue term in each case is fixed by the sphere of experience” (247) that appears in its 'thin' definition, there will be one correct answer to the question of what each virtue consists in. Notably, this looks rather like the way that science is (often, on realist models) supposed to proceed<sup>71</sup>. Thus,

We can understand progress in ethics, like progress in science, to be progress in finding the correct fuller conception of a virtue, isolated by its thin or nominal definition. (248)

How would this function as a grounding for capability theory? The immediate problem is that, although this may be quite all right as an account of virtue – of how people ought to behave – what we are looking for is an account of the human good, and a view of global justice that can cohere with that. However, one of the 'spheres' that Nussbaum thinks is universal, a part of every human experience, is 'justice', cashed out as 'the distribution of limited resources' (246). Presumably then, the idea is that some particular capability theory will be (at least a significant part of) the correct account of this particular virtue. Whether or not this methodology will support a given view of the capabilities will ultimately rest on which conception of justice 'wins' empirically here; Nussbaum's contention (or, as now seems appropriate, *hypothesis*) is obviously that a capability-theoretic view like hers will result.

---

71 Cf. for example Boyd 1988.

Unfortunately, there is a problem lurking. What justifies the selection of 'spheres' that Nussbaum makes (largely following Aristotle)? Secondly, what meta-level status do *those* judgements – judgements about which spheres of human life are relevantly *essential* to humanity – have? In this way, Nussbaum's project in this paper looks as though it just pushes back the question about the justification of the capabilities to a lower level. This is the level of essentialism about the human: of reflection on what is essential to human life. The core of Nussbaum's view is the claim that in thinking about the human good – that is, in addressing the kind of questions that are necessary to give content to capability theory – we are engaged in reflection on what it is to be human at all<sup>72</sup>. *Prima facie*, this might be a thoroughly scientific, indeed potentially a reductionist, conception on a metaphysical level. After all, the best (perhaps the only) way to discover the characteristic or essential features of the species *homo sapiens* is to engage in biology, perhaps aided by the social sciences. The link to evolutionary selection that biology's relevance provides might even provide the resources for a reduction of morality to *that-which-would-be-selected-for-in-the-relevant-environment*, or something of that kind.

Fortunately however, Nussbaum makes explicit that this is not what she intends. Rather, the link to reflection about humanity itself is supposed to work in *both* directions: as well as revealing that we can think about ethics in terms of what makes a life fully human, it reveals that reflection on what makes a life fully human is itself ethical. The account of what it is to function humanly is provided by an 'ethical investigation of a certain sort' (1987, 2). It is a matter of deciding 'which functions...are so important, so central, that their absence will mean the absence of a human being or a human life' (1995, 94)<sup>73</sup>.

72 This looks like a paradigmatic instance of neo-Aristotelian naturalism, which I examine in detail in Chapter 5.

73 Although she notes that, strictly, what makes a being human is a matter of what they *could* do (and be), if

This is an attractive way of justifying the capabilities, at least in the abstract. It begins with a conception of the human good, moves from that to an image of justice, and is capable of making sense of why, when we turn to justice itself, we should be interested in people's well-being again: it would be a strange kind of justice that was an essential part of the good life, but did not clearly make its bearers better off! Not only does this provide a method for thinking about normative issues, then, it provides a method that is well placed to justify a capability approach. In addition, this appears to be a realist methodology, and – because of its explicitly ethical basis – plausibly a culture-laden one, in the sense I gave above.

Worryingly for my project, though, even here there are signs that this realism is less than full-fledged. For example, the point of this method is to argue against cultural relativism. It does so on the basis that there are universal features of human lives. But, officially at least (see §2.3.3 below), these features are not external to human tradition *per se* - “there is no 'innocent eye', no way of seeing the world that is entirely neutral and free of cultural shaping” (260). As a result “[t]here is no Archimedean point here, no pure access to unsullied 'nature' – even, here, human nature – as it is, in and of itself” (265). To claim that there are universal features of human lives can only then be to claim that human lives *as experienced* have universal features. This is not constructivism, exactly – it does not hold that the facts about what human beings are (if they are, universally, anything at all) are *constructed* from subjective states of individuals. But it does come quite close; in particular, it suggests that were something *not* to form a part of an enculturated individual's self-conception, it could not be a potential part of their lives, something that they *should* pursue. This is a clear violation of culture-laden

---

only the conditions were right, and not, as this sentence suggests, what they actually *do* do. (Nussbaum 1987, 20) Compare my view in Chapter 5, §1.2, on this point.

realism, if not of realism *simpliciter*.

### §2.2.2 The Late Approach: Informed Desire and Substantive Good

The most important alteration in Nussbaum's work on the capabilities has come from her explicit acceptance of a Rawlsian 'political liberal' conception of the basis for political norms. In §1.3 above, I provided some reasons to move away from a political-liberal framework for a capability approach. Insofar as Nussbaum's view implicates such ideas, then – and (*a fortiori*) insofar as she may be committed to some level of constructivism as a result of them – it will be unsatisfactory. However, there is more to say, even if this 'political liberalism' is ignored<sup>74</sup>. This is made easier by the fact that Nussbaum's (2000a) work *Women and Human Development* contains significant discussion of methodology, most of which makes little or no reference to 'political liberalism'. Most importantly, the role that she gives there to 'procedural' methodologies is a step away from the clearer realism of her earlier work, and a step towards constructivism. It is this apparent combination of realist and constructivist elements in Nussbaum's recent work that is of particular significance for understanding the problems her methodology highlights.

The discussion in *Women and Human Development* is centred around a distinction between 'procedural' and 'substantive' methodologies<sup>75</sup>. Essentially, the difference is that 'procedural'

---

74 Notably, some have accused Nussbaum, quite reasonably, of just endorsing 'political-liberal' ideas without really making the changes to her theory that these would require. Most obviously, if you want to ensure that your prescriptions will be compatible with all 'reasonable' conceptions present in the global *polis*, surely the best way to do this is to follow some largely constructivist method, building only from premises that you *know* are universally endorsed. Without being too cynical, I cannot help but suspect that the reason that Nussbaum doesn't do this is that she knows that the output of such a process would not be friendly to her capability theory's ambitions, if it managed to generate a positive result at all (which I also doubt). Compare the discussion of Jaggar's critiques below.

75 I discuss this further in §1 of the next Chapter. Given that I go over the application of procedural and substantive methodologies to the topic of adaptive preferences there, I shall not include it in my exposition here, although it is in that context that the distinctions and arguments discussed in this section are made.

methodology aims to provide a way of thinking about normative issues that builds in as little substantive content as possible; the purest procedural model would build in none. These methods usually aim towards a disinterested and impartial viewpoint that is able to take on board as many different contributions as possible. 'Substantive' methodologies, by contrast, insist that we must begin with a significant amount of normative content already in place, and then work from there to discover answers to the rest of the questions that arise. As a result, substantive approaches will be characterised by a relatively *direct* focus on discussion about normative subject-matters (such as the nature of the human good), because once conclusions about the good are established these have a double significance: they will inform our future methodology<sup>76</sup>, as well as being important in their own right<sup>77</sup>.

Paradigmatically, procedural approaches work by stipulating some form of relevant data; the two most relevant being intuitions about the good, and preferences<sup>78</sup>, and then proceed by aggregating that data to arrive at the correct result. So, the policy that most people *prefer* or *think good* is the one that we should say *is good*, for example<sup>79</sup>. The relationship between proceduralism and constructivism should be clear: it will not make much sense to be a *pure* proceduralist, at least, unless one is a constructivist. At least *prima facie*, it is only if objectivity in beliefs about the good is a construction from (e.g.) preferences that the correct epistemic method would involve the aggregation of (e.g.) preferences alone. As one allows

---

76 See especially §2.4.2 of the next Chapter.

77 Cf. also Nussbaum 2001a, 68-71.

78 Nussbaum tends to talk about *desire* as much as preferences; I take it that preference is *relative desire*, such that to prefer x to y is to desire x more than y, and such that if there were only one conceivable option, it would make no sense to say that a person *prefers* it, although it would still make sense to say that they *desire* it. This doesn't appear to be how Nussbaum uses the word, however; cf. 2000, pp. 119-121.

79 There is a potential difference between this (meta-level) attribution of epistemic significance to preferences, and the first-order significance that *preferentists* such as H.E. Baber – for which see §1.2 of the next Chapter – propose. As I say in fn.1 of the next Chapter, one *could* be a first-order preferentist without believing that a preference-based proceduralism was correct, although it would make little sense to do so. Nussbaum's discussion in *Women and Human Development* doesn't distinguish between these two levels, which makes things more difficult for me throughout.

more and more substantive content to enter into one's methodology, constructivism becomes less and less obviously appropriate.

Nussbaum, however, doesn't focus on such *pure* proceduralism, for two good reasons. Firstly, pure proceduralism isn't well-motivated unless we assume constructivism, and Nussbaum generally seems unwilling to do that. We have certain beliefs about how people ought to be treated that are very basic, and we are not willing to give these up; on the other hand, if we came to think that objectivity *necessitated* that we put these beliefs to one side, we would be willing to do so. For example, feminists (at least) can agree that women should, fundamentally, be treated equally with men. If we are faced with someone (or perhaps many people) who (explicitly or implicitly) deny this, and would prefer women to be treated worse than men, we will not want to include their preferences as evidence that we (objectively speaking) ought to treat women unequally. If it is necessarily the case that the objective standpoint is a construction from preferences, however, we will *have* to include these preferences too (and just hope that they will be outweighed in the final analysis).<sup>80</sup> If this is right, all feminists should hope that constructivism (at least of a preference-based kind) is false. Secondly, very few serious people, upon inspection, really *are* pure proceduralists. Relevantly, Nussbaum discusses two broadly proceduralist models (those of John Harsanyi and Richard Brandt) and identifies a number of respects in which they fail to be relevantly 'pure' (122-135).

It is against this background that Nussbaum sets out her preferred account of normative methodology, and, thereby, of the justification for a capability approach that will result. This

---

<sup>80</sup> See (Nussbaum 2000a, 117-9), where she makes much the same point in introducing 'subjective welfarism', as a kind of procedural approach.

methodology has two strands: the informed desire approach, a weak form of proceduralism; and the substantive good approach, which (unsurprisingly) is strongly substantive. To begin with, the latter is quite straightforward, and much of it should already be familiar from the motivation for capability theory I provided in Chapter 1. We begin with the strongest beliefs or immediate judgements that we have about the good life and global justice, bringing together a conception of human dignity, treating each person as an end with an individual good, and working up to a list of fundamental entitlements of justice, without which people will not live lives of adequate dignity. The methodology here is, essentially, wide reflective equilibrium – the movement towards an ever-more coherent set of normative beliefs, and one that is consistent with an ever-better set of relevant empirical data. Included in this is an interpersonal, cross-cultural element, through which a wider set of ways of living can be understood and evaluated. (2000a, 70-110).

On the one hand, this may seem to be a sufficient method on its own. In addition to the scepticism towards procedural alternatives that I have already mentioned, Nussbaum adds other claims that indicate the superiority of this sort of explicitly substance-laden approach. She says that '[a]ny procedural approach that we can accept will be so laden with normative values that it collapses, in effect, into a substantive good approach' (119). Also, she believes that '[o]ur intuitions about the central capabilities are at least as secure as our intuitions about what constitutes a good procedure' (150). These claims obviously tell against the tenability of a procedural alternative that would fail to securely and explicitly build in the values that her capability theory represents.

On the other hand, Nussbaum presents three reasons to add a procedural justification to the

substantive good method, and these seem to be what motivates her to incorporate such an approach, despite her worries. Firstly, there are epistemic reasons. We may well neglect to properly consider alternative ways of life if we fail to cast our net wide. Additionally, she holds that wide-ranging independent agreement with the results of the substantive good approach constitutes evidence for its truth. (151). She claims that 'we would not conclude that....items on the list were central human goods unless we heard women saying so, when questioned in a [suitably value-laden] procedure' (160). Secondly, any prescriptions we make for the global *polis* will have to be *stable*: it cannot be the case that significant numbers of people would inevitably revolt or be deeply unhappy in the world we recommend. In part, she holds this thesis for Rawlsian reasons, according to which a political order is not justified unless people actively accept it, rather than it's being a mere *modus vivendi* (152-3). However, this element does not seem essential – after all, a world filled with chaotic, if not violent insurrection, or with hatred and mistrust for central institutions, would hardly be a good one, and this alone justifies *some* kind of stability requirement<sup>81</sup>. Finally, we have a normative reason to give weight to people's desires. She says that 'desire is an intelligent part of the human being that deserves respect in itself in any procedure of justification we would design' (154).

As a result, Nussbaum proposes that the substantive good approach be supplemented by a more strictly proceduralist method that involves consulting 'informed desires'. The idea here is that we examine what desires people in general have in procedurally-idealised

---

81 If we refuse to accept 'political liberalism', and endorse the radical transformation of existing perspectives, we can say that there would be stability, but only in virtue of the fact that people's desires had comprehensively changed. The only real impact that a stability requirement would then have would be the requirement that we not endorse anything that significant numbers of people are *innately* predisposed to reject. I suspect that, empirically, this would be a near-negligible category, but see §2.3.3 below on this point.

circumstances. The informed desire methodology yields a result when the following occurs:

[P]eople from a wide variety of backgrounds, coming together in conditions conducive to reflective criticism of tradition, and free from intimidation and hierarchy...agree that [some conception] is a good one, one that they would choose...When people are respected as equals, and free from intimidation, and able to learn about the world, and secure against desperation, their judgements about the core of a political conception are likely to be more reliable. (151-2).

Nussbaum believes that the results of a properly-conducted informed desire proceduralism such as this will converge over time with the result of the substantive good method. When this occurs, the justification for the capability approach will be even more robust than the substantive good view alone, and its objective soundness will be clear. (155-66).

## §2.3 Sketch of the Problem

### §2.3.1 Unclarities, Inconsistencies, and Lacunae

Now that Nussbaum's conception of the justification for her capability theory is in view, I can begin to criticise it. To a large extent, the motivation for my project here is quite simple. Nussbaum's account of the background to her capability theory is incomplete, inconsistent, or unclear in a number of dimensions, and these are worth remedying for their own sake. This motivation operates apart from any more pressing problems – e.g. with the practical ability to offer good prescriptions – that her theory may possess. This motive directs particular attention to the meta-level aspects of Nussbaum's theory; this is an area where I perceive a need for substantial revision, and my aim is to provide such revision, largely in the next three chapters. It should be sufficient to motivate an investigation of this kind to note that we have to have *some* meta-level view, and therefore that a *complete* specification of a capability theory will have to include meta-level theses as well as first-order ones.

I have operated, and will continue to operate throughout the Thesis, with a quite broad

interpretation of what is included in the meta-level. It certainly goes beyond the core questions contained within the dominant post-Moore line of meta-ethical inquiry, which focuses on debates between various species of cognitivism/non-cognitivism, realism/anti-realism, and naturalism/non-naturalism. Indeed, I shall not have a great deal to say about those general topics here, although as I have said, a truly comprehensive capability approach would have to do so. Instead, the meta-level theses that I shall be most concerned with have to do with differing images of normative practice, in a much looser sense than is generally standard in contemporary meta-ethics. Central issues with which I *shall* be concerned include the interconnected questions of how we can think about objectivity; how we should respond to diversity of normative attitudes; and what kinds of methods we should use in thinking about the good life and justice. With this in mind, in the rest of this subsection (§§2.3.2-2.3.3), I shall proceed to bring out two problems within Nussbaum's account that appear to ultimately have their source in her approach to meta-level questions. Firstly, I identify some points at which she advocates deference to procedures like local specification without any clear rationale in terms of the substantive good approach. Secondly, I describe an area of tension within Nussbaum's work which has to do with the need for methodological convergence that she posits.<sup>82</sup>

### **§2.3.2 Plural and Local Specification: Bounded Constructivism? Or An Acknowledgement of Epistemic Limits?**

Nussbaum's work on the capabilities approach is characterised by an occasional tendency to leave aside the detailed answering of questions to which the approach gives rise. Of course, a

---

<sup>82</sup> A third issue concerns the possibility that the absence of convergence between the two methods would give rise to the same sceptical aporia that constructivism faces almost immediately. I won't develop this here, however. See §2.3 in the next Chapter for further development of the point, applied to constructivist approaches generally.

certain reticence to seek a complete set of answers to all such questions is understandable for a number of reasons. Firstly, no one thinker could hope to deal comprehensively with every normative issue that her approach throws up in her lifetime, let alone in single articles or books. Secondly, there are many questions for which answers cannot be given because we do not have adequate evidence. We just do not know exactly how the global *polis* should be ordered, to take a significant example, because our understanding of economics, sociology, political science, etc., is obviously far from adequate. Giving answers to normative questions at the global level, at the level of precision that we might one day hope to achieve<sup>83</sup>, will then not be possible today. Alternately, we may know that we will *never* be in a position to justifiably include the answers to *certain* questions in our theory, because the evidence that would be required to justify those answers is bound to particular, fine-grained contexts in a way that prevents us from accessing it in advance<sup>84</sup>.

However, Nussbaum's typical response, when faced with a large number of issues for which such exculpations might be available, is instead to appeal to what she calls 'plural' and/or 'local' specification<sup>85</sup> (1990a, 235-8). Plural specification occurs whenever space is left for individuals to choose ways of living that are not specifically endorsed by the conception of the good that the theorist offers, but which remain within bounds set by that image. Local specification allows for a diversity of good ways of living that are linked to the context in which the relevant individual finds herself, but which are similarly restricted by the general, universal image of the good. The difference between the two seems just to be that the latter

---

83 Disregarding of course the fact that as our knowledge of the way things are grows, they change, ensuring that fully adequate knowledge is likely to remain a tantalizing prospect.

84 This being what is involved in Nussbaum's (relatively modest) particularism. See her 2001b, 290-317, esp. 298-306. Cf. §1.4 in the next Chapter.

85 Her use of this terminology seems to have lapsed in more recent work, but the same idea seems to be at work in her later work too.

involves reference to contexts<sup>86</sup> in which *many* individuals will find themselves (paradigmatically, cultural settings, and perhaps extending to things like disabilities), whereas the first leaves open a space for specific paths that might be unique to an individual. Both these theoretical devices function to restrict the degree of normative prescription that can take place in advance of individuals' and (perhaps) groups' choices, and so, to the extent that they can justifiably be employed, they will excuse the thinker from providing a complete theory in the same way as the two reasons I gave in the previous paragraph. But what kind of rationale is at work here?

There are several options. On the one hand, Nussbaum's use of these ideas could just be an application of one or the other of the two reasons I noted above; she might merely be acknowledging the limits of what she can be expected to do given constraints of time and available evidence. This would be well justified for a wide variety of cases, and would be compatible with a culture-laden realism of the kind I propose later. On the other hand, it might be motivated by a tacit, if partial, acceptance of some meta-level thesis that would contradict the realism that she elsewhere seems to accept.

This possibility is made more plausible by some of Nussbaum's language when talking about plural and local specification, and by the role that this kind of indeterminacy is supposed to play in making it more likely that the procedural and substantive methods will converge. Most critical here is Nussbaum's use of limited specification as a means to argue that her capability theory is not paternalistic; that it does not force people to live in a certain way for their own good. She is particularly concerned that her view not seem like “one that simply

---

<sup>86</sup> Bear in mind that 'contexts', as I'm thinking of them here, include facts about the constitution of individuals (e.g. disability status or personal history) as well as of their (e.g. social) environment.

tells distant people what wisdom requires”. Instead, we should always seek to “imagine a possible transition from the current ways to some specification of the vague conception [of the capabilities] that fits that history and those circumstances”. (1990a, 236-7). Given the level of particularity this would require, it is not surprising that limited (in this case, local) specification is recommended. Quite plausibly, such tailoring of prescriptions to particularity would often be warranted by a robustly realist view. But it seems that the only rationale for local specification that could *guarantee* non-paternalism, and refuse to *ever* engage in prescription that is not rooted in the addressees’ conceptions of the good, would be constructivist<sup>87</sup>. Surely, no realist can rule out that *some* people's norms might be too far from the truth to get to it without radical transformation? If so, realists cannot rule out the possibility that paternalistic interference might be warranted, unless they happen to already know that paternalism is always wrong<sup>88</sup>. Nussbaum's lack of a clear account of precisely when and why paternalism constitutes an objection to a claim about someone's good may well be rooted in her unclarity about basic meta-level issues, and this warrants an investigation into those issues directly.

### §2.3.3 What explains the Convergence?

A third issue concerns *why* we should expect convergence between substantive and procedural methods over time. This is not a trivial or obvious hypothesis, and if true it requires some explanation. The most straightforward explanation concerns human nature. If humans have a universal nature that causes them to desire certain things regardless of cultural

---

87 Paternalism, after all, is the bypassing of an individual's will for their own good – a view like Nussbaum's can only fail to be paternalistic to *some* extent if it *always* concurs with what people will, or would will, for themselves. Compare the introduction to the next Chapter.

88 Again, it is hard to distinguish first-order and meta-level species of constructivism/proceduralism here. Arguably, a person could technically be a culture-laden realist and yet still hold that all claims that cannot be procedurally derived from a persons perspective will be illegitimate, on the first order ground that this constitutes an impermissible paternalism.

setting, this would provide one of the ingredients necessary for convergence to occur.

Nussbaum seems to espouse such a view. For example, she claims that:

the human personality has a structure that is at least to some extent independent of culture, powerfully though culture shapes it at every stage...[some desires] seem to be relatively permanent features of our makeup as humans, which culture can blunt, but cannot altogether remove. (2000a, 155).

If we combine this with the strong epistemic role that human desire is supposed to play, this suggests that what is most fundamentally at work here is a non-culture-laden realism. The explanation for why people come to converge in their moral views over time might be that all humans, regardless of cultural location, will, when in the relevant environment, come under the causal influence of the mind-independent good, and will come to desire the right things as a result. As well as explaining why human desire, regardless of enculturation, will be good evidence, this account will explain why the substantive and procedural methods would converge<sup>89</sup>. If Nussbaum's substantive reflection provides her with access to certain truths about the good, and the consultation of others desires will tend to do the same in virtue of their natural orientation, the outputs of the two processes would converge over time.

However, as will become clear in the next Chapter, we have good reasons to seek a culture-laden realism. Insofar as the interpretation I have given here is correct, and insofar as that restricts Nussbaum to non-culture-laden forms of realism, it will conflict with the position I put forward later. Resisting Nussbaum's view will involve denying that desire can have the intrinsically evidential role that she ultimately seems to advocate, and holding that we should not expect desire-based proceduralist methodology to be particularly accurate. This is of wider significance, insofar as the prospects of a universalistic capability approach will depend

---

<sup>89</sup> Compare Sturgeon 2006; Enoch 2011, Ch.7. Arguably, the sort of empirical, explanatory considerations that are brought in by notions of convergence are likely to militate against culture-laden realism, at least when read as requiring explanations to be given in scientific terms: as we shall see in the following chapters (esp. Ch.5), to make sense of a plausible Aristotelian culture-laden realism we need to deploy autonomously ethical concepts on every level of our explanations.

on whether it can justify divergences from local (including individual) beliefs and desires. What is more, in the next Chapter I shall give several reasons for scepticism about ability of constructivisms to sustain themselves when faced with certain kinds of epistemic difficulty.

## §2.4 Jaggar's Critique

### §2.4.1 Jaggar's Epistemic Critique

So far, I have outlined Nussbaum's views on normative methodology and the justification of her capability approach, and provided a variety of reasons to be skeptical that this is, in the end, a coherent and plausible account. In this final section, I turn to one of the most comprehensive criticisms of Nussbaum in this area, Alison Jaggar's 2006 paper "Reasoning about Well-Being: Nussbaum's Methods of Justifying the Capabilities". Her critique is essential here because it incorporates and continues in spirit a number of others that came before, especially those of Okin (2003) and Ackerly (2000); and because it raises in a quite general way a number of distinct criticisms of Nussbaum, some of which I find compelling, and some of which are not cogent, but are mistaken in a way that points towards clarifications and revisions of Nussbaum's meta-level approach.

Many of Jaggar's criticisms are of a broadly *epistemic* kind: they aim to identify features of Nussbaum's way of thinking about justification that make it unlikely that it will result in objectivity. In this sub-section, I will describe Jaggar's criticisms that seem to be of this broad kind. Through both this section and the next, I will not assess the cogency of the criticisms I survey, leaving this task until §2.5 below.

Firstly, Jaggar argues that Nussbaum's claim that there is a good chance that a procedural

informed desire approach would, in fact, result in backing for the capabilities approach is unsupported, at least on the evidence she provides. She notes that Nussbaum has had very few direct and reasonably comprehensive conversations with people outside of academic or NGO circles (perhaps only two!) who come from non-Western backgrounds. Also, when large-scale conversations such as conferences have been organised, they tend to include mainly Westerners, and there is little evidence that existing inequalities in power have been compensated for. Finally, the limited evidence of other's views that Nussbaum provides is almost always filtered through her interpretations, and it is plausible that others might not interpret what has been said as supporting the capabilities approach over alternative proposals. (313-5).

Secondly, many of these same criticisms, especially those that trade on a general concern about the non-inclusion of all perspectives, also apply to the substantive-good approach. Given that this is her primary approach, relative to which the proceduralist method is merely a supplement, and given that it is something like the substantive good approach that I want to defend, it is this to which I give most attention.

Consider the following two quotes:

Establishing that a set of intuitions or preferences is in balance shows only that the set is internally coherent. Unless the individual or group creating the set has considered all available intuitions, it remains possible that anomalous intuitions may have been excluded or that other equilibria, expressing alternative moral perspectives, may exist. Since the reflective equilibrium approach offers no guidelines for preferring one of these possible alternative equilibria to the others, it does not dispel the specters of subjectivism, if the coherent set of intuitions is held by an individual, or of relativism, if the set is held by a community...In order to address such concerns, Nussbaum must explain why the intuitions of those who accept the values on her list of capabilities are more reliable than the intuitions of those who do not. (315-6).

Not only does [this approach] provide no safeguard against philosophers selecting only intuitions that are in agreement with their own, while dismissing dissenting views; it explicitly recommends that philosophers should select only intuitions with which they agree. Thus rather than providing independent confirmation of the political conception of the capabilities, it simply begs the question...Insofar as the...substantive-good approach lacks both a requirement that philosophers check their own intuitions against those of non-arbitrarily selected others and a procedure to ensure that this checking is done fairly, it is unable to assure philosophers that they are not projecting their own ideas on to other people and simply rationalizing their own pre-existing values. (317).

There are a number of epistemic criticisms here, many of which seem compelling. Firstly, Jagger notes that an *individual*, or *restricted group* method of reflective equilibrium does not seem satisfactory, because there may be more than one coherent, stable position, and it will be difficult to rule this out without unrestricted discussion. Secondly, if there are multiple equilibria, reflective equilibrium *per se* offers no reason to select one over another<sup>90</sup>. Thirdly, any justification that this method provides for a normative conception cannot be *independent*. And finally, if the process is very individualistic, there will be nothing in place to prevent misinterpretation of others and the invalid appropriation of their responses as evidence for theses they do not support.

#### §2.4.2 Jagger's Normative Critique

In addition to the above epistemic criticisms, Jagger also argues that Nussbaum's substantive good approach is *normatively* problematic: that it treats people unjustly. As before, I will quote the key passages, and draw out the criticisms afterwards:

The non-platonist substantive-good approach is problematic in several overlapping respects. It runs the risk of exclusiveness because it fails to mandate that everyone should participate in developing the list of capabilities; instead, an unidentified 'we' draw on the ideas of other vaguely identified 'people.' In addition, it is explicitly non-egalitarian: some people, the same unidentified 'we,' assume the authority to decide whether or not those people's desires are

---

90 It is not clear why the absence of a procedure that guarantees the establishment of a single correct result should lead to subjectivism or relativism, however, and Jagger does not expand on this here. Even a cursory glance at the history of epistemology should indicate that such guaranteed certainty is at least extremely rare, if not completely absent.

'informed' or 'corrupt' or 'mistaken.' Indeed, even when the desires of the others are accepted as informed, they are given only limited weight. Nussbaum writes that the 'informed desires' of other people play 'a large role in finding a good substantive list,' but the role she assigns to them is primarily that of assisting in *discovering* the capabilities; their role in *justifying* them is merely 'limited and ancillary'. (318)

If the desires of members of an oppressed group are incompatible with the capabilities, the non-platonist substantive good approach recommends dismissing them as “adaptive preferences” without taking seriously the possibility that “our” own preferences might be corrupt or mistaken. (319)

In addition, the approach asserts that people whose desires are judged to be deformed (shown by their failure to endorse “our” values) are not yet ready for proceduralism, moral epistemology’s version of liberal democracy. Only if they agree with “our” intuitions, are they allowed moral self-government. A picture emerges of morally enlightened philosophers, equipped with prior insight into the universal values, roving across the moral resources of the global South in search of whatever will enhance “our” theories. If this reading is correct, then the non-platonist substantive-good approach is not only an illiberal approach to moral justification; in many global contexts, its use by Western philosophers is likely also to be neocolonial. (320)

Firstly, Jaggar notes that the substantive good approach runs the risk of being 'exclusivist', and that 'we' (that is, members of a group that includes Nussbaum and is centred around her own normative views) make decisions about which perspectives are worthy of inclusion. Second, other people are used as sources of information about ways of living and differing evaluations of ways of living; what these people think is not a *justificatory* input though (at least, not just *per se*). Third, Nussbaum is willing to claim that other people's preferences are unsound, but does not place her own in play in the same way. Fourth, Nussbaum disrespects the self-government of those who she treats as incompetent, and this is illiberal. Finally, this disrespect smacks of colonialism, because as a matter of fact, many of those whose views or preferences will be thus treated as corrupt or mistaken will be from developing countries where there is a history of colonisation, and Nussbaum is a white American. Jaggar's earlier reference to 'oppressed groups' suggests that this colonialism claim is really only a special

case of a broader accusation, and the substantive good method will be problematic whenever the intuitions of someone from an oppressed group are said to be unreliable, in contrast to those of someone of a different background.<sup>91</sup>

## §2.5 What is Required in Response?

So, what impact do these criticisms have, and what changes should Nussbaum make in response to them? The epistemic critique has some cogency. If we take the procedural approach seriously, as Jaggar prefers, she seems right that we have few epistemic reasons not to include a broader range of participants than Nussbaum's engagement with others typically has, and to use those participants' contributions in a less thickly-interpreted way. Some of these criticisms are well-taken in relation to the substantive good method too – there too, we should be careful not to misinterpret the contributions of others, and we have good reason to increase the diversity of interlocutors if we *can*. Ultimately, though, most of the criticisms here raise further questions, at least if what is at work in the substantive good approach is a culture-laden realism. Jaggar's argument seems to require that *every* point of view be given equal evidential weight, but for culture-laden realists, views that originate from *some* processes of enculturation may be just plain unreliable. Responses that were the result of such enculturation would not bear even an approximate relationship to the truth. Far from making the process more reliably objective, then, they would be likely to at best make it more confused than it needs to be, and at worst lead to entirely false conclusions. Reflective equilibrium is a path-dependent method: where you end up depends on where you start. *Contra* constructivism, it is vital to restrict the responses that are given epistemic significance to those that seem to have a good chance of being accurate. As realists, we should also be

---

91 Notably, much of what Jaggar says relies on the concept of 'assuming authority'. I return to this in the next Chapter, §1.3.

sceptical of the possibility of multiple equilibria: reflective equilibrium is supposed to lead us to the truth, and, unless relativism is true, the truth will not be plural. I shall return to these issues in more depth in the next Chapter (§2), but this provides some motivation to articulate a culture-laden realist account, as I begin to do there.

In comparison, it is even more clear that the normative critique draws on a constructivist metaethic. The core of the normative critique is the idea that we treat people unjustly if we do not take all of their responses as evidence. The most plausible explanation for the idea that all responses must be treated equally, regardless of the cultural background that shapes them, is the idea that all responses *are* fundamentally equal. Although non-culture-laden forms of realism might endorse that<sup>92</sup>, this idea is most likely to be a result of an assumed constructivist meta-level theory. If the target of such criticisms is a culture-laden realism, they will not amount to arguments so much as blank rejections. Culture-laden realisms hold that some cultural positions may just be irreparably corrupt, and may not contain sufficient resources for constructive rational argument to a good end<sup>93</sup>. For these perspectives, then, the idea that *every* point of view must, for the sake of justice, be represented on its own terms is entirely unmotivated. I shall say more about this in the next Chapter (§2.1), where I argue more thoroughly that normative critiques such as those of Jaggar are largely question-begging, flowing from an assumed constructivism that fails to engage with the possibility of culture-laden realism, and that they cannot properly account for the practical necessity of epistemological 'fixed points', such as the fundamental equality of women, and the badness of adaptive preferences.

---

92 Although, if they do, this will most likely be because they do not take *responses* to ever be of epistemic significance as such.

93 Compare especially the discussion of reasons externalism in Ch.4 §2.2 on this point.

Although I shall not be returning to consider Nussbaum's meta-level approach in detail again, my reasoning for my ultimate conclusion regarding it is not yet in sight – much of that takes place in the next chapter. However, from what I have said so far, it ought to be possible to make out at least vaguely what my conclusion must be. I conclude that Nussbaum's approach is torn, when faced with the problems I have set out, and especially in the light of criticisms like Jaggar's. One horn of this putative dilemma is a culture-laden form of realism. I shall show that this approach has good prospects for filling in the gaps in her theory, and responding to constructivistic criticisms. However, it does not (*pace* Nussbaum's optimism about the prospects for overlapping consensus) build in the thought that people regardless of cultural position will tend to desire the good, and does not leave room for the Rawlsian model of political justification that she now prefers. The other horn preserves those elements of her view, but forces her – in the face of Jaggar's critique – to abandon the substantive good method and leave the capability approach at the mercy of a proceduralism whose results are necessarily far out of sight, and which is unlikely to be able to coherently sustain its own egalitarian motivations. My belief is that the correct response to this dilemma is plain: we should choose culture-laden realism. Hopefully, the argument of the next chapter will make that belief compelling.

## Chapter 3:

# Justification, Objectivity, and the Human Good

In this Chapter, following on from my conclusions in Chapter 2, my primary aim is to account in more detail for the division between ways of looking at objectivity that I have identified, and to give reasons to favour one over the other. In large part, I shall continue to operate with the broad distinction between 'culture-laden realist' and 'constructivist' conceptions of objectivity that I have used up to this point. However, what will also come to the forefront is a connected dispute about the proper role of preferences in our normative thought. These questions are related in that some views attempt to justify normative claims with reference to preferences *alone* – these views are forms of *preferentism*. Capability theory has traditionally opposed strict preferentism on the grounds that preferences are often *adaptive*. The basic connection between preferentism and constructivism can be put simply: since preferences are subjective states of the individual will, a view that refers to those alone may well be able to claim a constructivistic kind of objectivity just *per se*. If a view of the human good is a construction from preferences, and preferences are elements of the subjective will, some sort of constructivist objectivity about that view of the good will *trivially* be available<sup>1</sup>. Preferentists, therefore, would have a powerful incentive to be constructivists. If preferentism is implausible as a justification for capability theory, meta-level constructivism *about* the justification for capability theory will be weakened also, simply because it will be harder to see how the capability approach could be objectively

---

<sup>1</sup> This is a relatively modest claim. It leaves it open that one *could* be a preferentist and yet a realist. One *could* (*prima facie*) think that there is a real, mind-independent good for humankind, but that it is exhausted by the satisfaction of preferences.

justified on a constructivist account.

In §1, then, I discuss a number of issues that bear on this question. I begin (in §1.1) by quickly introducing the phenomenon of adaptive preferences, the recognition of which typically informs much of the motivation for a capability approach, in its usual global context. I also throw initial doubt on whether a preferentist solution to the problem of adaptive preferences is possible. Next, in §1.2, I describe a recent preferentist account, by H.E. Baber, that aims both to solve this problem, and to provide a justification for capability theory along the way. I then argue against this account.

In §2, I assess in greater detail, and with reference to a greater number of sources, the constructivist critiques of the realism-based meta-level conception and methodology that I propose, and begin the process of responding to them by identifying significant problems with the actual-process constructivism that these critiques seem intended to point towards. In §2.1, I begin with a somewhat generalised version of the normative critique of Jaggar – some elements of which I will already have identified in Baber's preferentist work – which I identify as a worry about the assumption of *authority* over others. I identify two forms that this worry could take. The first (§2.1.1) concerns significant first-order normative disputes that I cannot fully deal with here, although I say a little in response. The second (§2.1.2) goes deeper, identifying an issue that, if cogent, would necessitate a turn towards deliberative-democratic constructivism. However, I argue that this critique rests on a conflation between authority and expertise. In §2.2, I turn to the epistemic critique, identifying its heart as the idea that methods like Nussbaum's introduce standpoint-based biases. §2.2.1 examines the claim that the individualism of the substantive good method is at fault, reaching a mixed

conclusion. In §2.2.2 I turn to the more specific question of how to accommodate marginalised groups within the substantive good method, arguing that a reasonable response here makes better sense within a realist framework. Next, in §2.3, I point out that constructivism, including the kind that Jaggar seems to propose, is potentially self-effacing, and thus internally unstable. In §2.4, I follow this by outlining the widely-accepted thesis of epistemological 'naturalism', and claim, once again, that this cannot be made sense of unless a realist meta-level conception is at work in the background. The section, along with the primarily negative part of the Thesis, concludes in §2.5, where I focus on Rawls' 'political constructivism', and argue that it, as an influential example of a hypothetical constructivist approach, is similarly unpromising for my purposes. Finally, in §3, I discuss more constructively some ways in which epistemological ideas friendly to the form of realism I shall endorse have been deployed in practical feminist work. Once the ground has been cleared and the motivations established for a culture-laden realism, I can begin to explicate a suitable view. This provides the task of the next Chapter.

## §1. Adaption: Realism or Constructivism?

### §1.1 Adaptive preferences

Perhaps the most important motivation for a culture-laden realist<sup>2</sup> conception of objectivity, in relation to the capability approach, results from the phenomenon of *adaptive preferences*. This phenomenon, which was brought to prominence by John Elster (1983), and elaborated in relation to her capability approach in Nussbaum (2000a, esp.136-45), occurs when individuals in circumstances of deprivation or oppression come to endorse, desire, or prefer

---

<sup>2</sup> In general, I have been emphasising a distinction between realism *simpliciter* and constructivism. However, it is important to note that at least some of my arguments in this section relate to culture-laden realism specifically, and may not be relevant to realisms that do not share that view's essential feature.

things that they otherwise would not. People's affective sensibilities – their persistent tendencies to respond emotionally to worldly stimuli – and their beliefs have been shown to systematically 'adapt' to the environments in which they find themselves, conceiving of them as 'normal' or even 'natural' parts of life. When such adaptation is present, individuals often do not have any desire for change – it is not a realistic possibility in their minds, and even if it is presented to them as a matter of immediacy, they will often refuse it. Important instances of this phenomenon include the lowered demands for education, food, and medical care that are observed in disadvantaged groups, relative to their advantaged peers. Even within relatively localised contexts, such as between regions of India, women tend to exhibit less desire for additional education, food, and medical care than men, despite the fact that in objective terms they get less of each. (2000a, 139). On any reasonably egalitarian view, this looks problematic – not only are the disadvantaged deprived of the basic goods to which they have an equal right, the environment of disadvantage even leads to an acquiescence in this inequality. In addition, this topic has importance quite apart from any concern with equality (as a component of the human good or otherwise). Adaptation causes people not to seek their good, and causes others to be unconcerned with their lack of attainment of it. For anyone who cares about the quality of people's lives at all, it must then be a significant issue.

For example, consider the case of Jayamma, as recounted by Nussbaum in *Women and Human Development* (2000a, 17-21, 64, 106-113, and 140). Jayamma was (circa 2000) a woman in her mid-60's from Kerala (in the south of India). For around 45 years, she worked as a brick-carrier for a brick kiln. She carried 500-700 bricks per day, 20 bricks at a time, stacked and balanced on a plank on her head, holding them using the strength of her neck alone. She never earned more than 5 rupees a day for this work. Because she worked in a

'cottage industry', there was no protection from a labour union. As a rule, men begin working at the kiln doing this form of labour, but they soon progress to other jobs (e.g. moulding the bricks) that require far less physical exertion but earn twice as much. No woman is ever promoted to these better jobs, and they are not allowed to learn the additional skills that would be required in any case. Jayamma has never had any opportunities for formal education, and there are few other labour opportunities available. While Jayamma's husband was alive, almost all of his wages went on tobacco, alcohol and extra food for him, while she had to support herself and their children. Despite all of these injustices, however, Jayamma is and has been “very acquiescent in a discriminatory wage structure and a discriminatory system of family income sharing. When women were paid less for heavier work...and denied chances for promotion, Jayamma didn't complain or protest. She knew that this was how things were and would be.” (113).

Recognising this as a phenomenon also raises questions about what the causes of adaption might be, and what features of an individual's environment may be implicated in the process. In particular, it seems as though it is not merely the fact of lasting deprivation itself that results in problematic adaption – not, for example, the persistent lack of food on the part of one sociological group relative to another *alone*. Rather, many cases involve people adapting to features of the surrounding *cultural* milieu in ways that look just as problematic. It is not solely the fact that women are observed to consume less food, as a statistical generalisation, that gives rise to a lack of desire for adequate nutrition on the part of other women, but also the presence of ideological elements that directly discourage such desire. People's affective sensibilities are obviously affected by enculturation, and if the surrounding culture and wider society contain ideas that lead to a lack of desire for substantive equality, this is adaption no

less than in the non-cultural case<sup>3</sup>.

The question that arises here is whether or not constructivist views, restricted as they are to the space of existing desires and preferences themselves, can acknowledge those elements of people's affective sensibilities that are problematic *as* problematic, and include some plausible explanation of *why* that particular set of responses is so. Doing this while remaining within the space of existing people's responses will be difficult, in two ways.

Firstly, the justification for realism arising from adaptive preferences can flow from considerations of simplicity. Even if it is possible to fully account for adaptivity within a constructivist theory – weeding out affective responses that are flawed without appeal to a mind-independent standard – it is clear that realist conceptions will be simpler. Incorporating all the ways that affective responses can be problematically adaptive, from the involvement of an existing inequality that lowers expectations, to the presence of an oppressive ideology, would require an extremely complex constructivist procedure. The constructivist, given that she must identify the perspective of normative objectivity with a perspective rooted in each individual's existing affective sensibility<sup>4</sup>, will need to manufacture a procedure for sorting existing preferences that eliminates those that are problematically adaptive *without* smuggling in normative claims from outside. They cannot do the latter, of course, because substantive claims originating outside of the affective sensibilities of the relevant existing subjects are illegitimate for constructivism. Realism, by contrast, can account for adaptive preferences

---

3 If, that is, there actually are any cases where *no* moulding of people's dispositions by enculturating processes is involved, which seems doubtful.

4 Strictly speaking, it is possible that some (e.g.) Kantian versions of constructivism would disavow this requirement, since the subjective base from which construction is supposed to take place is not relevantly affective, on these views. If such views can genuinely keep affective elements at arms length (which I doubt), they would not be vulnerable to the argument at this point. On the other hand, of course, there is no *prima facie* reason why such views would be concerned with preferences at all.

straightforwardly and (probably, depending on the content of the realism) quite simply. If a preference results from the wrong kind of background, whether that be a background of material inequality or of inegalitarian or otherwise false ideology, it can be rejected as unresponsive to the normative facts<sup>5</sup>. Of course, it is important to note that the considerable theoretical advantages of realism here will likely not mean that the real-world task of evaluating the normative admissibility of preferences will be remotely easy in itself. Just because realism assures us that, were we to know fully the content of the correct theory of the good, we could straightforwardly generate a procedure for assessing preferences, this does not eliminate the major source of practical difficulty, which lies in deciding on the precise content of the good itself. However, this could not be a greater problem (indeed, I shall suggest, it is a lesser one) for realism than constructivism – if the constructivist is to identify the same set of preferences as problematic that the realist does, they will have to make claims about *procedures* that are functionally identical to those that the realist makes about *content*. Any indeterminacy or uncertainty in the latter will (at least *prima facie*) be reflected in the former.

Secondly, it is doubtful whether constructivist views can motivate the rejection of all and only those preferences that are problematically adaptive at all, regardless of whether they can do so simply. Nussbaum has claimed that an adequate answer to this problem *requires* an account of the human good that does not automatically take as evidence the existing beliefs and desires of contemporary humans, even those that have been put through some procedure of idealisation (2000a, 148-9). The problem here is not merely that no proceduralist account

---

5 Of course, it is only within the context of some *particular* realist normative view that we can say what the 'wrong kind' of background is here. My point is merely that *any* realism will have some resources available to do this. A further advantage of this approach is that it dispenses with any fraudulent veneer of neutrality that might accompany alternative theories.

seems to have been offered that could do this job. That might, if defeasibly, justify a prediction that no account *will* be forthcoming. Rather, the principal problem is that reasons have to be found for why each procedural rule is included, and the pure proceduralist cannot, in giving such reasons, refer to substantive normative content, e.g. 'without this rule my procedure will generate repugnant moral views'<sup>6</sup>. Given this, it is not clear what a pure constructivist can say, particularly about cases of adaption that involve the cultural background. The particular problem here is that all beliefs/desires that are relevant to normative judgement take place against *some* background of enculturation, so the challenge is to specify what it is about particular kinds of enculturation that is problematic without referring at any point to the content of the belief/desire itself.

Taken together, these two considerations are sufficient to motivate a realist version of objectivity for capability theory. Without a theory whose normative standards can stand outside of people's existing beliefs and desires<sup>7</sup>, it looks as though adequately responding to the problems raised by the phenomenon of adaptive preferences will be (at least) highly complicated on a theoretical level, and may even be impossible. In the next section, I shall examine one recent preferentist account, that of H.E. Baber, which explicitly aims to avoid an appeal to a conception of the human good that stands outside of existing desires in the way I have urged is necessary.

---

6 See, however, §§2.2-3 below. If a constructivist concedes the necessity of 'fixed points', they can reject preferences on the ground that they contradict them. The question, as I say there, is how any constructivist could justify such fixity.

7 Arguably, this is what it all comes down to. Two broader connections are notable here. Firstly, constructivism has a revealing relationship to anti-realism, construed (as elsewhere) as a refusal to countenance any concept of objectivity that would transcend our subjectivity (e.g. our possible or actual evidence, beliefs, desires, or our phenomenology generally). Cf. Miller (2003), which cashes out a requirement for realism in terms of whether or not the facts it postulates are 'constitutively independent of human opinion' (6), and Dummett (1991) for a detailed exposition of the idea. Secondly, constructivism in the normative domain is *voluntaristic*: the truth is dependent on states of the will. This raises a connection with the Euthyphro problem, the classic statement of which is obviously Plato (1997, 1-16).

## §1.2 Baber's Trans-world Preferentism

As I have said, Baber's account is a form of preferentism: it holds that *all* preferences, and *only* preferences, are the correct ingredients in a theory of the human good. Her 2010 paper, 'Worlds, Capabilities, and Well-being' describes a preferentist view, and argues that it has the resources to solve the problem of adaptive preferences and justify a capability theory. Given that this is a fairly comprehensive attempt to ground a normative theory, there would be much to say about it. However, I shall not be concerned with issues concerning Baber's account of the foundations of capability theory<sup>8</sup> that don't relate to the cogency of a constructivist approach to adaptivity.

Baber's preferentism is distinguished by its inclusion of possible preferences, as well as actual ones, when making an assessment of an individual's good. Specifically, preferences that a subject *could easily have had* – that is, those that beings relevantly similar to them have, in close possible worlds – also count as benefiting the actual-world subject if they are satisfied. This (arguably) counts as a capability theory because one way of understanding what a capability is appeals to that which could easily happen but won't, or could easily have happened but didn't. On this account, it benefits a person to open up a broad space of possible functionings that 'centres' on their actual preferences, even though they may not (actually) prefer the possible states of affairs that are thus enabled. Classic capability-creating policies do not merely impact what happens in the actual world, e.g. by enabling an agent  $A_0$  to choose to do things she otherwise couldn't. They also impact a range of possible worlds, by

---

<sup>8</sup> For example, her definition of capability differs from the one I gave in Chapter 1. On her definition, having the capability to  $\Phi$  is not conceptually related to the capacity to effectively choose to  $\Phi$ , rather it is a function of how metaphysically close one is to a world at which one  $\Phi$ 's, regardless of whether one's choices are what distinguishes that world from our own. (Baber 2010, p.381.)

making it the case that in those possible worlds possible versions of  $A_0$  ( $A_1, A_2, A_3 \dots A_n$ ) satisfy preferences that  $A_0$  may or may not have, but could easily have had. For a set of possible worlds, which are centred on subjects  $A_0$ - $A_n$ , such that each version of  $A_0$  lives in a unique world, a given world will be *closer* to the actual world to the degree that the 'psychology and circumstances' of the relevant subject  $A_n$  resembles that of  $A_0$ . (379-380)

This enables definitions of three key concepts for Baber: relevance, fruitfulness, and capability. A possible preference is *relevant* to the well-being of  $A_0$  to the extent that the world in which that preference occurs is close to  $A_0$ 's world; that is, to the extent that the value of 'n' for that particular individual  $A_n$  is relatively low. The satisfaction of a preference held by a close possible version of  $A_0$ , such as  $A_1$ , makes  $A_0$ , substantially better off, whereas for high values of n, the difference made may be negligible. *Capability* is constituted by a relatively small distance between the world of  $A_0$  and the world in which a given state of affairs occurs.  $A_0$  only counts as having capability relative to a state of affairs (i.e. the instantiation of a 'doing' or 'being' functioning) if the world at which that state of affairs occurs is metaphysically close. Finally, states of affairs are more *fruitful* if they are compatible with the satisfaction of lots of other actual and possible relevant preferences. Finally, then, it is possible to see how this theory evaluates a given state of affairs in terms of its contribution to an individual's well-being. Having capabilities but *not* functioning is valuable, on this view, just to the extent that the capability is achieved: to the extent that the individual *could* function; to the extent that the capability is relevant, such that some close version of the individual prefers it; and to the extent that the capability is fruitful, such that having it is compatible with a wide range of other capabilities. (379-83)

This amounts to a grounding for a capability-involving theory of well-being in preferentist terms. Baber also claims that it has the resources to account for adaptive preferences. It can do this in two ways. Firstly, some preferences that people have may be *fruitless*, and if that is so, satisfying them may be of neutral or negative value overall<sup>9</sup>. Baber's example here is the preference of someone who wishes to have a healthy limb amputated (386). Secondly, it may be the case that people *would* have preferences for some things if only their circumstances were slightly different, and her account allows this fact to have normative weight. It appears that this theory will generate the right verdict about a large number of the cases that we think problematically adaptive, and will do this without building in (covertly or otherwise) any objective, preference-external theory of the good.

As yet, I have not argued that trans-world preferentism does not have good prospects for doing the work that we require for (e.g.) the formation of a conception of dignity. The best way to begin to do this, to bring out the substantive differences between Baber's view and my own, is to consider the positive case she gives for her theory, relative to non-preferentist proposals. Unfortunately, she does not offer a great deal in the way of explicit argument. However, she does make three comments that imply criticisms of the realist approach that I prefer. The first criticism can be seen in this passage<sup>10</sup>:

Of course, it is a matter of empirical conjecture whether the preference for better working conditions, higher pay and control over her earnings is a nearby preference for Jayamma or other poor women. It could be that they are so beaten down that their desire for a better life is remote. Nussbaum's stories about the response of poor women in developing countries to improved prospects however suggest otherwise. When the poor Indian women whose stories she tells see the benefits of joining

---

9 It should be noted that standard preferentism can also designate preferences as deformed 'on net' in this way; the only difference here is that trans-world preferentism factors in the possible effects of relevant possible preferences as well as actual effects of actual ones.

10 I quote all three of the critical passages at length because I think they get remarkably close to the heart of the issues that I have been considering, and in the light of that, I don't want to run the risk of misinterpreting a key opponent via a paraphrase.

women's co-ops and taking advantage of microcredit schemes they jump at the chance to participate—no elaborate therapeutic intervention or extensive personality remodeling is required. Their preferences for improved conditions are nearby and that is why we think it is worthwhile to improve the conditions of their lives. *Mutatis mutandis*, when the preference for “improved” conditions is not nearby, when it would take aggressive therapy, propaganda and extensive character remodeling to elicit or manufacture them, we question the value of life-improvement programs: we are skeptical about projects intended to reform preferences by advertisers, missionaries and political ideologues. (388-9)

The criticism present in this passage seems to be very similar to the normative strand in Jaggar's criticism (as discussed in the last Chapter). Those who believe that sometimes people are *just radically wrong* about what would be good for them are compared to 'missionaries' and 'political ideologues'. Again, it is hard to see what the criticism here really is – we may condemn missionaries or advertisers, but generally if we do so it is on the grounds that their beliefs about the good of their proselytical targets are *false*, not on the grounds that proselytism is somehow bad in itself. But if that is the content of the criticism, it is blankly question-begging – it assumes precisely what is at issue: the truth of some (preferentist) claims about the good, and the falsity of some (non-preferentist) others. The same goes for 'political ideologues', except that this has connotations of dogmatism. This, however, would be an epistemic accusation, not a normative one, in the sense I have described; the claim would be that the 'ideologue' is insufficiently open to the possibility that she is wrong.<sup>11</sup>

In the second instance, Baber presents an example:

I have come into some money—not much, but enough to pay for either breast augmentation or a course in electronics technology at my local community college. Each alternative opens a range of career options for me. Breast augmentation opens possibilities for working as a stripper, waitressing at Hooters and sex work; the electronics course opens a very different range of career options. Personally, I prefer a career in electronics and regard stripping, waitressing and whoring as bad, awful and gruesome alternatives, so I'm better off with the electronics course. But that's just me. There is nothing inherently

---

<sup>11</sup> A third possibility is that Baber is trading on concerns about alienation or paternalism here. See §2.1.1 below. Thanks to Kirk Surgener and Sarah-Louise Johnson (respectively) for emphasising these as possible motivations.

better in being a technician (or a philosopher) than in being a stripper: *de gustibus*. The current account accommodates that intuition. (386-7)<sup>12</sup>  
With the second from last sentence, we arrive at the heart of the matter. There are two points to make here.

Firstly, this is clearly a substantive claim, and a contestable one. If Baber's argument relies on the positive thesis that all ways of life, preferences aside, are equal in value, this removes it from the merely procedural kind of theorising that I oppose, and which coheres with constructivism. One of the principal attractions of such an approach is supposed to be that it remains neutral in principle between particular normative claims, providing a procedure within which people can resolve their differences on allegedly equal terms. If we have to begin with the idea that all ways of living are on a par (an idea that, on the face of it, vanishingly few accept), all perspectives will not be treated equally.

Secondly, is Baber's claim here really intuitive? Surely not – and a turn back towards our starting point in the problem of adaptive preferences should show us this. We *begin*, when looking at a typical case of adaptation, with the thought that the preferences that we consider to be adaptive – such as (potentially) the preference for breast augmentation over education – are *bad preferences to have*. The most straightforward way of explicating that response is precisely as a belief that they will *not* make the person better off. Baber's aim here then seems to be to *explain away* this response, rather than to explain it by setting it in a wider context.

---

12 There is a rhetorical turn in Baber's writing here. Presumably, she focuses on the case of stripping and breast augmentation because she expects us to be 'liberal', 'pro-sex' feminists who wouldn't dream of regarding as problematic women's choices to do these things. Whether her audience matches her expectations on this count is, however, beside the point. Her theory, as is clear from the first paragraph quoted above, entails just as equally that e.g. preferences not to eat properly in order to feed one's gluttonous and domineering husband are also just fine, so long as the psychology of the individual who has them is set up in the right way; so long, perhaps, as the person in question has been sufficiently 'beaten down'. It is far less likely that people will find *that* intuitive.

On her account, what is really at work when we recognise adaption is an implicit empirical hypothesis about the preferences that a person would have if only circumstances were different (if she is to remain true to her doctrine that all preferences are inherently equally good, she cannot say 'if circumstances were *better*'...<sup>13</sup>). We hypothesise that in most slight permutations of the person and her situation, she would prefer nutrition for herself over the indulgence of her husband. As she then points out, this is compatible with its *not* being the case that the original preference is bad *simpliciter* – it just happens to be outweighed in trans-world comparison. So, we are supposed to discard our initial intuition of the badness of the preference, in favour of this theory-motivated alternative. Even so, it remains the case that we *did* initially think that the preference was bad *simpliciter*, and so the idea that we would find it intuitive in its own right that all preferences are inherently equal is implausible on its face<sup>14</sup>.

Baber's third argument is contained here:

Arguably, *ceteris paribus*, we should prefer an account of well-being that is monistic and subjectivist to an irreducibly pluralistic objective list theory like Nussbaum's. Unified theories are to be preferred to piecemeal explanations, and objective accounts of well-being that purport to ground the requisites for the good life in the structure of Being or essential human nature seem suspiciously obscurantist. (379)

Firstly, it is not clear why it would be better, *ceteris paribus*, to have a subjectivist theory than an objectivist one; if anything, it seems to be the other way round.<sup>15</sup> Secondly, although the

---

13 There is a further criticism that would flow from this point, and which would damage the prospects for trans-world preferentism to recognise almost *any* cases of adaption, although this is less germane to my purposes in this Chapter. On the face of it, the preferences of all close versions of subject A count equally, if they are equally fruitful. Baber gets the result she wants vis-a-vis adaptive preferences by arguing that there will be close possible versions of most agents that would prefer things we think unproblematic. However, she cannot conclude (without metaphysical argument) that there would not be an equally large number of equally close possible people whose preferences are *more* problematic than those of the actual person. If so, she will not be able to condemn adaption without abandoning preferentism after all. I thank Sam Shearn for inspiring this line of argument.

14 This part of my argument is likely to generalise to most other forms of proceduralism.

15 If it is just as easy for us to justify an objective theory as a subjective one, as the *ceteris paribus* condition implies, it must be preferable to be able to claim mind-independent objectivity for it, rather than mere justification from our/my particular perspective. Probably, Baber has in mind alleged metaethical difficulties with normative objectivity here.

account I shall eventually propose (in Chapter 5) does have to do with 'essential human nature', I hope it is not obscure. Finally, although it may be that monism and unity are theoretical virtues *ceteris paribus*, I think that there are sufficient reasons to think that all things are not equal here – that Baber's view will very likely not provide us with the resources that we need after all.

That is because Baber, from what she says against the non-preferentist, must think that there are substantive differences between the set of preferences that she acknowledges are problematically adaptive, and those which she expects her opponent to find problematic. A residual substantive dispute remains, even if Baber succeeds in accounting for *some* preferences' adaptive badness, over whether or not this exhausts the category of 'adaptive preferences' entirely. Ultimately, there is not *one* 'problem of adaptive preferences' that different people are trying to solve in different ways. There are potentially as many distinct 'problems of adaptive preferences' as there are preferences that a person might think problematically adaptive. For the preferentist to stand on a level playing field with her opponent, she must first isolate some specific set of preferences whose badness must be explained. Plainly, Baber and myself (and, probably, Baber and Nussbaum) disagree about the content of that set. As a result, the meta-level dispute about methods of theoretical accounting will collapse into a first-order dispute about what is good, and what is bad. Anti-preferentists can be quite content with this state of affairs, because they want to stop there – with the idea that some preferences are bad, and others good. The satisfaction of some preferences will make their bearers better off; the satisfaction of others will be of significant disvalue, or neither valuable or disvaluable. While we remain at this substantive level of disagreement, the non-preferentist seems to have the upper hand: all they have to do is identify some set of

preferences that are bad, and try to see what connects them together to build a more general image of the bad life. They do not, as the preferentist does, have the additional task of explaining away our initial sense that some preferences are normatively irrelevant or positively perverse.

Overall, then, we seem to have a number of reasons, with respect to the phenomenon of adaptive preferences, to pursue a realist account. It does not appear that even a highly sophisticated preferentism like Baber's will be able to explain the badness of all the preferences that are problematic. The most likely explanation for this is that Baber doesn't actually think that the phenomenon of adaption is nearly as wide-ranging as non-preferentists (and, perhaps, disinterested others) likely will do. Given this, it is best to pursue a realism-congruent theory, on which the explanation of the badness of adaptive preferences is simple and direct: adaption occurs in those cases where people do not know the good, as a result of their social environment and the enculturation to which it has given rise.<sup>16</sup>

## §2. Arguing about Constructivism

I have suggested that realism provides the most straightforward, and perhaps the only, theoretical response to the problem of adaptive preferences. We need a way of thinking about normative subject-matters that makes the good independent of the space of existing human subjective states, rather than requiring that a conception of the good be built exclusively within it. I have also suggested that some of the criticisms of this realist conception (at this

---

<sup>16</sup> From the McDowellian Aristotelian point of view that I develop in the next Chapter, the only person whose preferences will not exhibit adaption at all will be the virtuous person. Her preferences will not be adaptive because they will be the result of a process of habituation, within a cultural setting, that is individuated exactly by its conduciveness to human flourishing. From this perspective, the phenomenon of adaptive preferences is entirely unsurprising – people have bad preferences because they have no exposure to the good. The next Chapter fills out the meta-level story here.

high level of generality) are unconvincing. In this section, I shall evaluate these criticisms more closely, and argue either that they are not cogent, or that they can be effectively accommodated by some realisms. As a result, in this section I aim both to further motivate a realist foundation for capability theory, and to make some progress towards the identification of what *kind* of realism is appropriate.

Before I continue, a reminder may be necessary. I began, in the previous Chapter, with Nussbaum's model of capability theory. Having noted a number of problems with it, I opted to provisionally endorse the substantive good model, construed as a form of culture-laden realism. Having already adopted Jaggar's criticism of the substantive good model as the paradigm case, I continue in this section by concentrating on constructivist criticisms that have a (broadly) feminist origin<sup>17</sup>. Feminist epistemologies are still not entirely mainstream<sup>18</sup>, and as a result, the critiques of my chosen meta-level conception may not be the most common ones that would be appealed to. However, this narrowness of focus is defensible for a number of reasons. Firstly, it is of course impossible in a project of this size to consider every criticism. Secondly, criticisms from within the feminist tradition are especially relevant to capability theory because the capability approach is, arguably at least, unusually consonant with feminist concerns. Examples would include the emphasis on effective rather than merely formal (i.e. legal) equality, the willingness to pay close attention to human diversity and difference, and (to some extent at least) the rejection of public/private distinctions. Finally, the form of culture-laden realism I shall propose has some interesting features that an introduction via feminist-constructivist criticism will highlight (see especially §3 below). The

---

17 As should become clear, 'feminist' epistemology here is not narrowly concerned with issues arising from *gender* inequality, but equally with race, class, and other structural inequalities. If this is regarded as obscurantist, readers can substitute 'critical-theoretical' for 'feminist' at most points.

18 Or, as many of the contributors to the Antony and Witt (2002) volume say, 'malestream'.

two most important of these are scepticism about (or even total erasure of): A) the dichotomy between reason and emotion<sup>19</sup>, and B) the idea that knowledge about the normative (and, potentially at least, knowledge *per se*) involves the elimination of contingent particularity from the believer's standpoint<sup>20</sup>.

## §2.1 Privileging one's own Perspective – Authority

The core of Jaggar's normative criticism in her 2006 paper concerns an alleged assumption of authority on the part of a normative theorist (in that instance, Nussbaum). I won't repeat myself too much by quoting examples (see §2.4.2 in the last Chapter), but perhaps the clearest is the following: “A central theme of my discussion has been concern about the philosopher’s moral authority. With respect to each methodological approach, I have argued either that it assigns undue authority to the philosopher or that the philosopher, in this case Nussbaum, assumes undue authority in using it.” (Jaggar 2006, 320). In this section, I identify two worries that this criticism may well trade upon, and respond to each in turn (although my response to the first worry is necessarily brief).

### §2.1.1 Paternalism and Alienation

As I noted in my previous discussion of Jaggar's critique, one of her central normative criticisms is that Nussbaum's substantive good methodology violates individuals' 'moral self-government' (Jaggar 2006, 320) by failing to give their preferences intrinsic normative or epistemic weight. If Jaggar's point was *just* to note that preferences are not given such intrinsic weight by the (culture-laden realist) substantive good conception, this would not be

---

19 See §2.1 in the next Chapter for the claims about this which I endorse.

20 Cf. Jaggar 2000a (235-40); it is clear from the definition of culture-laden realism itself that this is a consequence, but the next Chapter provides the background for this thought.

an argument but an instance of begging the question: not giving such subjective states inherent weight is precisely what distinguishes culture-laden realism from its rivals. However, Jaggar also invokes the idea of self-government. I take it that this – as well as her references to the substantive good method as 'illiberal' (320) – marks a concern that using this method will be morally wrong<sup>21</sup>. There are two commonly-used moral concepts that might plausibly be in play here.

Firstly, Jaggar might be concerned that using the substantive good methodology will involve *paternalism* – it will claim to know the good of individuals better than they do themselves, and make prescriptions concerning them accordingly. Secondly, she may be concerned that there is a threat of *alienation* looming here<sup>22</sup>. The substantive good conception refuses to take people's existing self-conceptions and normative perspectives as read, most obviously because those may be problematically adaptive. In engaging *some* people in discussion, or developing actions or systems of policy that might affect them, the substantive good method will not be able to begin by engaging them on their own terms. The political change that the substantive good methodology would recommend – which would necessarily include a change in people's perspectives – will then cause some people to be alienated from themselves, or from aspects of themselves<sup>23</sup>. Both of these concepts *are* potential problems for a substantive-good-based capability approach, in part for reasons Jaggar notes. Substantive good approaches cannot rule out the possibility that individuals' perspectives will

---

21 If this is right, Jaggar's normative criticism seems to contradict some elements of her epistemic criticism. She relies here on the normative cogency of certain moral claims. However, her claim elsewhere seems to be that the substantive good method she opposes is guilty of the same thing – relying on certain moral theses in thinking about the future path of normative methodology. See §2.3 below for a generalised account of this threat.

22 The quotation from Axel Honneth which Jaggar uses in her (1999, 318) also seems to involve a claim that alienation is a serious wrong, which may support my interpretation here.

23 Compare the discussion of alienation in the context of Williams' argument for his reasons internalism in the next Chapter (§2.2.2)

be judged to be defective (i.e. adaptive). That being so, others might have better knowledge about an individual or group's good than that individual or group themselves do. Similarly, the correct course of political change might then involve alienation. If alienation is a very serious harm, or if paternalism is extremely disrespectful, these will represent powerful objections to the substantive good methodology<sup>24</sup>.

I have limited space to argue about first-order normative questions here, and I certainly cannot adequately defend the view that concerns about paternalism and/or alienation are not cogent. In lieu of that, I shall restrict myself to two points. Firstly, although I have certainly not shown that alienation or paternalism do not pose problems for a substantive-good-based capability approach to justice, Jaggar has not shown that they *are* problematic either; she seems to simply assume it. Secondly, is it really so implausible, *prima facie*, that individuals or even groups can be so corrupt that they would be better off undergoing paternalistic interference, or alienating social change? Examples such as the de-Nazification of Germany, or of intervention into self-destructive religious cults, suggest otherwise.

### **§2.1.2 Methodological Exclusion; Authority and Expertise**

The second worry that seems to be present in Jaggar's normative critique concerns the exclusion of some people from the methodology that, in application, is supposed to govern their lives. This differs from the first worry in that it does not seem to involve a claim that using the method will make people's lives *worse* or violate rights to independence or negative liberty. Rather, the thought is that failing to include every perspective as an equal influence in

---

24 I say 'very serious' and 'extremely' because even if these things are bad- or wrong- makers to some extent, this would not automatically show that the method that involves them is indefensible. It may well be the case that *any* reasonable course of action in the contemporary world will involve *some* bads/wrongs. Cf. Ch.6, esp. §§2-3

a methodology is problematic independently of such harms or rights-violations.

Jaggar's claims here are powerful, if they are on target. The substantive-good methodology *does* involve treating some perspectives, or elements thereof, differently from others. Engaging in substantive thought about the good, we begin with plausible theses about the good, modifying them as a result of engagement with the experiences of epistemic peers. But who *counts* as an epistemic peer is itself affected by whether, and why, they accept or reject enough of these theses; if they don't accept some of these theses, their perspective will be adaptive, and to that extent they will not be an epistemic peer. As a result, the account of the good (and, relatedly, of justice) that is the result of substantive-good methodology is likely to apply to people who have not had an equal influence on its formulation. Jaggar holds that this methodology is elitist, in that it privileges theorists over everyone else (2006, 318-9), and prestigious, assertive, and domineering<sup>25</sup> theorists over other theorists (317-8), and probably neocolonial, in that it is likely to privilege 'Westerners' over global others (319). Elsewhere, she juxtaposes the conceptions of moral justification she opposes, described as 'covertly elitist and authoritarian', against reformed conceptions that would be 'transparent and democratic' (2000b, 241). As I noted above, what seems to unite these elements of Jaggar's normative critique is the idea that using methodologies like the substantive-good conception implicitly privileges the user, setting them up in a position of *authority* over others that those others have not accepted, and likely would not accept. If this is a correct description of what such realist conceptions involve, it seems clear that they should be rejected.

However, this criticism seems to elide a very important distinction, and I shall argue that it is

---

<sup>25</sup> These being, typically, male-gendered traits. See Code 1991.

only because of this elision that the features of substantive-good methods that Jaggar identifies look normatively problematic. The distinction that Jaggar's discussion elides is between *expertise*, which is a concept that realism needs to make use of, and *authority*, which realists need not say anything about, at least at the level of epistemic methodology. The agents who are, relative to some particular topic, better situated to possess knowledge about the good are claimed to be *experts*. Someone could be an expert for some ethical/political subject-matter because of (e.g.) the possession of certain skills, or a large amount of pre-existing relevant knowledge. Some of these skills, such as the ability to understand the gender significance of some particular action or policy, may be easier to acquire in some cultural contexts rather than others<sup>26</sup>. Likewise, some pieces of knowledge, such as knowledge of the ways in which gender or race are not normatively-relevant categories, may have an existing currency in some cultural contexts (i.e. feminist and race-conscious ones) that will make them easier to attain. Possessing such skills, or such knowledge, however, is not to be identified with the possession of authority over others.

This distinction is a fundamental marker of the difference between realisms and constructivisms. If one is a realist, one does not think that moral prescriptions are *commands* that have their force in virtue of a relationship of authority between commander and commanded. The voluntarism that makes rightness or wrongness a function of someone's will is a distinctively *constructivist* thesis.<sup>27</sup> Sometimes people know things that others don't due

---

26 It is important to note that talk of 'culture' need not suppose that there are internally homogeneous, unchanging, and uncontested cultures to which individuals either wholly and exclusively do, or wholly and exclusively don't, belong. To talk of 'culture' in the broad sense that I intend is just to talk of ideas that are present within human societies, and then of the patterns of enculturation that development and ongoing life in such societies involves. I see no barrier to accepting the view of culture that is briefly proposed in Jaggar 1999 (314-9), or the sophisticated model that is elaborated in Anne Phillips' *Multiculturalism Without Culture* (2007).

27 See fn.7 above. Most of my negative claims about constructivism in this Chapter can profitably be compared to the excellent discussion of Gaus' (2011) constructivistic view in Enoch 2013.

to a difference in circumstances that includes, potentially, a difference in cultural location (a difference in the enculturation that has formed them thus far). But these differences are not relevant to the force that prescriptions have; any force that they have derives solely from their relationship to the truth, not from the status of the prescriber.<sup>28</sup>

If we bear this in mind, the authority-based criticism of realism that I have sketched is either mistargeted or question-begging. Jaggar finds the substantive good method problematic because she thinks that the person pursuing it claims the *authority* to decide which intuitions are good evidence and which are deformed. To the extent that these references to authority rather than expertise are ineliminable, the relevant criticisms – concerning elitism, or concerning neocolonialism – will be mistargeted. Suppose, however, that we can replace every reference to authority with a reference to expertise. In this case, Jaggar's criticisms may, at least largely, be on target. But, to the extent that these are not *epistemic* criticisms, which would work within realism by urging that we cannot safely treat cultural location as being of epistemic significance, they will be question-begging. All of the claims that the constructivist would be attributing are central theses of culture-laden realism; bringing them out does nothing to damage it unless there is some sound prior argument that the concept of moral expertise is normatively problematic. Jaggar provides no explicit argumentation here; she just seems to think it obvious. Where constructivist criticism takes this form, it does not amount to a critical engagement with realism so much as a failure to take it seriously at all.

## §2.2 Privileging one's own Perspective – Bias

---

<sup>28</sup> As well as the sources I note elsewhere, Jaggar 1998 explicitly endorses the Habermasian political constructivist tradition, defining moral justification in terms of actual (rather than hypothetical) consensual agreement under idealised conditions (see esp. 8). This is clearly a voluntaristic conception of objectivity.

I have argued that Jaggar's normative critique is not very convincing as it stands. In this section, I move on to discuss her epistemic critique, noting a number of points at which it seems to mandate alterations in the substantive good methodology (at least as practised by Nussbaum). However, I shall argue that these warranted reforms do not amount to a transformation from culture-laden realism to constructivism; indeed, I shall suggest that it is only within the framework of realism that some of them make sense. The section is divided into halves; in §2.2.1 I discuss the general concern that Jaggar raises about the reliability of the substantive good method, while §2.2.2 addresses a specific issue about the epistemic contributions of marginalised groups. In the first part, I shall argue that in general, Jaggar's arguments fail to justify a move away from culture-laden realism; in the second I concede that, while Nussbaum's specific approach is lacking in relevant respects, realism is a more promising approach to issues of marginalisation than constructivism in general.

### **§2.2.1 Individualism and Epistemic Path-dependence**

Jaggar's central complaints about the substantive good method concern the extent to which the method's starting point in the user's own, individual, normative perspective might bias its future course. If this claim is sound, all of the judgements that the followers of the method arrive at will be suspect, because (as Jaggar also notes) the substantive good conception is a path-dependent method – where you end up depends on where you began. This is guaranteed to be the case, because this method is a form of (wide, interpersonal) reflective equilibrium: every change that occurs is the result of an alignment of old elements with new ones, or dissonant elements with one another. Path-dependent methods, if they are to be epistemically sound, will need to ensure that the starting points are as good as possible. One way to achieve this is might be to widen the starting point to include as many theses as possible.

It is against this background that Jaggar's epistemic criticisms have their force. Jaggar suggests that, in various ways, Nussbaum's substantive good conception fails to be as sound as it might be, by failing to be as *inclusive* as it might be. Instead, Nussbaum's methodology makes it likely – given some of the insights of feminist epistemology – that important normative truths will be overlooked. Jaggar mentions a number of mechanisms by which this could occur, including systemic hermeneutic inaccuracy (2006, 319)<sup>29</sup>; marginalisation as a result of linguistic difference (1998, 12-3), or pre-existing social hierarchy that gives some agents relative prestige (2006, 317; 1998, 10); and, above all, individual theorist-centred methodology (2006, 320). My aim here is not to fully specify an epistemology for capability theory, so I shall not go into these topics at length. Instead, in the next section, I shall focus on perhaps the most important specific issue that Jaggar raises – the question of the extent to which substantive-good methodology might reinforce the marginalisation of certain groups.

Before I turn to consider that, however, there are two very general points that Jaggar makes that again bring fundamental meta-level differences between constructivism and realism to the fore. The first comes as part of her claim that the substantive-good method is 'monological' (317-8) in a Habermasian sense. As a result, it cannot constitute 'independent' justification: in every case where new content is introduced into the method, its introduction is shaped by pre-existing items of content, and so no confirmation from outside for existing content can ever truly be independent<sup>30</sup>. This claim is question-begging, however, in its insistence that the (flexible and changing) resources of a not-all-inclusive perspective could

---

29 See also Fricker 2006, for a compelling development of this concept.

30 Jaggar's invocation of this idea is strange. As I note later (§2.4), she, like many feminist epistemologists, accepts a Quinean post-positivist model of epistemological naturalism. Against that background, surely Jaggar's demand for independence here presupposes the theory/observation dichotomy, which Quineans must reject?

not be sufficient for justification. Jaggar, likely because of her discourse-ethical affiliations, presupposes that objectivity requires a model of *dialogue* – no one 'person'<sup>31</sup> can achieve justification on their own. It is not clear why anyone who is not already a Habermasian constructivist has any reason to accept this. The second general point concerns Jaggar's claim that the substantive-good methodology is suspect because there could be 'multiple equilibria' – multiple coherent, but differing, normative systems. As she says, this cannot be ruled out unless we have considered all possible responses (2006, 315)<sup>32</sup>. Arguably, given that the central mark of realism<sup>33</sup> is the view that objectivity is not a function of subjective evidence alone, any realist must hold that the choice between two *apparently* maximally-coherent and equally complete conceptual systems will always be illusory – one will always be closer to reality, and therefore superior, even if we do not at present have the resources to identify which. If Jaggar is denying this, her argument is, again, question-begging.

### §2.2.2 Marginalised Voices

Apart from this general claim that substantive-good methodology is not epistemically sound, Jaggar raises an important specific issue, concerning the treatment of marginalised perspectives. Take the case of stigmatised and socially dominated minorities, such as (in most societies and my own) women, ethnic minorities, and minorities of gender and sexuality. Jaggar's claim is that theorists, who are in the contemporary world likely to be from dominant sociological categories, are not likely to share much cultural background with disadvantaged

---

31 I am extending Jaggar's speech metaphor here – of course, the users of a given instantiation of the substantive-good method will not be a single person *literally*. Separate people here represent putatively separate forms of enculturation (or 'traditions', or *bildung*).

32 Strictly speaking, it seems that we could not infallibly *rule out* the possibility of other equilibria even *if* we had considered all possible responses: equilibrium is not merely a function of the set of inputs, but of the path that is taken – that is, of the order that the inputs are considered.

33 According, that is, to the Dummettian view that I am working with in the Thesis. See Miller 2010; compare Wright 1992.

others. Culture-laden realism could then provide a license to dominate those others, by creating a category of cultural settings that do not conduce to understanding of the good, and placing them within it<sup>34</sup>.

This could occur, but as Jaggar recognises (2006, 310-11), no method can *guarantee* success or render abuse impossible. It will be necessary to put in place epistemic practices to help prevent such abuses, and many of those that Jaggar recommends, including most importantly the practice of deference to the testimony of well-placed members of such groups<sup>35</sup> on questions that relate to ways of life of which they have experience<sup>36</sup>, look like good candidates. In this way, much of the detail of Jaggar's suggestions can be incorporated into a reformed substantive-good methodology. What no culture-laden realist can accept, on the other hand, is the idea that the acceptance of people as epistemic peers has no limits: that we should be *unboundedly* inclusive. Jaggar's arguments may suggest that Nussbaum's use of the substantive-good methodology, at least, is not inclusive enough – she does not seem to be committed in practice to engaging with a wide range of diverse perspectives. But widening

---

34 Technically, it looks quite unlikely that a user of the substantive-good method would be justified in doing this, even within that method's own terms: in most cases, the overlap between conceptions of the good and sociological categories like these is relatively loose. The rejection of capabilities approaches, for example, doesn't seem very likely to supervene on gender categories like 'woman', or broad racial categories like 'Black American'. The sociological categories relative to which this sort of consideration is more likely to be relevant are cultural or (especially) religious categories membership of which is *constituted* by attachment to certain norms. (It remains an open question whether there are any such categories.)

35 A second issue, which Jaggar discusses in much more detail in her 1998 (esp. 11-4), concerns whether there even *are* members of marginalised (including cultural) groupings who are well-placed to articulate what is of value and what is of disvalue in their particular ways of life. If a group has traditionally been subject to domination from others, there may not be any view of the good that is *theirs*, but which is not contaminated by such domination. Jaggar concludes that it might be necessary to allow such groups to retreat into exclusive spaces ('subaltern counterpublics'), so that they might develop and endorse such a view. Only then will it be possible for others to engage with them on equal terms. Again, both this possibility and Jaggar's consequent prescription appear to be entirely compatible with realism.

36 Jaggar references Lugones and Spelman 1983, and Thomas 1992-3, in this regard. See also Mohanty 1991. At least most of the recommendations contained therein seem to be compatible with realism, as part of an epistemically improved (because more inclusive and with less potential for hermeneutic injustice) methodology, but I can't engage further with the details here. Wylie 2003 provides a pre-eminent statement of standpoint theory, the generalised epistemological framework within which these sorts of claims fit.

the range of considered perspectives and imposing no bounds at all upon that range are very different things, and Jaggar has not done enough to require the latter response.

In addition to the extent to which realist methodology can create barriers to such abuse, the tables can be turned on the constructivist here. Where it seems as though we have reason to defer to the self-descriptions of groups over the descriptions of others, for example, culture-laden realism's paradigm of expertise can make sense of why this should be so in a way that constructivism's paradigm of authority cannot. After all, culture-laden realists *already* accept that knowledge about normative subject-matters is only accessible by those who have been enculturated in some particular way (probably, in actuality, some very broad *set* of ways).<sup>37</sup> However, it is not necessary to think, even though objectivity is available to us today, that *total* objectivity<sup>38</sup> can in practice be obtained by some single individual. It should go without saying that I, as a white British male, am not going to be able within a lifetime to incorporate all the normative insights about the plural human good that may be embedded in traditions I have little or no immersion in. *A fortiori*, I do not possess such insights here and now, and of course no other individual will either<sup>39</sup>. From a constructivist point of view, however, where authority is the central concept, and authority is equally distributed across humankind, it is mysterious why anyone should be deferred to. If we begin from the position that all preferences/normative beliefs/etc. have equal weight, what could the rationale for deference to others in specific circumstances be? Why should the preferences of those with greater familiarity with a way of life be accorded greater significance, if all have equal authority to

---

37 See the next Chapter for an exegesis of McDowell's culture-laden realist claims in this area.

38 I have not, as yet, given any account of what I take objectivity to be (apart from gesturing towards some form of culture-laden realism. To the extent that I have space to provide an adequate account, this will come later in this Chapter (§2.4-3.2).

39 Robeyns 2005 uses this uncontentious fact as a basis to criticise the construction of lists of capabilities. It is not clear, however, why the fact that a method is fallible, when followed by any given individual, should lead us to give it up.

decide for all, and there is no mind-independent matter of fact to which some might have better access?<sup>40</sup> Realism gives a much better general explanation of the concrete grounds for testimonial deference than appears to be available to the constructivist.

### §2.3 Constructivism as Self-Effacing

I have suggested that although culture-laden realism can (and should) incorporate many of Jaggar's suggestions, the ultimate end-point of her critique seems to be the idea that, contrary to Nussbaum's methodology, no perspective may be excluded from normative discourse on substantive grounds. Jaggar endorses the idea that our methodology should be morally-laden (2006, 310-15; 2000b, 235-40) – insofar as it should not allow some voices to dominate over others, for example through threats of violence – but rejects Nussbaum's thesis that this moral content should inform not only the way that discussion takes place, but also what it is reasonable to discuss. In this section, I argue that constructivism is self-effacing<sup>41</sup> in a number of ways, and that these threaten to render incoherent Jaggar's position here.

To see how this might be so, it may be helpful to turn back to consider Nussbaum's own view, as explicated in the previous Chapter. A central thesis of hers is that the capability approach is justified because it is ethically plausible – it is something that seems true increasingly as we progress towards wide, interpersonal reflective equilibrium: as we take in more and more information, and the insights of more and more others. On the other hand, however, she

---

40 Of course, we could imagine a constructivism that *didn't* begin from the view that all (e.g.) preferences have equal weight; a kind of culture-laden constructivism (cf. the discussion of LeBar's view in the next Chapter on this point). But again, in the absence of a belief in some mind-independent truth to which some set of people has easier or unique access, why would we decide to begin by treating one preference or belief differently from another? My discussion here is very quick, but this argument is parallel to that of David Enoch, which I discuss in the next Chapter; see there for a more adequate treatment.

41 The use of a self-effacingness objection here follows Claassen 2011. The contemporary *locus classicus* of this sort of self-effacingness critique is probably Parfit 1984, Part 1.

concedes that we have reason to leave the substance of the view quite vague, not merely for pragmatic reasons (so as not to alienate potential allies), but also for normative reasons, which she associates with Rawls. There are a number of positions that distinguish capability theory from its competitors. All capability theorists as such accept that there are some constraints on a theory of justice – that it emphasise capabilities as well as functionings, that it be pluralist about the good, that it tailor its prescriptions to the actual circumstances of individuals, etc. This sets up a distinction between general or structural claims, about which capability theorists agree, and issue as prescriptions for the global *polis*, and specific items of content, about which there is little agreement, and about which Nussbaum does not issue prescriptions. If such theorists then endorse a constructivist view of value, according to which the facts about goodness/rightness are determined by what people (usually, as a group) think or want, this presents a challenge. Why hold that the general *structure* of the correct theory of justice is evident, when its specific *content* requires a deliberative process (actual or ideal) to be decided upon?<sup>42</sup>

To flesh this out, consider an opponent of the capability approach's response to it. One of the lines of argument they might use, exemplified by Jaggar's critique, would trade on the extent to which capability theories might involve a disrespectful disregard for the thoughts and feelings of (many, at least) of the very people that they are designed to help. Perhaps such an opponent would begin by disputing some detail of the specific capability approach under attack – if Nussbaum is the target, perhaps her description of the capabilities list. But, if what they are really concerned with is a failure to engage with the worldviews of others, it will make little sense to stop there. After all, most capability theorists, even those who most

---

42 This line of argument is similar to a criticism of capability theory by Richard Arneson (2010, 112-8). Cf. Claassen 2011, 497-8.

emphasise public deliberation and democratic input, do not actually show that the capability-based theories of justice that they propose are congruent with the existing beliefs and affective responses of impoverished (or, indeed, other) people. Some of them, some of the time, make moves in this direction (e.g. Nussbaum 2000a, 148-161)<sup>43</sup>. This notwithstanding, it remains the case that the primary motivations that they give for capability theory stem from normative claims made in abstraction from survey-taking or the implementation of deliberative-democratic procedures. When Sen, for example, criticises resourcist models for failing to pay sufficient attention to the specific needs of each individual, or welfarist ones for failing to respect the importance of choice and agency, comprehensive consultation of people's *actual* beliefs and desires is nowhere to be seen. This line of argument amounts to a dilemma for advocates of the capability approach – either endorse some constructivist view, and defend all of the distinctive claims of capability theory on that basis (which is likely to be extremely difficult to do given the diversity of normative perspectives in the world), or endorse realism, accepting that not *all* normative claims (at least) require a grounding in the existing desires of human agents as such.

However, this is not merely a problem for those who half-heartedly endorse constructivism. Even avowed constructivists such as Jaggar face contradictions of this kind. Firstly, Jaggar's constructivism *itself* can hardly be construed as theoretically neutral – after all, Nussbaum disagrees with her about it. If it is not neutral, then surely people should have their say about it, and this opens up the prospect of constructivism collapsing into self-contradiction, for people might well not come to agree that constructivism is the correct meta-level view after all! Secondly, apart from the prospect that constructivism *per se* might be rejected, we require

---

43 Cf. Jaggar (2006, 311-5).

a more specific method – some *version* of constructivism – if we are to actually engage in thought about the good or justice.<sup>44</sup> Does Jaggar have particularly good reasons to think that her preferred version of constructivism is uniquely accepted by a wide range of epistemic agents? If so, she does not offer them. Finally, it should be noted that some of these methodological factors within constructivism are, or implicate, controversial normative claims. Jaggar holds that the ideal speech situation involves a level playing field between men and women, for example. Is there widespread agreement across the world that women ought to have an equal say on normative questions with men<sup>45</sup>? Even if there is, history tells us that this is a sadly contingent and fragile state of affairs. And if not, constructivism seems to have to come to an end at some point, just presupposing answers to questions that could be, or actually are, still controversial. If it does this, how is it fundamentally different from culture-laden realism?

My argument here has recent parallels in feminist-epistemological work. For example, Daniel Hicks has recently argued that Helen Longino's conception of objectivity (intended primarily for the practice of science) is unable to distinguish between substantively feminist and anti-feminist perspectives. What is more, her account of objectivity requires that enough 'alternative viewpoints' be 'actively cultivated' for it to be the case that all widely-shared

44 See Antony 2008, esp. 574-6. Compare also Claassen 2011, esp. 495.

45 Jaggar has said: “Feminists are certainly unwilling to concede that the subordination of women may ever be morally justified and this core claim has a status something like that of an analytic statement in the feminist “web of belief.” Apart from this minimal moral conviction, however, feminism’s views about the processes, methods, and conclusions of good moral thinking are sufficiently varied, contested, and negotiable that each can provide a useful check on the others.” (2000a, 465). This suggests that Jaggar's objection to Nussbaum – that her view is exclusionary because it allows items of substantive content to have an impact on which perspectives are taken as evidence – is somewhat disingenuous. This makes it appear that the only difference between Jaggar and Nussbaum in this respect is that Jaggar's substantive requirement is relatively vague, and so does not introduce quite so *much* normative content. If so, Jaggar cannot coherently object that it is an immediate, fatal flaw that “Nussbaum's test for determining whether or not a desire is informed seems to be precisely whether or not it can be interpreted as a desire for one of the items on her list” (2006, 316). At most, she could claim that the capabilities list, unlike her even vaguer commitment to work against 'the subordination of women', is not sufficiently justified that it is reasonable to use it as a practical test for epistemic peerage.

beliefs within a community are subject to the critical scrutiny that she thinks objectivity requires. As Hicks points out, this plausibly means that feminists would be committed to not only including but actively cultivating substantively anti-feminist viewpoints, up to and including neo-Nazis. Hicks then argues that the only way to avoid this criticism is to mandate, for participants, a liberal-style explicit commitment to the principle of epistemic inclusion itself, which would exclude (e.g.) most Nazis, but would also exclude many communitarian and non-liberal thinkers including (as it happens) Jaggar herself (2011). In this way, the use of (allegedly) feminist values of inclusion and anti-marginalisation to shape the course of ongoing inquiry can function to undermine those very values<sup>46</sup>.

A second important example comes from Kristen Intemann (2010, esp. 789-93). There she posits a division between 'standpoint' and 'empiricist' feminist epistemologies<sup>47</sup>, and argues that the primary difference between them concerns their conception of, and attitude to, diversity. Briefly, the former view the epistemically valuable sort of diversity in terms of differences in position within social hierarchies, while the latter identify differences between people's values and interests as being of relevance.<sup>48</sup> Intemann then argues that the approach to diversity of the 'empiricist' camp is untenable because:

On the feminist empiricist account, the content of values and interests is irrelevant to

---

46 Hicks' argument is significantly different to mine, and I note it only as an emblem of the wider problem for proceduralistic approaches to methodology. Firstly, it is targeted at a much more liberal version of feminist epistemic theory than Jaggar's own. And, secondly, its target is a view of scientific objectivity, which brings in some different issues to the ethical/political objectivity that is my concern.

47 This distinction is potentially quite misleading, and Jaggar is a good example of this. She is a standpoint theorist (Jaggar 2004), but is also extremely friendly to the appropriation of Quinean 'naturalist' ideas that are distinctive of feminist empiricism (e.g. in Jaggar 2000b).

48 Jaggar's account muddies the waters between these two kinds of exclusion: standpoint exclusion and valuational exclusion. In her critique of Nussbaum, the most objectionable implications are consequences of her attack on valuational exclusion; however, she also discusses narrower standpoint-based difficulties, and is more usually identified as a standpoint theorist (e.g. Jaggar 2004). My account actively denies the distinction: unless standpoints are individuated in such a way as to preclude them from representing evaluative perspectives, many standpoints will do so (if only accidentally). If so, the criticisms I have been making will apply to proceduralist versions of standpoint epistemology just as much as explicitly evaluationally-inclusive approaches.

the purported epistemic benefit they provide in contributing to diversity. That is, all values and interests are equally beneficial in contributing toward a diverse community that will be likely to identify when the value judgments of others are operating as background assumptions. This is worrisome, as it suggests that the ideal scientific community is one where all values and interests are represented, including those that are potentially problematic. For example, do we need to include members of the Flat Earth Society in research in astronomy? Do we need to ensure there are representatives of chemical companies in research on environmental toxins? Should the interests and values of tobacco companies be represented in cancer research?

(791-2)

Further feminist scholars have endorsed similar restrictions on the practice of inclusion of value-laden perspectives in ethical enquiry<sup>49</sup>. The conclusion of all these arguments tends in a single direction – against the pure procedural principles which characterise constructivist meta-level positions, and towards the explicit and clear incorporation of substantive judgements into ethical methodology.

## §2.4 Making Sense of Epistemological 'Naturalism'<sup>50</sup>

Most of my claims in this Chapter thus far have been negative, arguing that constructivist framings are implausible. In this section and the next, I shall begin to concentrate on the positive case for a culture-laden realist understanding of methodology. This section is concerned with some theoretical ideas which have informed much, if not most, feminist epistemology over the last two decades. Discussing them will have two benefits. Firstly, it will give a final recapitulation of my claim that constructivism does not adequately explain the epistemic methodology that is necessary for (e.g.) answering questions about the constituents of human dignity, while culture-laden realism does. Secondly, it will provide a

---

49 Cf., for instance, Haslanger (2000, 46): “what is important is not whether a particular account ‘marginalises’ some individuals, but whether its doing so is in conflict with the feminist values that motivate the inquiry.” And cf. Benhabib (2006, 286): “all demands cannot be said to be valid because they are articulated by groups with a history of marginalisation or oppression. Oppression and marginalisation are no guarantee of progressive or just politics.”

50 'Naturalism' requires scare quotes here (if not everywhere!) because, as Anderson (2011) notes, although feminist epistemologists often follow Quine in many ways, they do not (*pace* his attempts to the contrary) aim to produce a psychologistic, non-normative reduction for epistemology. Cf. Code 1996, and see Held 2002 for critique in this vicinity.

basis for the more thoroughly metaethical claims of the next Chapter, much of which will presuppose an understanding of 'epistemological naturalism' and its role in feminist thought.

As I have said, the core concept in question is called 'epistemological naturalism' by most of those in the feminist-epistemological literature; I shall occasionally also refer to it as 'Neurathian', for Otto von Neurath and his eponymous boat (1944, 47). The function of this concept is to register the way in which epistemological subjects and processes can be examined empirically, and the results of this examination can feed back into the epistemic practices themselves. Thus, for example, learning that a particular method excludes women by bringing to salience features of women's appearance or behaviour, which will then cause testimony to be dismissed because those features are associated with negative implicit biases, can function to discredit the method epistemically. Quite generally, epistemological naturalism involves a feedback loop between existing empirical methods and the judgements that rationalise them, and the new judgements that may be generated if parts of the methodology (i.e. theory-construction and observation) are deployed in critical reflection on other parts (i.e. the broader practice of knowledge-creation, in which gender (and other social) norms are often observed to be operative). This kind of empirically-based critique is necessarily holistic, because its Quinean practitioners reject the idea that there are any deep, principled divisions within background theorising about the world which preclude some content from bearing on the interpretation of the empirical observations which methods (and the broader belief-generating practices within which they sit) produce. Not only is there no deep epistemic distinction between theory and observation – because all observations depend on existing judgements to be incorporated as empirically meaningful; there is also no in-principle reason why one region of the existing 'web of belief' ('theory') can be made use of

in such incorporation, while another cannot. Thus, Quine and his followers conclude that, strictly speaking, the background theory is only ever confirmed or disconfirmed as a whole; individual observations can only be thought to justify individual beliefs if an enormous amount of (in-principle alterable) background content is taken for granted.

A vital and extremely relevant strand within epistemological naturalism then concerns the positive role that feminist values themselves can play in improving epistemic practices. The general thought is that values such as the equality of investigators can, because they form part of the existing background 'theory', motivate and guide empirical investigations which will reveal biases and other unreliabilities in operative methodology, and thereby have epistemic benefits. Louise Antony, in inaugurating the tradition of 'empiricist feminism', identified what she called 'the bias paradox' here (2008, 554-6). The core thought here is that, although feminist values are used to *expose* (e.g.) gendered biases operative in existing practice, there is also a sense in which they constitute biases themselves; like all elements of the background theory (in principle, at least), they shape ongoing inquiry, while being themselves substantive claims, and thereby (in some sense, at least) revisable. The general response from feminists must involve an argument that the egalitarian 'biases' are better, epistemically speaking, than the gendered biases which they are used to critique. Different feminists have given different versions of this critique; some are more consonant with my endorsement of culture-laden realism than others.

Antony's own response, however, is congenial to my general claims about methodology and the meta-level frameworks which best make sense of it. Her main claim is that 'the real problem with the ruling-class [androcentric, racist, etc.] worldview is not that it is biased; it's

that it is false' (574)<sup>51</sup>. With reference to examples such as the design of experiments on language development which could (or could not) include both boys and girls, or people of different races, she argues that the assumptions which lead to sexist conclusions are reliably implausible (empirically speaking) when considered in the clear light of day, and usually reflect dubious essentialisations and/or naturalisations of categories which our best empirical (sociological, anthropological, etc.) theory tells us to be socially constructed (576). Thus, Antony's defence of the use of feminist values in guiding inquiry is that, when we use these values, we arrive at a more generally coherent, more complete, or more empirically adequate theory, than we would do if we instead incorporated other values. For Antony, this is just to say that these values are likely to be part of the true theory of reality.

As I have noted, Antony's positive resolution of the 'bias paradox' is broadly realistic. Her paper also contains useful thoughts about the deficiencies of a constructivist-sounding alternative. As she says there:

Naturalized epistemology tells us that there is no presuppositionless position from which to assess epistemic practice, that we must take some knowledge for granted... a genuinely open mind, far from leading us closer to the truth, would lead to epistemic chaos. (571-2).

The problem is that – as Quinean empirical holism suggests – if no substantive presumptions are made, epistemology becomes impossible; there are too many ways to interpret the observations which are made, and so no conclusion can be securely justified (2008, 571-3). So, it seems that any constructivist approach that seeks to evade the need for a total *aporia* will need to concede the importance of certain (probably not universally accepted) values, at least in order to orient the direction of future inquiry. However, constructivism requires an explanation of the role that such orienting ideas would play. Notably, realism has such an

---

51 Cf. also Langton 1993.

explanation – we make use of these value-laden beliefs because they are places where we have particularly good evidence, places where we may be confident that we are relatively close to the truth, such that taking contrary testimony seriously as something with epistemic weight could not fail to worsen our epistemic position, making the resulting judgements less objective, rather than more so (Antony 2008, 556). It is not clear what parallel move a constructivist can make, however, other than simply insisting that *these* orienting ideas – rather than all the possible alternatives – are constitutive ingredients of objectivity<sup>52</sup>.

My negative case against constructivism in the context of epistemological naturalism is substantially complete. However, before I conclude, it will be illuminating to consider one other means of defence for the presence of feminist 'biases' in epistemology: that of Elizabeth Anderson. Her interpretation, unlike Antony's, is not especially coherent with culture-laden realism, and it will be helpful to make clear why I reject it.

Anderson, in arguing that feminist values can play some role in methodology<sup>53</sup>, does not appear to want to claim that this is because those values are good in their own right. Rather, she argues that *some* values are needed to bridge the gap between theory and observation (1995a, 27-32), and then that 'our values, interests, and aims, some of which have moral and political import' (37) do this by shaping the course of enquiry in a variety of (she argues) thoroughly legitimate ways (37-55). However, she never actually specifies what 'our' values, interests, and aims are, which are supposed to do this work<sup>54</sup>. In her other prominent paper on

---

52 Of course, people can stipulatively use terms like 'objectivity' as they like. But if constructivists are merely stipulating that objectivity requires feminist values, they are conceding that they have nothing rational to say to anyone who disagrees – they will be doomed to talking past one another. Cf. Douglas 2004, on the many meanings of 'objectivity'.

53 Her paper is aimed at scientific methodology, and it is not clear to what extent it is intended to generalise, so my rejection of her proposals for the ethical/political case may not apply to her actual target in the paper.

54 Intemann 2005 provides a good critique of this content-neutral aspect of Anderson's proposals.

the topic, 'Feminist Epistemology: An Interpretation and a Defence', her defence again turns on considerations distinctive of naturalised epistemology: the idea that some epistemic 'processes are such that, once we reflect on how they work or what they do, we lose confidence in the beliefs to which they give rise', because they are revealed to be unreliable (1995b, 55). Like Antony, she is happy to talk of truth in explicating this notion of reliability/unreliability. However, some of her language is revealing:

Feminist studies of sexism in theories explore the prospects for alternative scientific theories that meet criteria of empirical adequacy while seeking to serve women's interests and to promote universal equality. (1995b, 58)

Here, the epistemic aspect of the defence of the involvement of feminist values is merely negative; some values are needed, and feminist ones are as good as any, since they meet criteria of empirical adequacy. However, there is no attempt to link the ethical/political aims of feminism with the epistemic role of these values more positively; there is no attempt to claim that feminist values actually do any *better*, epistemically speaking, than alternatives. The epistemic role and the ethical/political one are still functionally separate. It does not appear to be the case, on Anderson's view, that feminist values are more keyed in to the real world than other values. This is unsurprising, since all she urges, when generating an ethical/political component to fill the epistemic gap, is an apparently content-neutral 'reflective endorsement test' (1995b, 52-3). This would not be a very robust defence of feminist epistemology, even for the scientific case, because it is far from obvious how a view like hers can respond on epistemic terms to (e.g.) anti-egalitarian values, so long as the relevant individuals have reflectively endorsed them (which is hardly a high bar). It would have even worse prospects for the case of ethical epistemology, amounting to a form of proceduralism that disavows any claim to the better standing of the values which orient the inquiry. Such a proceduralism would be especially clearly vulnerable to the criticisms I've made throughout the Chapter. Thus, I hold that a realism-congruent view like Antony's has

much better prospects for making sense of the good standing of feminist values within the framework of epistemological naturalism.

To conclude this section, it will be helpful to bring the discussion back to the bigger picture. The dispute within feminist epistemology has tended to be between those who see the impulse of feminism as being towards an epistemology that is traditional in its fundamental premises, but more aware of ways that social inequalities might distort people's beliefs, and those who want more radical change, fundamentally altering (or even rejecting) the ideas of knowledge, reason, truth, and reality themselves. Although I have resisted one radical revision – meta-normative constructivism – this does not imply that the realism I endorse is not consonant with many of the more revisionary theses that feminist epistemology includes. As I have shown, culture-laden realism is actually conceptually well-equipped to incorporate many feminist insights. I conclude that feminist work should have a major impact on our conception of moral objectivity, and the normative methodologies within which it functions, but that there is no compelling case to abandon realism. Indeed, the upshot of feminist epistemology in some ways *supports* the kind of culture-laden realism that Nussbaum's substantive good methodology implicates, rather than undermining it as Jaggar has claimed. To end this sub-section on a conciliatory note, it seems that this conclusion may not actually be too distant from Jaggar's overall view. For example, her survey paper 'Ethics Naturalised: Feminism's Contribution to Moral Epistemology' (2000a), spends a great deal of time outlining the feminist theses that I am urging<sup>55</sup> are in harmony with culture-laden realism, and just a fragment of two pages making claims against 'objectivist realism' (463-4). This

---

55 To be more explicit, here I have in mind the emphasis on socially marginalised voices (subject to case-by-case ethical evaluation!), the recognition that social location can more broadly affect objectivity, and the recognition that affective response is not irrelevant to epistemic standing. Some of these have already been discussed; some will be discussed either in §3 below, or in the next Chapter.

suggests that, in the end, there is more scope for agreement here than might be apparent.

## §2.5 Rawls' 'Political Constructivism'

Throughout this Chapter, I have given several reasons to suspect that (some form of) culture-laden realism has better prospects than constructivism for playing the crucial background roles that are my focus in the thesis. The preliminary barrier to constructivist success involves the problem of adaptive preferences: how can a constructivist prevent preferences that have (*prima facie*) been warped by the unjust circumstances of their formation from distorting moral epistemology? I have argued that even very technically sophisticated preferentialist forms of constructivism – such as that of Baber – do not promise a satisfactory resolution of this problem, and so are likely to be unreliable as bases for epistemology. Despite this, it would be too quick to conclude that constructivism *as such* has poor prospects: many proposed forms of constructivism have not (at least explicitly) been preference-based, and some aim to face the problem of adaption head on, excluding or transforming the adaptive elements of subjectivity through idealisation or a constrained real-world dialogical process. In this section, I fill in the lacuna in my argument that this leaves. Firstly, I clarify the scope of the criticisms I make of constructivism in the thesis, and how they may be taken to demotivate it as a plausible source for a capability approach. Secondly, I provide some argument specifically against Rawls' political constructivism, with the aim of illustrating some significant challenges that such a view would have to overcome, and thereby demotivating a further possible way of providing a constructivist background for a universalist<sup>56</sup> capability

---

<sup>56</sup> One obvious challenge I shall not make involves the fact that – as Rawls himself (1999b) recognises – it is even more difficult to see how the idea of a mutually acceptable conception of justice could be established at the global level than it is at more local levels. This throws great doubt on the prospects for a genuinely universalistic 'political liberal' constructivism. However, my arguments will aim at the plausibility of Rawlsian constructivism as a method of justification for *any* realistic political space, rather than only this narrower issue.

approach.

Constructivism takes many forms. The most basic pick out some subjective category, and provide a completely content-neutral way of aggregating or otherwise integrating the various items within that category to produce prescriptions; preferentisms such as that of Baber fall within this type. They are precluded (by stipulation) from incorporating into their constructive procedure any test or constraint which does not flow straightforwardly either from the individual subjective elements alone, or from (*prima facie*) neutral sources such as logic<sup>57</sup>. For example, Baber's test of 'fruitfulness' satisfies this requirement, because it is simply a measure of the compatibility of the satisfaction of one preference with the satisfaction of numerous others. Constructivisms of this type are straightforwardly vulnerable to the problem of adaptive preferences (or analogues to it), and I take my argument so far to be sufficient to demotivate them as such.

Alternatively, a constructivist can attempt to design a procedure such that the (*prima facie*) adaptive elements will be prevented from having an influence on the result, by incorporating constraints which do *not* flow simply from the subjective basis alone. Two prominent ways of doing this exist. The first involves an actualised dialogical procedure, which may build in value-laden constraints, such as a constraint on gendered power inequality, but which allows the discursive process which (eventually) generates prescriptions to flow relatively freely.

---

<sup>57</sup> See Enoch 2006, Velleman 2009, and Enoch 2011b for an exemplary exchange concerning whether a (purportedly normatively significant) constructivism such as Velleman's can be regarded as flowing from a neutral subjective base (in that case, what is common between all instances of human agency). Rawls' strategy is similar to that of Velleman, except that its means of response to arguments such as these would necessarily be weaker: Rawls does not (even more implausibly) pretend that his norm of impartiality is constitutive of agency in the *polis per se*, only that it is implicated by normal ('reasonable') agency in modern *liberal* polities. I do not need to consider views such as Velleman's, because – as I have noted in Ch.2 (fn.62) above – they overwhelmingly tend towards relativism, and thus would be incapable of supporting a capability theory of global justice.

The combination of limited value-ladenness and inclusive flexibility is supposed to generate an expectation that adaptive elements, even if they are introduced throughout the process, will 'come out in the wash', so to speak; they will not ultimately be able to have a detrimental impact on the eventual conclusions. It appears that Jaggar's own preferred epistemology would be of this type, and many, if not most, proposals for deliberative democracy are also likely to fit this bill. The problems I have identified for such views stem from the choice of values to be incorporated into the procedure, and the necessary weakness of the grounds for any particular incorporation, given that all such values are far from being the subject of an unconstrained consensus, and given that no realist rationale can be offered which would bypass the need for such unanimity. In the previous two sub-sections, I have introduced this issue, but the deepest form of it can be located in the work of David Enoch, discussed in the next Chapter (especially §§1.4-5). This deep and unresolved problem suggests that realism is ultimately better placed to provide the background for an epistemology, even if it draws extensively on resources developed by actual-process constructivists in doing so. The existence of this serious problem for such constructivisms – which realism, by contrast, can avoid – suggests that they, too, can be considered poorly motivated for my purposes here.

The second strategy involves building a hypothetical procedure which aims to idealise away the problematic (/adaptive) elements directly, by drawing (explicitly or implicitly) on exogenous values to justify their exclusion from a reflective process (rather than an actual-world dialogical one). Because this approach may not need to incorporate feedback over time between the outputs of its procedure and the influences on the procedure itself, the meta-level concern about making sense of epistemological naturalism is less obviously pressing for it<sup>58</sup>.

<sup>58</sup> The concern does not really go away, however, as my discussion below will reveal: the question of *why* it is legitimate to privilege the value of political impartiality, and the liberalism-friendly perspectives of those who strongly endorse it, is still a significant one.

The most prominent example of a theory like this is Rawls' political constructivism. I proceed, then, in the rest of this section, by discussing Rawls' account in a little more detail, making clear where some of its weaknesses lie, and, thereby, throwing direct doubt on at least one of the main opponents to the species of claim that I prefer. The specifics, at least, of this critique of Rawls are unlikely to generalise to all hypothetical-process non-preferentist versions of constructivism *per se*. However, the weakness of this paradigmatic attempt to ground moral (or, in this case, political) epistemology without appeal to realism helps to demotivate constructivism, and thereby motivate the alternative I shall offer. I shall begin with a brief exegesis of the core ideas of Rawls' method (*circa Political Liberalism*), and an explanation of the way in which these are constructivist. Then, I shall make clear that the methodology which Rawls proposes is extremely demanding, both in that it builds in constraints that many contemporary individuals do not, on the face of it, accept, and in that even if those constraints are accepted in principle, it privileges them over the particular values that people espouse to an extraordinary extent. This makes it very difficult to see why many individuals would find Rawlsian argument compelling; if this is correct, Rawls' methodology must fail, on its own ambitious constructivist terms, to provide justifications suitable for guiding action in the *polis*<sup>59</sup>. Finally, I shall note that a method such as Rawls' also necessarily invokes a public/private contrast that is likely to restrict the content of (i.e.) conceptions of dignity, to the detriment of those whose unjust treatment disproportionately occurs in the 'private' sphere<sup>60</sup>.

---

59 To make it explicit: the point will not be that it is a failure of any justification *per se* that it will not be compelling to its audience, but that it is a failure of any *constructivist* justification to be so, because constructivist objectivity is premised on the availability of agreement without prior substantive transformation.

60 I mentioned this point in my discussion of the public/private issue in the previous Chapter (§1.3.2), but I have not substantiated it at greater length until now.

I shall begin, as I've said, with an outline of how Rawls' constructivism is supposed to work, focussing only on its most central ideas. The core of Rawls' methodology for thinking about the proper form of justice remained much the same from the original publication of *A Theory of Justice* in 1973 until his death; the differences concern the justification that Rawls offers for choosing this methodology, which evolved from relative ambiguity in his early work to a much more clearly constructivist approach. The central idea is that objective judgements about justice can be attained through reasoning about the proper basic form of society which has been purified of biasing, partial influences by a process of abstraction. Individuals are to imagine that they are designing principles both for constitutional design, and for the distribution of goods in society, but from the 'original position', which is distinguished from the actual world by a 'veil of ignorance', which keeps from individuals any particular knowledge which, it is argued, might improperly influence them. The agents in the original position are conceived of as otherwise – apart from the effects of the veil itself – representative of the psychology of 'normal' human reasoners<sup>61</sup>. In his late works, this method of hypothetical reasoning against a background of abstraction is justified as necessary for the achievement of objectivity, constructivistically construed.

Rawls' overarching constructivist conception of objectivity, as it figures in his late works, can be seen in the following:

Political convictions...are objective...if reasonable and rational persons, who are sufficiently intelligent and conscientious in exercising their powers of practical reason, and whose reasoning exhibits none of the familiar defects of reasoning, would eventually endorse those convictions...provided that these persons know the relevant facts and have sufficiently surveyed the grounds that bear on the matter under conditions favourable to due reflection.

This is then simplified:

---

<sup>61</sup> This core methodology is described and redescribed in almost all of Rawls' major works, but a relatively quick exegesis is given in his 2001, 10-26.

To say that a political conviction is objective is to say that there are reasons, specified by a reasonable and mutually recognisable political conception... sufficient to convince all reasonable persons that it is reasonable. (Rawls 2005, 119)

The specifics of Rawlsian methodology – the assembly of a 'veil of ignorance', etc. – are then to be viewed as a model of what it would take for a conception to be 'reasonable' and 'mutually recognisable', via an interpretation of the impartiality that 'reasonableness' itself is thought to entail. As Leif Wenar has convincingly argued (1995, 36-8), the concept of reasonableness that does the vast majority of the actual work throughout Rawls' late theory ultimately boils down to a set of features of the persons he stipulates as 'reasonable'. Two of the most significant and controversial among these – some, like the possession of 'the intellectual powers of judgement, thought, and inference' (Rawls 2005, 81), pose relatively little difficulty – are discussed in the next two paragraphs.

Firstly, 'reasonable' persons recognise 'the burdens of judgement', which, as well as including an uncontroversial admission of general human fallibility, require people to explain a wide range<sup>62</sup> of existing disagreement about politically-relevant questions in terms of that mere fallibility, rather than in terms of specific, attributable epistemic failures. If disagreement is explained in this way, a gap is opened for the justification of an attitude of principled toleration towards the opponent: the situation is symmetrical, because you are as fallible as your opponent, and so you have no special justification to incorporate your belief into the political structures which govern you both (2005, 54-8). If, on the other hand, the disagreement of the other were attributable to 'prejudice...bias, self- [or] group-interest, blindness [or] willfulness' (58) on their part, the way would be open to appeal to that

---

62 It remains entirely obscure exactly what Rawls intends for this range to include and exclude. He appears to be content with appealing to the sense of his audience (if such a sense exists) that most of the major accepted systems of belief can be endorsed by some without this being explicable by those who disagree other than in terms of fallibility. As he says soon after (60, fn.14) 'a reasonable doctrine is one that *can* be affirmed in a reasonable way' [my italics], rather than one which always, or even usually, *is* affirmed in that way.

diagnosis in justifying (i.e.) a policy which took greater heed of your perspective than theirs. The problem here – apart from the fact that it is obviously a matter of considerable general dispute whether or not many popular doctrines (think of theist/atheist controversies...) can genuinely be believed without particular epistemic failure – is that this criterion for reasonableness is highly exclusionary. Wenar points out, in particular, that religions typically attribute the failure of other- or non-religionists to accept the truth as they see it to such things as 'wordly temptation, demonic possession, divine predestination,' etc. (1995, 41-8). Catholic orthodoxy, in particular, holds that 'those religious truths which are by their nature accessible to human reason can be known by all men with ease, with solid certitude, and with no trace of error, even in the present state of the human race' (44). It does not appear to be only religious perspectives which are likely to refuse to accept such a stringent and broad-ranging requirement of agnosticism about one's opponents' epistemic credentials, however. As my discussion of feminist epistemology in this Chapter demonstrates, there is great contemporary demand for explanations of disagreement about a wide range of politically-relevant questions which could hardly be satisfied with a nod towards the fact that 'the evidence...is conflicting and complex, and thus hard to assess and evaluate', or that 'there are different kinds of normative consideration of different force on both sides... and it is difficult to make an overall assessment' (2005, 56-7). Instead, if anything, many individuals are *especially* likely to make assertions of bias about the large-scale politically-relevant controversies that are Rawls' concern. Such people too, then, will be excluded from the category of the 'reasonable'.

Secondly, Rawls requires that included individuals possess a 'reasonable moral psychology'. The controversial element here is the claim that reasonable people possess certain 'conception-dependent desires', in particular the desire to live up to Rawls' 'idea of

citizenship'. If someone possesses such a desire, they will be motivated to act in accordance with the conception, no matter how much the conception may have evolved, just so long as it has evolved compatibly with its internal logic (Rawls 2005, 82-6)<sup>63</sup>. This may be necessary for Rawls to include because of the potentially radical nature of the revisions which might be required of an individual's practical identity, and which would – on this theory – be expected to flow solely from the intellectual recognition that some hypothetical process had delivered that result. Unlike actual-process discursive views, which can realistically postulate change across the full range of psychological categories which might play a role in ethical thought and action, hypothetical methods can offer only narrow intellectual change based upon logical reasoning. It is only if a distinctly Kantian theory about moral psychology is true (at least of 'reasonable' agents), then, that Rawls' methodology could be expected to reliably issue in altered political *praxis*. As Wenar notes, this is a deeply controversial area. For Rawls' stipulation to be effective, the full range of Humean theories of moral psychology, according to which reason is the slave of the passions, must be false. Likewise, those who agree with Hobbes that humans are primarily self-regarding, and those who agree with Bentham that we are driven solely (or even just overwhelmingly, on occasion) by pleasure and pain, are likewise required to be wrong. (Wenar 1995, 48-51). This second aspect of reasonableness, too, then, is highly problematic as a requirement for inclusion in Rawls' process.

Rawls' political liberalism is a constructivist approach of the third kind. The subjective basis

---

63 The extremity of this is clearest in his fn.31 (2005, 82), where Rawls makes explicit that his claim is that the desiderative power of a conception-dependent desire depends entirely on the logical power of the conception 'to which it is attached'. If a desire is attached to a principle with enormous logical power – one which could, consistently, generate almost any conclusion, given the right assumptions (e.g. about the form of society) – that desire will overwhelm absolutely any other, so long as the conception in question is granted lexical priority by the agent concerned. This is an extraordinarily strong Kantian thesis.

– the politically-relevant psychology of typical contemporary human agents – is to be purged of bad (including, at least potentially, adaptive) elements by a method of hypothetical judgement whose guiding principles are determined by a conception – in this case, of 'reasonableness' – which is not drawn from that subjective basis alone. Instead, Rawls claims to have located the demands of reasonableness, rather vaguely, in 'the public political culture of a democratic society', including 'the political institutions of a constitutional regime and the public traditions of their interpretation...including historic texts and documents which are common knowledge' (2005, 13-14). Although Rawls repeatedly postulates that the core values of political liberalism are compatible with the actual perspectives of most contemporary agents – the fact that he never substantiates this with the requisite sociological evidence notwithstanding – he makes clear that he does not think that just any way of drawing out a methodology for building consensus would do. His choice of methodology is explicitly based upon substantive beliefs about the right way to build such a consensus, which are not themselves presented as being uncontroversially present in contemporary society (Rawls 2005, 134-44). It remains obscure whether, and how, these meta-methodological values are supposed to be objectively justified. This lacuna would provide an entrance point for general anti-constructivist argument of the kind given by Enoch, and discussed in the next Chapter: what, apart from realist considerations, could justify the selection of meta-methodological principles, given that these stand outside the space of construction itself? However, I'll supplement this general avenue for critique with more specific criticisms in the remainder of this subsection.

In line with this general method, a Rawlsian constructivist account of 'dignity' – as I have been deploying the term – would justify the putative elements of that conception in terms of

their being subject to an 'overlapping consensus', between all of the subjects of justice whose lives the conception is to regulate<sup>64</sup>, with this being evidenced by their having been successfully subjected to a process of idealisation designed to ensure Rawls' distinctive kind of impartiality. For something to count as an element of 'dignity', it would have to be shown that representatives in the original position might reasonably be expected to choose to incorporate it, presumably on the grounds that it is necessary<sup>65</sup> for the attainment of social justice, as those representatives would (consensually) conceive of it. Because this idealisation draws on values that are not simply part of the subjective base on which the constructive procedure is performed – the subjective base does not include an explicit, overwhelming injunction against political 'unreasonableness' – the view as a whole cannot be counted as immediately comparable to preferentism, and might be invulnerable to some of the criticisms I have made in this Chapter thus far.

On the other hand, from even the highly condensed accounting I've given, a deep problem can be identified for Rawlsian methodology. It seems unlikely, in reality, that conceptions of impartiality – 'reasonableness' – suitable to do the heavy lifting that Rawls requires would very widely be given the primacy of normative place that is required for this strategy to work, even where they are endorsed at all<sup>66</sup>. If even one of the core elements of 'reasonableness' is

---

64 This is intentionally ambiguous: the results of the constructive procedure only *directly* regulate the political agency of the 'reasonable', for Rawls. However, in an indirect sense, everyone who is subject to political power (that is, absolutely everyone) would be regulated by the conclusions the procedure produces, if (as Rawls intends) the 'reasonable' possess at least the bulk of such power.

65 Thom Brooks has provided an argument that a minimum threshold level of a range of classic capabilities can, in fact, be justified using Rawlsian considerations alone (Brooks 2015). Nussbaum, in most of her later work, also says things which suggest that she thinks this is plausible, although the actual methodology she uses is significantly different to Rawls'. See especially Nussbaum 2006a, for explicit consideration of her similarities and differences to Rawls.

66 Most of Rawls' discussions on this issue are focussed on the question of whether an effective, stable overlapping consensus is *possible*, rather than whether it is realistic in a more immediate sense. However, only the latter question is straightforwardly relevant to the question of methodologies for thinking about global justice here and now.

rejected by, or fails to hold of, an agent, they will not count as capable of such impartiality, and Rawls' methodology will not be available to them. This would not be a deep, philosophical difficulty – as opposed to a brutally practical one – if it were not for the fact that political constructivism leaves itself few or no resources to engage with those who are thus categorised as 'unreasonable'. If – or rather, when – faced with someone who explicitly rejects the 'political liberal' construal of the political problem, or someone who more meekly *wonders* whether this way of setting things up does not amount to targeting us at justice so much as giving up on its possibility, the Rawlsian constructivist appears to be left with little to say. What Rawls thinks his methodology does allow at this point is argument 'by conjecture', in which a 'reasonable' person attempts to argue that a (putatively superficially) 'unreasonable' individual does in fact endorse the values that would enable political liberalism to have rational purchase on them (1997, 786-7). However, this is little more than fishing for common ground where none may exist, and can only be rationally arbitrary from the constructivist perspective itself (it is not, as Rawls acknowledges (786), a form of 'public reasoning', the kind that is amenable to objectivity as he conceives it).<sup>67</sup> After all, there is no reason, other than Rawls' frequently asserted but sociologically unsupported *confidence*, to think that practitioners of constructivist methods would know the minds of the 'unreasonable' better than they do themselves<sup>68</sup>. Apart from this thin, arational form of engagement, the Rawlsian can do nothing but ignore such respondents, and hope that they do not turn out to be so numerous as to upset the feasibility of the project.

---

<sup>67</sup> See also Besch 2013, 64-65, here.

<sup>68</sup> The situation is worse for the Rawlsian than it is for practitioners of Marxist, or Freudian, or other programmes incorporating ideas like 'false consciousness', which are often accused of the same hubris. At least those bodies of thought include diagnoses of why individuals might be systematically deluded about (e.g.) their interests, values, or practical commitments. Rawls appears to offer nothing which might play such a role.

As Wenar's paper (1995; cf. Hurd 1995, 819-23; Friedman 2000, esp. 28-32; Besch 2013, 65-7) suggests, by building such strong restrictions into his conception of 'reasonableness', Rawls is essentially guaranteeing that he can only – without hypocritically violating his own constructivist *bona fides* – preach to his (much-diminished) choir. Ross Poole has pointed out that this circumstance is also paradoxical, from Rawls' own perspective: he needs his audience to be sufficiently committed to the political-liberal project to give the outputs of his idealising methodology precedence, and yet the methodology itself functions by abstracting away from the extant sense of justice and/or other-concern that must underlie that commitment (Poole 2005, 63-8). The result is that the methodology is unlikely to be able to do any real work – those who would already be inclined to take part in Rawls' project may well be alienated from its prescribed course; those who would not can be given no reason to participate in the first place.

For them, as Poole says,

Rawls' assumption that a sense of justice already has a cultural existence occupies the place that should be filled by an argument why people should be moved by, and give priority to, considerations of justice [as he conceives them]. (67)

These tensions, once their full depth is recognised, make the entire project of political constructivism presented in *Political Liberalism* and other late works very likely to be a failure on its own terms. As Heidi Hurd says, Rawls' exclusions comprise not only 'hate-mongering members of fringe cults', but

libertarians and classical liberals... act consequentialists... egoists, altruists, Catholics, Protestants, hedonists, perfectionists, communists, socialists, feminists, communitarians, and religious fundamentalists, who [all] for one reason or another part ways with the fundamental liberal tenets that define Rawls' reasonable person ...  
Once we realise the scope of the audience that Rawls is *not* addressing, one has to

wonder what remains of...[his] project. (821-2)

Hurd's focus is on the pragmatic absurdity of a claim to be justifying a form of liberalism, when one has excluded many, if not all, non-liberals (and – for that matter – even many liberals) from one's argumentative targets, but a more pressing tension for my purposes is evident on the level of methodology itself. Rawls' exclusions on the first-order level involve him in something of a performative, dialectical contradiction, but when the focus is instead on the constructivist nature of his methodology, he is involved as well in conceptual contradiction. The nature of his political constructivist project is to draw out ideas that are supposed to be implicit in the accepted institutions and/or common values of the *polis*, and thus supply constructivistic objectivity to norms which can gain commensurately general allegiance. But, if his audience is this limited, the scope of the norms that could gain objectivity will not be remotely general; if, instead, he *were* to direct his arguments to the *polis tout court*, any objectivity they might have could not be construed as constructive, at least in the way he proposes, because any basis it might have in antecedently endorsed ideas would remain entirely obscure. Rawls' 'political constructivism' appears to be either not political, or not a form of constructivism.

The second major problem that I'll discuss concerns the limiting effects of Rawls' restriction of political constructivism to the 'basic structure'. As I have previously noted, Rawls' constructivism limits its target to what he calls the 'basic structure' of society. Some move such as this seems to be necessary for him, in order for his claim to be a constructivist to have any plausibility: it is only because the range of discourse falling under his account is limited that he is – even potentially – able to avoid making controversial claims about such a wide range of things that it would become entirely unclear how he could still aspire to ground

political truth in the existing worldviews of the general public. As this is a private/public distinction, it necessarily falls under the warranted suspicion I argued in the previous Chapter (§1.3): it is not clear what principled reason can be given for drawing the line between public and private in any particular place – even if a reason can be given for drawing it at all – and, more importantly, private/public distinctions have a notorious history of being used to cover up injustices by relegating them to the 'private' sphere. There, I noted Okin's classic work on justice within the family, as an attack on Rawlsian restrictions to the basic structure. Here, instead, I'll draw out a more recent critique by Andrew Koppelman, which supports the same conclusion: that restricting the space of thought about justice to the basic structure is likely to produce extensionally inadequate results, and will do so in a way that reveals a bias against the concerns of oppressed or marginalised people.

Firstly, Koppelman examines the case of sexual freedom for sexual minorities. He notes that, since it is not clear that sexual freedom is a prerequisite of the development of the capacities necessary to engage on equal terms in the democratic process of reasonable disagreement with others, nor is it necessary for the development of the ability to use deliberative reason in formulating an individual conception of the good, it is not clear what special grounds there are for protecting such freedom from a Rawlsian point of view. If one of these 'two moral powers' were in danger (Rawls 2005, 18-19), political liberalism would prescribe protection, because the capacity for use of such powers is essential to the fundamental political freedom and equality that Rawls thinks central to the values underlying it. If one of these fundamental publicly recognised interests is not at stake, however, the interest of (e.g.) a gay couple to have the liberty to have sex with one another is on a par with the interest of others in living in a society where sodomy is subject to public sanction. For Rawls, there is 'a general

presumption against imposing legal or other restrictions on conduct without sufficient reason... But this presumption creates no special priority for any particular liberty' (2005, 292). It is similarly difficult to see how any special rights to enjoy specific forms of relationship could have a particular justification under political liberalism, since incorporating these would be accorded great importance by some general 'comprehensive' worldviews, but little or no importance, or actively negative importance, by others. (Koppelman 2009, 463-6).

Secondly, Koppelman notes that in the case of female genital mutilation (FGM), too, Rawlsian constructivism has difficulty accounting for the injustice of the practice. Koppelman notes that many practitioners or supporters of FGM do offer (putative, according to their comprehensive conception of the good) reasons for the practice. It is bound up with widespread (at least in certain locations) conceptualisations of hygiene, womanhood, and marriageability. There are clear, countervailing reasons, of course, many of which could legitimately be appealed to in making a case for prohibition. However, because the harms that FGM does, which generate these reasons, do not (at least *prima facie*) affect the two moral powers, they are not relevant to considerations of Rawlsian justice. What is more, because the power of legislatures, insofar as they do not violate the principles of justice for the 'basic structure', is unconstrained by Rawlsian constructivism itself, it is even possible that forcible FGM could legitimately be mandated in a polity. Koppelman considers a number of Rawlsian counter-arguments to his thesis. Against the complaint that FGM (if performed on children) constitutes child abuse, he notes that Rawls has no consistent means of justifying a conception of 'child abuse' which would go beyond abuses that preclude the development of the moral powers. As examples of adult women who underwent FGM, and who have achieved clear excellence in the use of their moral powers clearly imply, however, it is

implausible that FGM will always have this effect. Against the complaint that FGM damages health, he notes that health is a contested concept between comprehensive conceptions that Rawls apparently must regard as 'reasonable', including some which regard genital integrity as irrelevant or even contrary to health, and that no reasonably disputable version of such a concept may be built in to the constructed account of justice. Most importantly, against the response that FGM (at least, as a structural practice) violates gender equality, Koppelman notes that the only species of gender equality that Rawlsian justice can coherently require is one wholly concerned with gender-based *political* inequality; equality as a citizen. It is far from clear that practices such as FGM necessarily cause political inequalities of this kind. (2009, 466-74).

Koppelman then contrasts the strictly Rawlsian response with that of Nussbaum, who, while she draws extensively on Rawlsian-sounding concepts, and is, as I brought out in the previous Chapter, officially (if equivocally) a friend to constructivism in practice, does in fact bring in non-universal value judgements and interpretations of relevant concepts when formulating her arguments about these matters. As Koppelman notes, this may be consistent, because Nussbaum settles on a much weaker interpretation of the requirement for compatibility with 'reasonable' doctrines than Rawls does. Unlike Rawls' formal and strict criteria for 'reasonableness', Nussbaum requires only that claims about justice be susceptible to agreement between the holders of *many* diverse perspectives; she is not forced, as an unequivocal political constructivist must be, to accept all which meet some determinate criteria. (Koppelman 2009, 480-1). This suggests that an accurate account of the wrongs involved here can only be achieved by abandoning the most distinctive idea of Rawls' constructivism.

Koppelman's general point is that the Rawlsian restriction removes issues such as prohibitions on the freedom of sexual minorities, and the prohibition of gender-based (or – for that matter – gender-neutral) involuntary genital cutting, from the conceptual sphere of justice. It does this for different reasons in different cases, but the overarching problem is the same: the restriction to the basic structure, and the ban on 'private' reasons that is its necessary corollary, deprives the political liberal of the resources needed to articulate what is unjust about certain cases. Because the language of justice is not only necessary for accuracy, but equally for practical political organising and contestation<sup>69</sup>, adopting a political constructivist methodology would not only fail many oppressed people on a theoretical level, it would also be likely to harm their struggles for liberation.

It will be helpful, in closing, to consider explicitly a key argument that Rawls would likely level at someone who argues against the species of idealisation in methodology that he proposes, as I have here. Perhaps the most forceful argument that Rawls offers for his way of thinking about the methodology for thought and discourse about justice rests on its intended role as the basis for political order and stability. Rawls attempts to solve two problems simultaneously: the problem of securing social order by 'reconciling' people to the basic institutions of their society, and the problem of actually *justifying* those institutions and (over time) rationalising any changes to them without substantively transforming the people concerned. He thinks that performing the first without the second is unpalatable, while the

---

<sup>69</sup> There is a connection here to the more general feminist work that I cite in this Chapter, and elsewhere in the thesis. This suggests that there may be many instances in which reasons which have been rendered invisible to many by structures of oppression can come to be 'public' and shareable only by transformative experiences that (however partially) eliminate that oppression just *per se*. However, the pursuit of such transformation is precluded by restrictions, within the space of discourse about justice, to what can – at a time – be publicly shared. See especially the next section for relevant theory.

second without the first is impossible, since a social order that was not stable could not be justified. (2005, 133-57.) If what I have argued in this section is correct, however, this aspiration is likely to be a vain hope, and, unless individuals can be significantly transformed, the work of maintaining political stability may have little to do with public justification to one another. This removes much of the force of this key Rawlsian argument: if something probably cannot be done, the failure to do it is unproblematic.

Instead of Rawls' strict and exclusionary understanding of 'reasonableness', there are a number of alternative possibilities for thinking about reconciliation and its relationship to justification proper. Most of the authors who have taken a similar critical line to that given here propose a more minimal concept of the 'reasonable' which avoids the entirely amoral 'mere *modus vivendi*' alternative with which Rawls constantly contrasts it (Wenar 1995, esp. 61-2; Koppelman 2009, 479-82; Klosko 1997, 637-46). In all of these alternative proposals, additional kinds of controversy are permitted into the space of valid deliberation about the political, to varying degrees. As some of these thinkers recognise, this is likely to require (if not constitute) an abandonment of Rawls' constructivist program, along any number of dimensions. For Koppelman, this departure from constructivism is necessary because additional normative content is required in order to secure commonly-agreed standards of justice (i.e. the prohibition of FGM), but this will come at the cost of explicitly excluding some (those who demand FGM) from the relevant deliberative space on transparently first-order grounds (2009, 476-9). For Wenar, it is essential because the endorsement of constructivism about the political is itself an exclusionary threat to stability (1995, 52-7). Alternatively, for Besch, it is because views about the proper scope of justification cannot, without brutally begging the question, be given any constructivist rationale: every form of

constructivism simply assumes some constrained scope. Instead, for this question at least, a substantive, perfectionist answer must be given (2013, 68-70). The precise details can be left aside, although all of these lines of argument are consistent with my broader proposals in the thesis, because my purpose here is merely to demotivate Rawls' way of thinking about the issue of stability through justification, which is so tightly bound up with the motivation for his constructivist methodology; I do not need to provide an alternative account of my own.

Many non-Rawlsian political philosophers may continue to reject a mere *modus vivendi*, and seek, however partially, to satisfy Rawls' hankering after stability for the right kind of reason, for example on the grounds that a more satisfying kind of social unity is often possible (e.g. Koppelman 2009, 481-2). In any case, whether or not this aspiration is a noble one in principle, nothing can be gained by the Rawlsian pretence that many deep and lasting disagreements about normative matters can be hermetically sealed off from questions about justice by castigating many participants as 'unreasonable', and exhorting the rest to use only an impoverished and potentially biased set of political concepts. There may be hope for the actual-world construction of genuine agreement here and there, which might secure at least local political stability for reasons other than brutally Hobbesian fear of chaos or unreflective complacency. But this goal is unlikely to be served<sup>70</sup> by a rhetoric which pretends that those who seek different political ends should be opposed because they are fundamentally uncooperative, dogmatic, or malicious; rather, they should be opposed on the honest, realism-friendly ground that they are *wrong*.

### §3 The Locus of Epistemic Values; Contingent Virtue

---

<sup>70</sup> Compare especially Bohman and Richardson 2009, 265-74, which argues that the true grounds for hope of agreement in politics lie in the scope for actual common ground, and not on hypothetical consensus.

Finally, in ending this Chapter, I want to do a little more to reinforce my claim that culture-laden realism is far from being an alien presence within feminist epistemology. To that end, in this section I shall discuss work by Laura Ruetsche and Rebecca Kukla, which utilises ideas, within a feminist ethical/political epistemology, which are especially congruent with culture-laden realism, and particularly with the McDowellian view I bring out in the next Chapter. Before I do so, I shall also discuss one controversy within feminist epistemology, which concerns the correct location of epistemic values (such as objectively justified belief). This controversy will be important in what follows, because one answer to this question sits better than others with the version of culture-laden realism I endorse.

The controversy in question turns on whether or not we should regard individuals or groups<sup>71</sup> as the primary objects of epistemological theory. Many feminist epistemologists have argued that, because of the ubiquitous relevance of social-relational factors to epistemic standing, and because (on many views, at least) this implies that no individual can be epistemically self-sufficient, we must regard the community as the level at which epistemic explanations occur, and as the locus of knowledge itself. I understand this question about the individual-group relation in epistemology in a parallel way to the understanding I've previously adopted about that relation within normative political philosophy<sup>72</sup>. Here, as there, it is open to us to combine a 'normative individualism' – the view that the basic locus of relevant values is the individual – with an anti-reductionist social ontology which acknowledges the many

---

71 I am ignoring a third category of social-epistemic stances on the individual-group relation, which hold that individuals within specific relational settings are the primary base for epistemology (as with the community model, this is ambiguous between normative and explanatory interpretations of the relevant 'base' notion). Most importantly, I do not discuss Grasswick's own 'individuals-in-communities' account (Grasswick 2004). I think that Grasswick's (or some other) relational account is likely to be better at doing the explanatory work than a strictly community-centred view, but my key argument – that normative and methodological individualisms are separable, and only the latter are problematic – will apply to relational views no less than group-based ones.

72 See §1.3.3 in the previous Chapter.

(including some constitutive) ways in which social location affects how individuals stand with respect to those values. Thus, although justification (or objectivity, or knowledge) *depend* upon social facts and an individual's social history – that is to say, the community view is correct that individuals are not self-sufficient – the justification itself (the epistemically valuable state) is primarily a state of individuals.

Fortunately, it is not clear that this claim contradicts very much of the community-centric view. For example, much of the discussion of community-first views about the locus of knowledge (e.g. in Grasswick 2013) does not pertain to the questions I am concerned with. For example, Grasswick identifies the view that “communities can be understood as knowers in the sense that a communal context is required for knowing. Individuals know only within communities.” However, this is not incompatible with the epistemic form of normative individualism (and is clearly congruent with culture-laden realism as well). Her overall summary is as follows:

Feminist social epistemologists who argue for a community model of knowers do not deny that individuals also know. However, their accounts challenge the possibility of understanding individuals as knowers in isolation from their communities, and press for a better understanding of the relative roles of individuals and communities in knowing. For example, communal accounts of knowing suggest that the epistemic responsibilities of individuals will need to be understood in relation to their communal memberships, which set limits on the conceptual resources and epistemic tools individuals have available to them. (Grasswick 2013)

On the face of it at least, this is perfectly compatible with normative individualism in epistemology: the complaints raised against traditional epistemology here argue that it is impossible to adequately explain how knowledge (or justification, etc., *mutatis mutandis*) comes about within the confines of methodological individualism. But, as I argued for the ethical/political case, normative individualism need not imply its more general methodological sibling.

What is more, we have some good reasons to continue to view justification (etc.) as located at the individual level. Most importantly, a central evaluative function of justification is to distinguish interlocutors as having epistemic standing relative to one another; it is hard to see how a view which regards communities as the primary locus of justification will be able to tell a compelling story about the conferral of such standing. At least, an individually-based account will be more straightforward. In addition, it is not obvious how to make sense of the role of epistemic norms (asserting the true, not believing the false, etc.) in guiding individual action, if those values apply to individuals only derivatively. With an individual-based account, we can more easily make sense of the role of such values. Why should I listen to the testimony of an agent? Because I (an individual) am justified in believing that she (an individual) knows something that I do not. Even if, to explicate my justification, and/or her knowledge, we have to refer to relational or community-implicating facts at every stage, it remains the case that the justification is mine, and the knowledge is hers. If this is correct, the ground is clear for a conception of epistemic standing which incorporates extensive reference to relational factors, but is centred on an individual-level concept: the concept of virtue.

In the remainder of the Chapter, I begin to introduce the notion of virtue I shall be supporting, with reference to the feminist-epistemological work of Kukla and Ruetsche. This work is distinguished by its uptake of a (roughly) Aristotelian notion of virtue alongside an acceptance of a significant degree of *contingency* in virtue (so construed). I shall concentrate on their joint paper, 'Contingent Natures and Virtuous Knowers: Could Epistemology be "Gendered"?', but many of the same points are present in their sole-authored work<sup>73</sup>. The

---

73 Cf. Ruetsche 2004; Kukla 2002.

central claim of the paper is that 'there exist some individuals for whom being gendered<sup>74</sup> in the way that they were had some relevant effect on their contingent history which in turn altered their second nature<sup>75</sup> epistemic capacities.' (Kukla and Ruetsche 2002, 396)<sup>76</sup> Later, they gloss this as the more precise claim that 'a gendered standpoint conditions a second nature yielding access to results duplicated neither by [purportedly location-neutral epistemic rationality] nor by other second natures as they are currently constituted'<sup>77</sup> (407). Although theirs is a nicely detailed account, it suffices for my purposes in this section to note a few general features which demonstrate the compatibility between feminist epistemology and an Aristotelian development of culture-laden realism. Firstly, they implicate a notion of virtue as a habituated capacity, with cultural influences playing a role in its development (esp. 413-18). Secondly, they (unsurprisingly) develop their account so that locations salient to feminism are included – gender, race, class, etc. Thirdly, they make clear that there may be some second-natural positions such that no rational process (i.e. of argument) can bring the relevant agent to a state such that they can access the (e.g. gendered) contingently-inaccessible reasons (410-12). Finally, they explicitly regard the epistemic benefits of the contingent virtue that they discuss as involving a relation to objective reality (409-10; cf. especially Kukla 2002, 334-8). These, taken together, constitute a version of culture-laden realism which is well-placed to incorporate feminist-epistemological insights; it is also a view which, as will become clearer in the next Chapter, is extremely close to that which I favour<sup>78</sup>.

---

74 It appears that Kukla and Ruetsche would generalise the account to social (including, presumably, more generally cultural) locations beyond genderings.

75 I discuss the notion of second nature, as it figures in McDowell, explicitly in Appendix B. At this point, an intuitive understanding will suffice.

76 Much of the (uniformly excellent) paper runs through several different interpretations of this idea. I discuss only the last.

77 This element is included to avoid 'epistemic separatism', either as the view that gendered reasons are gender-relative reasons, or as the view that gendered positions (and the reasons to which they give access) are permanently beyond the reach of differently-gendered others (Kukla and Ruetsche 2002, 403-5). The first is implausibly relativistic, the second is in tension with its own motivations – if gender is socially contingent, it should be possible for the perspectives it involves to be shareable.

78 This is not especially surprising, since they explicitly (e.g. 392-3) draw on McDowell's work in developing

To sum up the Chapter, there are two key reasons to reject constructivism as a meta-level frame for the methodology of the capability approach. Firstly, at least most forms of constructivism either motivate, or are motivated by, proceduralisms – such as Baber's preferentialist response to preference adaption – which are all, so far as we have seen, extensionally inadequate. Even if such views had meta-level advantages (metaphysical parsimony, simplicity, etc.), their inability to capture which desires are problematic and which not is devastating. Secondly, constructivism also has meta-level disadvantages, because it cannot adequately make sense of the epistemological feedback between procedural and substantive judgements. Since an appeal to such feedback is the only way to resolve the problems of extensional inadequacy, I conclude that constructivism is not a promising methodological position. Even in the face of arguments such as Jaggar's, we have reason to seek a culture-laden form of realism.

---

it. Cf. also Code 2014, 188-199, and Fricker 2003 (which provides a link between McDowell-type virtue-centric epistemology and her broader later work on testimonial (2007) and hermeneutical (2006) epistemic injustice), for other especially friendly approaches.

# Chapter 4:

## Shapeless Virtue

Throughout much of the previous two chapters, I have issued a series of promissory notes for an account of objectivity, about the normative topics that capability theories concern, that is capable of meeting a number of desiderata. It is worth collating these here. Firstly, we need a concept of objectivity worthy of the name: we need to be able to account for some people knowing something that others don't, in virtue of something different about the way they relate to the parts of the world they have beliefs about. This concept needs to make sense on the meta-level, and discussion concerning it cannot be conducted in isolation from meta-level discourse. Specifically, I have given preliminary reasons to doubt that a constructivist meta-level approach is likely to be plausible here, so we should instead seek some form of realism<sup>1</sup>. Secondly, we need to be able to make sense of the specific contours of the concept. If we believe that objectivity about some fact can depend upon cultural location or social position, for example, it needs to be clear how this can be the case coherently with a plausible meta-level framework. I have argued throughout much of the last two chapters that the requisite concept of objectivity will have features such as this – under the banner of what I've called 'culture-ladenness' – so these must be accommodated. Thirdly, we must be able to claim that, for some kinds of normative discourse at least (i.e. that involved in the capabilities approach

---

<sup>1</sup> I am not considering non-cognitivist or error-theoretic meta-ethical theses here. This is defensible given that my main dialectical opponents seem to be committed to there being normative truths. On the other hand, if constructivism is compatible with expressivism, say, (see e.g. Lenman 2012), or if expressivists can 'quasi-realistically' accommodate almost all ordinary phenomena of normative thought and discourse (Blackburn 1984 and 1993; cf. Dreier 2004 for the now-classic internal threat to this paradigm), or perhaps even if morality is thought to be a *useful* fiction (cf. Joyce 2006; Kalderon 2005), it might be possible to motivate at least many of the kind of things that constructivistic normative thinkers say coherently with a non-realist theory.

to global justice), the answers for which we claim objectivity hold universally. Even if the perspective of objectivity about some question is culturally-laden, as I have suggested it will be, this cannot entail that the claims it endorses are true only for those who have been similarly enculturated. Enculturation-dependent objectivity cannot be allowed to collapse into cultural relativism. Finally, it may help to identify some account of how the target discourse relates (semantically, metaphysically, epistemically etc.) to other discourses (natural-scientific discourse, for example)<sup>2</sup>. This is what is generally referred to within metaethics as the broad debate about 'naturalism'.

Some of these desiderata will be addressed in later chapters: I will not make any serious claims about naturalism, at least, here<sup>3</sup>. Others can best be dealt with now, however, by explicating and defending a general view of metaethics that can satisfy them. I find this account in the work of John McDowell, and it is his approach that I shall draw out here. Two large caveats apply, however. Firstly, I am not concerned with pure exegesis – my aim, as I have said, is to produce a plausible basis for a capability approach. Any errors of interpretation that there may be in the following should not damage this aim significantly<sup>4</sup>. Secondly, I do not have space to consider every criticism of McDowell, nor can I adequately trace the connections between areas of his philosophy, or fully describe many aspects even of the proposals I do rely on. Setting out on any of these tasks would quickly turn this into a Thesis about McDowellian metaethics alone, with any hope of contributing to the capability

---

2 Specifically, this might help because the (perhaps) more straightforward relationship that constructivisms might posit between the ethical and non-ethical levels might provide motivation for constructivism. Cf. fn.15 in the previous Chapter.

3 I discuss neo-Aristotelian 'naturalism' in Chapter 5. However, the full meta-level interpretation of that approach is of little significance for my purposes in this Thesis, and so I cannot devote significant space to it, although I concede that a genuinely complete framework of the sort I'm proposing would provide more on this score. To that end, Appendix B contains some argument on the topic, which may help to illustrate the direction of travel I think most plausible.

4 Of course, I do not think that the approach I describe here is correct *because* McDowell is (for the most part) its origin!

approach lost. Rather, my objective is relatively modest: I believe that McDowell's metaethical approach is independently compelling, but I also claim that it is an excellent complement to a capability approach, and it is this latter belief that I hope to support here. As such, I shall only engage in general defence of the McDowellian approach I favour when faced with problems so damaging that, if they were not responded to, would make this metaethics clearly inappropriate as a partner for a capability approach.

The structure of this Chapter is as follows. In Section 1, sub-sections 1.1 and 1.2 describe important elements of McDowell's general philosophical output that are especially relevant to the metaethical claims that are my focus. §§1.3-5 then explain McDowell's central idea: that moral facts can be understood as *sui generis* worldly dispositions. In so doing, I also bring into greater focus the similarities and differences my view has with constructivism. Section 2 turns to more general meta-level issues, such as moral psychology and epistemology, bringing out a number of respects in which McDowell's approach may complement a transformation of the way we think about ethics on a more practical level, and reiterating some points of agreement with themes from recent feminist work. §2.1 focuses on moral psychology, noting two ways in which McDowell's highly flexible approach to these issues brings advantages to the overall picture. §2.2 briefly engages with the debate about reasons internalism, and argues that here too a McDowell-inspired approach pays dividends in the context I'm working in. This section also provides the basis for an important link which I identify between different kinds of reasons claims and the controversy about ideal and non-ideal theorising in political philosophy, which I pursue further in Ch.6. Finally, §2.3 grapples with what may be the most interesting potential problem for this approach: in linking moral knowledge to a (set of) cultural position(s), while linking what moral knowledge is *about* to humanity itself, it may

render relativism not so much *false* as *unintelligible*. This may seem unsatisfactory, so I engage in further explanation to assuage worries in this area.

## §1 Secondary-Quality Realism

### §1.1 The General Picture: Idealism, Realism, Quietism

McDowell's aim, throughout all of his work<sup>5</sup>, is to show how we can be satisfied, when faced with an apparently pressing philosophical problem, without having to radically revise the viewpoints from which we begin. Rather, problems can be shown to be illusory, by clarifying our practices and discourses to show that no revision is required. This methodological stance is called *quietism*. (McDowell 2009e). It is important to recognise that being a quietist about one region of philosophy (broadly conceived) does not commit one to quietism about all such regions: in particular, first-order normative discourse still contains problems that will require substantive, non-quietist, solutions. However, large areas (at least) of contemporary metaethics may be ripe for such treatment, and McDowell's proposals for this region of philosophy should be seen in this light.

Within this expansive metaphilosophical framework, McDowell's most relevant theses for my purposes are also his most general. It will be necessary to explain some of these broad interventions to provide a context for the specific, metaethical claims. I shall begin with a summary of the central relevant conclusions of *Mind and World*; I do not have space to recount the arguments for these conclusions. Firstly, McDowell claims that the world is

---

5 Notice that this formulation does not assert that McDowell is committed to this sort of aim across all of 'philosophy' (for at least most construals of that subject): he has explicitly disavowed that view (cf. McDowell 2009e, 366-7). Those first-order areas of philosophy for which quietism seems most implausible – ethical, political, aesthetic, scientific-methodological, first-order epistemic, etc. – are ones in which he has done little or no work.

constitutively apt for conceptualisation. The stuff of which the world is composed must be such that some concept-user can place it within their cognition; the world must be *thinkable*. (1994, esp. 24-45)<sup>6</sup>. Secondly, the concepts that the world is apt for must be *ours*, in some (very weak) sense. At least, it must be the case that there is a rational route between (some of) the concepts that we now possess, and the concepts within which the world itself could figure. Assuming that all epistemic rationality is in some sense Neurathian, this rational route will be available just in case our current concepts allow us to know things about reality to any extent at all; that this is, in fact, the situation that we are in then seems to be a good candidate for something presupposed by ordinary human cognition and discourse.<sup>7</sup> This thesis, which might be considered a weak form of idealism, does not require that any of our *current* concepts can capture any aspect of the world fully. All it requires is that there be some possible conceptual community (some group of beings<sup>8</sup> that share some concepts), that is continuous with the contemporary one in the sense that one could move in a Neurathian way from our worldview to theirs, and whose concepts could do so.<sup>9</sup> Thirdly, whenever people have veridical experience of something, there are concepts at work in that experience (1994, esp. 3-23 and 46-65). As a corollary, coming to *have* experience of something will require the possession of some relevant set of concepts<sup>10</sup>. The immediate benefit of this idea is that it

---

6 McDowell's view on this has changed, since *Mind and World* was published. Originally, he fleshed out the idea of the conceptuality of the world in terms of *propositionality*, with an associated focus on what can enter into language and discursive thought. He has since relaxed this thesis, but I can't discuss any impact this might have on my argument here. See McDowell 2008a.

7 Along the lines of a Wittgensteinian hinge proposition, perhaps (Wittgenstein 1969), or a constitutive norm of thought itself *à la* Davidson (2001); cf. McDowell 1998g.

8 Technically, one being (a 'God' -type character) would do, but that is enormously implausible given the social way that knowledge arises in our species (our only solid reference-point), and, indeed, given the physical constraints of the universe.

9 This refers to the roughly Davidsonian argument that McDowell gives in *Mind and World*. (Cf. also McDowell 2009a, 134-151). As far as I can see, this very weak interpretation of the maximal conceptual community is all McDowell *requires* to do his work. However, there is some reason to think that McDowell holds a more strongly anthropocentric interpretation (cf. especially Haddock 2008). Even if this is so, it should not affect my argument.

10 Note that this doesn't entail that all knowledge requires some specific course of education, or the understanding of some theory; some concepts might either A) be innate, or B) be so fundamental to engagement with the world that they would be a part of any pattern of agential development whatsoever.

coheres with the overwhelmingly plausible thesis that the experiences which are epistemologically important for ethics will be so only because they are infused with certain acquired concepts – clearly people do not recognise the humiliation of others, for example, using the simple application of visual attention alone. Rather, the deployment of some conception of social standing (*inter alia*) will be necessary in order to be able to register the salience of those features present within (e.g.) the visual field that are criteria for humiliation.

An important pattern is visible here (and elsewhere in McDowell's work). He aims to show how the essential involvement of human capacities in our epistemological practice does not preclude its being about a reality that those capacities do not create. He aims to reconcile very weak (arguably, even trivial) versions of realism and idealism<sup>11</sup>. In doing so, the problems that (putatively) more substantial realisms or idealisms concerned themselves with should no longer seem pressing: we already have everything we need<sup>12</sup>. As I shall show, this same impulse is at work in McDowell's metaethical work; that this broader background is available helps to make that specific intervention more compelling.

A second aspect of McDowell's general work that is particularly important here also concerns an attack on a certain kind of theoretical sealing-off of the human mind from the rest of the world. As I have noted, McDowell's view involves a kind of very modest idealism quite generally, which erodes the presupposition that the concepts that are at work in human worldviews must form a markedly different kind to the facts that those worldviews are about, such that we could thoroughly separate out internal mentality from worldly materials. But

---

11 Haddock 2008 puts it this way, although he sees this resolution as in tension with a similarly weak (or trivial) thesis of *naturalism*, that he also identifies in McDowell. I cannot respond to this here.

12 Compare Naomi Scheman's allegedly Wittgensteinian invocation of 'real needs' in attacking 'metaphysical' realism in her (2001). The turn back to real needs is congruent with quietism, but her rejection of unconstructed reality is nothing if not a substantial, revisionary thesis. Cf. McDowell 1998a, esp. 260-2.

there is also a more specific commitment, to a variety of mental content externalism<sup>13</sup>: in this case, the view that the psychological states involved in virtue (epistemic, ethical, etc.) cannot be individuated using 'inside the head' ('internal') resources alone. This view is explicitly developed for the epistemology of non-ethical perception<sup>14</sup>; put simply, McDowell holds that someone who sees (e.g.) a bent straw in a glass, and someone who sees a straight straw that they consequently judge to be bent because the glass contains some water, do not have identical experiences of the world<sup>15</sup>, even if the only way for a third party to tell that this is so is to look to the broader environment outside their heads. (1998b; 1998c; 2011, 36-53). As McDowell puts it, there is no epistemic 'highest common factor' (1998b, 385-394) between illusory (or hallucinatory, or – perhaps – otherwise misleading) experiences and veridical ones, because veridical experiences are genuinely world-involving in a way that the others are not. As will become clear, this is an extremely useful epistemological feature, since it allows us to make sense of the evidential difference between a virtuous response to a situation and a non-virtuous one, and also forces us to grapple with the virtuous condition in all its contingent complexity, while still thinking of it as a unity. Even better, it does so in a way that does not require us to give any strictly *independent* motivation for the judgement that there is a difference between one experience and the other – we presuppose things about the state of the world and our relation to it in describing the experience as veridical, but we might also have to rely on that very experience in justifying these more general claims about the world. This is an excellent fit for culture-laden realism: our methodology must incorporate substantive ethical content with the thought that this is truth, using this to exclude (as

---

13 It may be arguable whether McDowellian disjunctivism really does involve mental content externalism, but I'll ignore this exegetical question here, in line with my general strategy in this Chapter.

14 As well as in giving a gloss on the Wittgensteinian notion of 'criteria' as it figures in e.g. the debate about knowledge of other minds. The classic source for both is McDowell 1998b.

15 Putting it this way assumes that the water itself is not independently visible to the latter agent at the time the mistake is made.

epistemic peers) the positions from which that content is inaccessible; this epistemological circularity is well-justified on realist grounds (and is unavoidable in any case). Just as – quite generally – we need to deploy all our substantively-laden thought about the world in order to separate epistemically-weighty perceptions from mere appearances, so we need to deploy substantive ethical thought in order to pick out the objectivity-involving deliverances of virtue. What enables the state of virtue to be visible as a unity, on this view, will be exercises of practical reason itself.

## §1.2 The General Picture: Rules and Rails

The second of the more general areas that it is important to describe concerns McDowell's response to the rule-following considerations that arise from the work of Wittgenstein. As before, I cannot go into detail concerning the motivations for this debate, bring out the nuances of McDowell's response, or consider criticisms. The rule-following considerations are a connected set of issues, largely dealt with in the philosophy of language, that have to do with the nature of rules. Rules are essentially normative, and they can be extremely complex (perhaps, in the case of some e.g. mathematical rules, infinitely so). On the other hand, rules can be followed by a human agent, and this sort of internalisation brings in a psychological component: necessarily, people who have internalised a rule are disposed to respond to events in accordance with that rule. The problem is that the latter considerations can seem to be in tension with the former. For one thing, the normativity of rules ensures that there are always, in principle, further cases to which they could apply (even where the rule itself doesn't contain the concept of infinity); there is always something that would constitute 'going on in the same way'. What is more, it always has to be possible to make sense of the idea that one could go wrong – if one could not go wrong, it is not clear that one can get things right, and

then the rule would no longer be normative. However, what dispositions hold of a particular person is a determinate matter of fact that can seem to have nothing inherently to do with the nature of the rules that they follow. This gives rise to several problems. For one, how could anyone know that a person understood a rule, when all that they could observe were the products of an individual disposition that the rule would inherently transcend (Kripke 1982, 22-37; McDowell 1998a, 226-8)? Alternately, how can rules be normative, if they can be entirely captured in the finite psychological dispositions that people can observe? (McDowell 1998a, 246-9).

McDowell's response is that we can dissolve any apparent problems that may arise here so long as we reject an assumption that is at work in the motivation for the problematic. Sceptics about rules, such as Kripke, assume that we have to be capable of characterising the personal dispositions that command of a rule involves independently of the content of the rule itself. The thought is that if people are latching on to a pattern in learning to follow a rule, this pattern must be discernible to those that do not have a grasp of the rule already. There must be a pattern in the world that the rule relies upon, and that people then internalise. Rules are thought of as something like 'rails', on this picture. They are already there to be followed, and rule-following individuals attach themselves like railway carriages. This image allows us to make sense of how people can follow rules, but it does so at a fatal cost (McDowell 1998d, 203-212). On the one hand, if following a rule is attaching to externally-characterisable rails, one cannot go wrong once one has done so, and this prevents rules being genuinely normative<sup>16</sup>. If it is the case that we can see the rail stretching out before we attach to it, there

---

<sup>16</sup> According to McDowell, this was an indirect consequence of Crispin Wright's community-involving view. (McDowell 1998a, 232-236). Wright appears to have largely come round to the conclusion that a package of views like McDowell's is ultimately the only way to secure empiricism in the face of the rule-following considerations (although he retains significant worries about it as a general philosophical approach) (Wright 2007).

is no longer any sense to be made of our being right in going on in one way rather than another, since staying 'on track' will require no evaluable acts of normative (logical, semantic, epistemic, ethical...) judgement. On the other, perhaps the right result (if the 'rails' picture is mandatory) is to say that rules can't really be said to exist – there are no determinate and yet normative things for rules to be<sup>17</sup>.

McDowell rejects the picture of rules as rails, and endorses the view that the worldly patterns and personal dispositions-to-respond that rule-following involves cannot be identified without using the content of the rule itself. McDowell's response is not a denial that there is any psychological disposition that holds of a person who knows how to follow a rule. There will be such a disposition. It is just that *which* disposition it is cannot be picked out by anyone who does not already understand the rule. Understanding the rule, and understanding the behaviour that internalising the rule produces, are co-dependent. A consequence of this view, which McDowell thinks people find unsettling, is that there is nothing that a person's rule-following consists in apart from the internalisation of some extremely specific, irreducible pattern. That a person is genuinely following a rule is, then, a highly fragile and contingent circumstance; their 'going on in the same way' is not something that can be guaranteed or known about without sharing an element of their personal constitution, their 'whirl of organism'. (McDowell 1998d, 206-9). As with the elements of *Mind and World* summarised above, this relatively general part of McDowell's work will be of great importance for the account of his metaethics that follows.

### §1.3 Worldly Facts; Human Responses

---

<sup>17</sup> This is Kripke's 'sceptical solution'. See (Kripke 1982; McDowell 1998a).

These general issues aside, I can return to the main focus of this Chapter – describing McDowellian metaethics, and clarifying how it can do background work in capability theory. In the remainder of Part 1, I describe the core idea of McDowellian metaethics, as I understand it: the idea that both moral facts and the states constitutive of objectivity can be characterised as *sui generis* dispositions.

The insight that we can use the concept of a worldly disposition to flesh out a conception of moral reality is not unique to McDowell. Other proponents of the idea include David Lewis, Mark Johnston, and Michael Smith (Johnston, Lewis, and Smith 1989), and David Wiggins (1987)<sup>18</sup>. However, McDowell's view is different, in some of the specific claims it makes, and in the connections those claims have to other parts of his philosophy. For McDowell, the function of the dispositional model is twofold. Firstly, understanding moral facts as dispositional helps us to understand how they can have certain properties which have been thought of as 'queer', without giving up on realism concerning them. Secondly, by exploiting a series of analogies that the dispositional model opens up, a number of 'companions in guilt' arguments are made available. These arguments show just how much of the world as we know it is unavailable to the moral anti-realist; just how impoverished such a perspective would have to be, if it were consistent. I'll go through this account briefly now, before noting a potential problem and describing a way of avoiding it.

Many of the major threads of metaethics in the late C20<sup>th</sup> can be seen as beginning with J.L. Mackie's 1977 book, *Ethics: Inventing Right and Wrong*. There, Mackie raises several

---

<sup>18</sup> Enoch 2005 includes Peter Railton and Bernard Williams as varieties of dispositionalist, too. Some of these thinkers are closer to McDowell than others, in various ways. See for example Smith 1989, 107; Johnston 1989, 147-8; as well as much of Wiggins 1987, for some not dissimilar thoughts. None of these views has all the benefits that McDowell's does, however, because none of them is as consistent in their anti-reductionism.

potential problems for the realist conception of ethics that may seem to be implicit in ordinary practice. The most important feature of these is that they function in the way they do because of the *categoricity* of ethical demands, as well as their *motivational force*. One problem is that ethical facts would have to have motivational force A) independently of a wide range of contingent features of individuals<sup>19</sup> and B) despite being mind-independent real existences (Mackie 1977, 40-2). Alternately, if moral properties are independent of humanity, it is unclear how we can come to know about them: the usual epistemic accounts do not seem appropriate, and traditional responses refer to a special faculty of intuition that seems obscure to many (38-40). It is problems of this kind that dispositionalism responds to.

It does this by substituting a somewhat different conception of moral facts and their independence from human subjectivity for the one that Mackie has in mind. Rather than thinking of such facts as generally inert, sitting in the cold of the physical world quite apart from anything human, we can think of them as precisely the kind of thing that *is* hooked in to human subjectivity. They can be seen as dispositions to elicit responses from agents who interact with them (McDowell 1998e). If these responses include actions, this could solve the problem of motivational efficacy. Similarly, if they include the formation of beliefs, this could solve the problem of epistemic access. Things are not quite as simple as this, however. There are a number of further features of these dispositions, which are required in order to respect one or another property of ordinary ethical thought and practice. Firstly, the dispositions have to be dispositions to cause responses in *virtuous* agents; those who get things right.

---

<sup>19</sup> This is required because of categoricity. If the motivational force of a moral property depended on an individual's contingent desires, for example, any individual that did not happen to have such a desire would no longer be subject to prescriptions generated on the basis of instantiations of the property. Mackie commits to motivational internalism here. As I shall describe below (§2.1), McDowell is also a motivational internalist (of a sort), and this prevents him from responding to Mackie in the way that some others have. Compare, for example, the classic response of Brink (1989, 37-56).

Otherwise, it would be unclear why there isn't a consensus about ethical questions – if everyone who came into contact with a particular situation was caused to have exactly the same response to it, general agreement would be the result (cf. Wright 1988, 5-11). Secondly, we have to be able to make sense of the idea that moral facts *merit* (or even *require*) certain responses from people, rather than *merely causing* them. (McDowell 1998e, 143-5 (esp. fn. 42)). Thirdly, if a single moral fact is to be capable of generating actions (so as to solve the motivation problem), and beliefs (so as to solve the epistemic one), it is clear that there will have to be some sort of unity at the psychological level between beliefs and action in this case. I discuss this third feature at greater length below (§2.1). Respecting the second feature, I take it to be clear that any plausible account of the phenomenology of virtue and its proper placement in social practices will have to accommodate this sense of meriting/requiring, but I do not think I need to say more about exactly how this can be done here. Finally, the first feature is accommodated straightforwardly, by introducing the restriction to the responses of the virtuous directly into the theory, in a way that will become clearer through the rest of this section.

Once the general concept of a subjectivity-involving worldly disposition is available, it is possible to link ethical facts (so construed) to a more-or-less wide variety of other facts that are subject to the same treatment. The thought is that dispositional treatments are available for those facts that have traditionally been thought of as secondary qualities. Examples that McDowell discusses include redness (134-5), funnyness (157-161), and fearfulness (143-6). As I have said, once analogies with properties such as these have been established, companions in guilt arguments can be formulated to show A) that moral properties are not 'queer' after all – rather, they are eminently ordinary, and/or B) that a very wide range of

people (including, perhaps, many moral anti-realists) are committed to the existence of properties that will stand or fall with those implicated by ethics. Given that my aim is not to fully justify this metaethical approach, but just to demonstrate its congruence with capability-theoretic concerns, I do not need to go into any more detail here.

However, one problem which looms at this point turns out to be of more significance for my purposes. This problem concerns the primary/secondary distinction – how sharp a dichotomy it is, and what really marks the difference. One of the ways that McDowell introduces the primary/secondary division that he wants to make use of is in terms of *intelligibility independent of subjective response*: this is what primary qualities possess that secondary qualities do not (1998, 133-4e). As Tim Thornton has pointed out, however, this way of making the contrast begins to make less sense when we bear the rest of McDowell's thought – especially the theses of *Mind and World* – in mind. In general, McDowell resists the idea that there are parts of the world that could lie beyond the possible scope of our concepts (for some maximal Davidsonian 'our', at least); this is what I referred to above as his 'idealism'. As a corollary, the world is not some alien place inherently beyond our ken, it is rather the kind of place that we do, or could come to, have experiential or experience-implicating knowledge of. (Thornton 2004, 65-99). This makes it less clear that the difference between primary and secondary qualities can be cashed out in this way: perhaps *every* empirical property can be seen as a disposition to give rise to *some sort* of experience in *some sort* of epistemic agent<sup>20</sup>.

---

20 This may seem to straightforwardly vitiate a commitment to realism, as the view that facts are not – in general – the kind of things that agents are capable of recognising whenever they obtain. McDowell appears to be committed to such a thesis (e.g. 1998f, 315). However, the appearances are deceiving here. We can make the idea that there is, for each fact, some agent whose psychology it could affect in an evidence-constituting way coherent with this Dummettian conception of realism if we recognise that a person can, on an occasion, 'have a knowledge of conditions that they are not, in general, capable of recognising whenever they obtain'. Here, as elsewhere, the way to dissolve a philosophical puzzle lies in avoiding a reductive insistence that the shape of a norm must be independently codifiable; in this case, the idea is that someone can have direct knowledge of some worldly fact not because they possessed an independent epistemological decision procedure that guaranteed them that access, but simply in virtue of the irreducible way in which

On the other hand, another element of McDowell's construal of the difference is not subject to this problem. This is the emphasis on the *specificity* of the responses that secondary qualities involve. Secondary qualities like colours or moral facts differ from primary qualities in that the agential capacities they recruit are not essential to experience of the world in general<sup>21</sup> – the emotional capacities that, as I shall suggest below (§2.1), are involved in moral cognition would be a good example of this. This specificity might also be reflected in the (possible) fact that there could be other concept-using beings<sup>22</sup>, unlike us in many ways, who would find ethical normativity entirely unintelligible and would be forever cut off from ethical reality, but who would be perfectly capable of engaging with the discourse of (e.g.) physics.<sup>23</sup>

The picture that we have arrived at is fairly simple, at least at this schematic level. The core thought is that there are ethical facts, and (complete) objectivity consists in the apprehension of all such facts.<sup>24</sup> However, these facts cannot be apprehended, or even comprehended, without the possession of ethical concepts, which is a matter of inculcation with thoroughly human (as it happens, thoroughly cultural) dispositions. As in the case of rule-following generally, we can deploy the notion of a disposition in two places. Firstly, the ethical facts can be thought of as (world-side) dispositions to produce certain responses in suitably-constituted subjects. Secondly, the subjects whose responses these are will instantiate a psychological

---

they relate to the world at this particular juncture. (Cf. McDowell 1998f).

- 21 A different way of capturing this might be to recognise that secondary qualities have a distinctively *narrow cosmological role*, in the sense elucidated by Crispin Wright (1992).
- 22 An interesting question to consider is whether this idea of divisions between kinds of being might profitably make use of the neo-Aristotelian naturalist species concepts that I discuss in the next Chapter. However, I won't discuss this further.
- 23 I do not ultimately want to take a stand on whether Thornton's criticism is cogent here or not, and my argument does not rely on it. Using the specificity-based conception of secondary qualities, however, will both avoid this criticism if it *is* cogent, and may clarify other interesting connections here, like the possibility noted in fn.21 above.
- 24 As I have noted before, it is obviously extremely unlikely that any individual will possess complete objectivity, but this does not matter for my purposes; all I require is that the broad first-order framework of universalist capability theory be objectively justified for (at least) me. Because this framework is fairly vague, it only implicates some small subset of all the specific ethical matters of fact that there are; thus, complete objectivity about it will not require complete objectivity *tout court*.

disposition – precisely the disposition to respond in that particular way to that particular worldly fact. Both these dispositions – the 'fact' disposition and the 'virtue' disposition – will presumably supervene on other facts (psychological, social, biological, physical, etc.), but which facts those are is not something that could be fully specified in advance of the possession of the 'agent' disposition itself<sup>25</sup>. It is this agent-side disposition that McDowell calls 'virtue'. In this way, the concept of virtue is brought into conceptual connection with ethical objectivity, conceived in a fairly traditional realist frame.

#### §1.4 Sui-Generality and Circularity

What I have just stressed is the sense in which ethical practice, and the specific set of capacities and concepts it involves, are *sui generis* – they stand on their own, irreducible to anything else. In particular, we cannot expect to ever be able to simply say what the good is – or which actions are right – without relying on an irreducible expertise. Rather, these things can only be shown immanently from within the developed virtuous sensibility<sup>26</sup>. In this section, I turn to address a common accusation against views like this: that they are viciously circular, trivial, or vacuous, which prevents them from doing any philosophical work at all.

One classic way of bringing out this sense of vacuity<sup>27</sup> is Crispin Wright's in his 'Moral Values, Projection, and Secondary Qualities'. Wright notes that dispositional accounts of properties involve 'basic equations'. For example, consider this basic equation, for 'red':

---

25 Of course, it is also possible that complete possession of the 'agent-side' disposition will require previous (or current) awareness of the world-side fact too; so this sentence will likely be equally true, *mutatis mutandis*, for the 'fact' disposition. That is to say, perhaps one cannot know what virtue looks like (psychologically speaking) until one knows how the virtuous would respond to which facts; perhaps the advent of recognition of the features and the responses is strictly contemporaneous with recognition of the psychological state.

26 I discuss the relation of this thought to McDowell's particularism in §2.1 below.

27 There are others: D'Arms and Jacobsen argue that the concept of the moral sentiments is theoretically circular in McDowell (2000, 729-32); Smith claims that accounts such as his cannot provide interesting analyses (1993, 247).

x is red iff for any [subject] S: if S knows which object x is, and knowingly observes it in plain view in normal perceptual conditions; and is fully attentive to this observation; and is perceptually normal and is prey to no other cognitive d[y]sfunction; and is free of doubt about the satisfaction of any of these conditions – then if S forms a belief about x's colour, that belief will be that x is red.  
(Wright 1988, 15).

This statement is plausibly true (or in the vicinity of a truth) and illuminates the secondary-quality nature of redness. Wright suggests that this sort of statement constitutes a paradigm case for dispositional accounts of secondary qualities.

However, compare the following (simplified) basic equation for 'good', which is consonant with the McDowellian line of thought I'm pursuing:

x is good iff for any S: if S is virtuous and S is fully aware of x's features, S will believe that x is good.

Like the first statement, this is plausibly true. However, Wright argues that there is a major difference between this sort of claim and the one concerning 'red'. Here, unlike there, it is not likely to be the case that we can independently specify the conditions that the subject needs to be in such that their beliefs are guaranteed truth. Both statements fit the form:

x is F iff for any S: if S operates under conditions C then (if S forms a germane belief then S will believe that x is F)

For the case of redness, but not for the case of goodness, it is plausible that we can give a substantive account of the conditions C independently of particular judgements about the actual distribution of F: we can say what normal vision is like – and perhaps also what normal visual conditions are – without depending on our substantive judgements about which objects have which colours. Even given such an independent interpretation of the C-conditions for colour, it remains plausible that the biconditional will be true. Wright thinks that this is because the extensions of colour terms, but not those of ethical terms, just *are* determined by

the normal reactions of colour perceivers. He concludes that

here is where the analogy between moral and secondary qualities most fundamentally breaks down: proper pedigree for visual appraisals of colour is a matter of meeting conditions whose satisfaction in a particular case does not directly depend on what the extension of colour predicates is; proper pedigree for moral judgements, by contrast, is a matter of meeting conditions the satisfaction of some of which is, irreducibly, a moral question. (1988, 23-4)

However, this does not have to have the significance that Wright imputes – that ethical qualities are not really secondary-quality dispositions after all. What it does entail is that ethical dispositionalism cannot be a reductive enterprise, and the basic equations it endorses cannot be regarded as giving ethics a grounding in a thoroughly non-ethical reality. Instead, the truth of these basic equations will be a trivial consequence of the tightly circular interaction between a conception of virtue and an ethical worldview<sup>28</sup>.

As a result, it is clear that McDowell's approach cannot possibly provide certain kinds of substantial theoretical resources that others potentially could. For example, the truth of the basic equations of ethics could not be used to elucidate, for a genuine outsider<sup>29</sup>, what is involved in ethical perception (ethical language, etc.). While a colourblind individual might come to gain a better appreciation of colour facts by learning the truth of the basic equations for the relevant colours, it is not plausible that a putative amoralist could learn anything about ethical truth by coming to understand its relation to virtue, since virtue itself will be just as much of a mystery to them. In other ways, too, the McDowellian perspective must leave

---

28 Compare, especially, the discussion of accusations of relativism in connection to response-dependence theories (e.g. that of Horgan and Timmons 1996) in LeBar 2005 (198-204). The greater clarity of the circularity in McDowell's version of response-dependence, compared to that of LeBar, would make this sort of counterargument even more effective for a view like mine. I cannot discuss the details here, however.

29 Clearly, someone is not an outsider in this radical sense if they merely lack the precise enculturated location necessary to appreciate *some particular* ethical fact, or deploy correctly *some particular* ethical concept (which may be the same thing); rather, they would have to be genuinely outside the space of ethical concepts entirely. Clearly, there are not many genuine ethical outsiders in the actual world.

many of the elements of the metaethical story vague (and so – one might suspect – incomplete). This sort of incompleteness might give ammunition to (e.g.) reductive naturalists (perhaps of a proceduralist bent). However, this is not a deep consequence of this sort of perspective. What the perspective implies is not, strictly, that there is almost nothing substantive to be said about (e.g.) moral psychology, but that saying more substantive things will necessarily involve deploying ethical conceptions which are substantive in their own right, and thus that the claims, and the fine-grained methodology, that such putatively meta-level enquiries involve will be strictly coeval with the enquiry of traditional first-order ethics. The thought is not that metaethics *per se* must always be empty, but just that meta-level inquiry, if it is to be substantive, cannot be regarded as separate from ethical inquiry<sup>30</sup>.

What is not clear, however, is whether the inability of this account to play such independent theoretical roles is genuinely problematic. Firstly, McDowell's general quietist approach will clearly preclude some such contributions – otherwise it would be hard to make sense of what the quietist opposes. There is a strong potential for a kind of grand metaphilosophical question-begging in making accusations of triviality of this kind. McDowell's quietist stance – where it is in play – involves the claim that no non-trivial, constructive philosophical truths are available to be discovered<sup>31</sup>; of course, every claim he does make in such domains will then be insubstantial. In using such triviality as a *criticism*, an opponent presupposes the falsity of quietism; accusations of theoretical shallowness are appropriate only if something

---

30 There is a second potential barrier to substantive abstract theorising in McDowell, which this response neglects. The extent to which epistemological particularism is true will be the extent to which generalities about meta-ethical matters – which kinds of consideration can count as evidence, for example – will not be reliable. However, as I note in §2.1 below, McDowellian particularism does not entail skepticism about all ethical generalities – McDowell is not committed to a strong form of epistemological particularism – so there might be some genuinely informative theory about (e.g.) the nature of virtue after all. All McDowell is committed to is the claim that we do not strictly *need* there to be any such theory for it to be possible for us to (virtuously) go on in the same way.

31 Traditionally, quietists have taken the view that this is because the supposed questions that constitute the domain are incoherent or in some way confused. I do not need to commit to this here, however.

deeper is available (1998a, 241-2). Secondly, it is not clear that being trivial, circular, or theoretically vacuous in the respect that McDowell's view is results in his view being totally uninformative, or (especially) practically useless. For a start, surely the realisation that there is nothing substantial to be said about a putative question is a pragmatic gain – it would allow us to divert time and energy to other questions. Apart from that, perhaps the *apparent* candidates for non-trivial truth – which turn out to be confused or otherwise implausible – have had pernicious effects on other, more substantial, areas of our discourse that quietist therapy can then remedy. The impetus towards weak practical responses to injustice that I have previously argued proceduralism supports would be one example.

To end this subsection, I'd like to note a respect in which this apparent weakness – the way that linking virtue and objectivity so closely makes most metaethical truths insubstantial – may actually be a benefit. As I have stressed in the last Chapter, it is necessary to incorporate a reciprocal justificatory link between correct methodology and correct conclusions – this being what was frequently referred to as 'epistemological naturalism'. What I want to note in closing this section is the extent to which the McDowellian emphasis on the theoretical co-dependence of feature and response makes sense of this<sup>32</sup>. For McDowell, there is a clear rationale for us to alter our concrete methods of going about ethical discourse and practice when we come to see that following a previous method has not lead to success: it is only by following the right method – becoming virtuous, perhaps in a community of virtue<sup>33</sup> – that we

---

32 Cf. James 2007, 312 on this point.

33 This way of putting it brings to the fore the normative individualism built in to McDowell's approach. However, as I've already stressed, there does not seem to be any reason to think that this precludes a social epistemology in any substantial way. The social environment would need to be incorporated into the conception of individual virtue, such that no-one can be fully virtuous on their own, although the virtue in such a case presumably remains *theirs*, rather than belonging to some mysteriously free-floating communal entity. Cf. Ch.3 §3.

could *possibly* come to act well with true reliability. As David Enoch has suggested<sup>34</sup>, reductive dispositionalists or constructivists do not have this rationale, because they cannot invoke truth without invoking methodology *first*, and what is in dispute here is substantive truth, with methodological consequences following depending on which particular judgement is reached. Non-dispositional ('robust') realists like Enoch himself can make sense of altering methods or other 'agent-side' elements in the name of truth, but it is not clear for them why good practice – virtue – is so deeply important for getting ethics right: for them it is a mere means to an end of which it is constitutively independent. Likewise, robust realism has problems explaining the epistemic and motivational connection between facts and (epistemic and practical) agency, given this complete metaphysical independence. It is only a view that posits a constitutive *but mutual* relationship between ethical virtue and ethical reality that can be entirely satisfying here; and this is what McDowell's approach – perhaps uniquely – does.

### §1.5 Constructivism Redux?

Hopefully, the basic form of the secondary-quality analogy, and the dispositionalism that it makes use of, are now clear. In this section I will discuss one potential challenge to my use of McDowell, which involves a claim that McDowell's theory is a form of constructivism after all – given that I have previously argued that capability theory is not well suited to a constructivist development, this would be fatal for my project if true. I shall argue, on the contrary, that the McDowellian perspective dissolves the Euthyphro contrast that drives the idea that we must choose between constructivism and a dubiously non-anthropocentric ethical realism. In doing so, it makes sense of the epistemological feedback loop between conception of expertise and conception of fact in a straightforward and satisfying way. Two applications

---

<sup>34</sup> See the next subsection, where I bring this critique out more fully.

of the Euthyphro contrast will be discussed here. Firstly, I shall note some close similarities between the view I'm endorsing and the avowedly 'Aristotelian Constructivist' approach of Mark LeBar, the critical issue about which appears to turn on a question about some kind of *priority*. Secondly, I shall bring my view into the frame set by David Enoch in his devastating criticism of the idealisations which any remotely plausible constructivism must involve; doing so makes the power of a McDowell-style realism particularly clear.

LeBar's metaethical stance has, on the face of it, remarkable similarities with McDowell's, as I've been describing it here. The core of his view is given in this passage:

The *goodness* of (say) a cold beer is not anything one recognizes and responds to. Instead, one recognizes and responds to its natural properties (these, of course, are not constructed), and to the extent that, in virtue of its natural properties, it contributes to one's living well, only in that way does it have value... the goodness of a good life is in turn crucially dependent on the exercise of practical rationality in the right way (that is, practical wisdom, or *phronesis*). This is true both because such lives require the direction of practical wisdom and because only successful practical rationality can determine which lives are genuinely good. On the other hand, the criterion for success in practical rationality — practical wisdom — just is the construction of a good life. Neither eudaimonia nor practical wisdom can be specified without essential reference to the other. (2008, 13-14)

This shares two important features with the view I'm developing here. Firstly, various goods are proposed to depend upon the integration of certain things into a life of virtue ('practical wisdom').<sup>35</sup> Secondly, and more importantly, LeBar also holds that the conception of the good and the conception of virtue are circularly interdependent.<sup>36</sup> However, LeBar seems to think

---

35 The version of this claim I advocate will become clearer in the next Chapter (esp. §3.2).

36 One potential difference between LeBar's view and the one I've been developing is that LeBar holds that the connections he identifies are connections between properties rather than concepts; *de re* identities between the properties picked out as ethically salient by virtue and those that actually have such salience, rather than connections *de dicto/de se*. He has argued elsewhere (2005), that the tendency to rely on an *a priori* and necessary connection has damaged the case for response-dependence theories. See also Wedgwood 1997 and Mišćević 1998 for similar claims. I accept (tentatively) that it is better to see the connections as operating at the level of properties rather than concepts. However, as LeBar points out, recognition of the identity between the properties could and should transform the concepts over time (2005, 181-3). Because McDowell's version of the response-dependence claim does not entail any positive first-order theses, as I acknowledged in the previous subsection, it would probably be fairly easy for people to reform their thinking to align with it, as I would recommend.

that accepting these claims amounts to being a constructivist. Some of the things LeBar has said suggest that the difference here is merely terminological:

The conception of welfare at work here is constructivist in that there is no “prior and independent order of objects and relations” providing a criterion for the principles governing what constitutes well-being, no “standpoint external to the parties’ own perspective” from which principles of welfare may be specified. (2004, 204)

If all that is required to be a constructivist is to think that there is no *prior* and *external* standpoint from which ethical truths can be understood, then McDowell is clearly a constructivist. But there are two other senses of constructivism which should be considered here. Firstly, constructivism can be construed as the positive thesis that 'the parties' own perspective' is prior to the truths they endorse, rather than the other way around. Secondly, constructivism can be viewed as the metaethical counterpart to first-order proceduralism – the view that there is some formulable procedure that is sufficient for ethical epistemology, including (e.g.) the epistemology of claims about human dignity. Viewing the constraints placed by practical wisdom as irreducible to any independent procedure or independently-identifiable perspective, as LeBar explicitly does, would clearly rule out both *conceptual* priority – or 'priority in the order of understanding' (McDowell 1998o, 157), and the *epistemological* priority – the privileging of a methodology even in the face of its apparent extensional failures – which typifies proceduralism. It appears, then, that if there is any major meta-level difference between the strongly anti-priority view of McDowell and LeBar's 'Aristotelian Constructivism', it will have to lie in a thesis about some *other* form of priority, which LeBar does not mention<sup>37</sup>. Otherwise, LeBar's claim that his view is constructivist appears to rest entirely on the fact that it fails to endorse the traditional 'realist' answer to the Euthyphro question. It is unclear why constructivism should be identified with the mere rejection of the 'robust' realist horn of the dilemma, rather than with the positive espousal of

---

37 Perhaps LeBar has some distinct form of *metaphysical* priority in mind, of the sort which has recently become fashionable? (See, for example, Schaffer 2009). It is unclear.

the anti-realist one. In the rest of this section, I shall suggest that the McDowellian view, as I understand it, is distinct from both constructivism and traditional forms of realism precisely because it dissolves the Euthyphro question itself.

On my understanding, what is truly distinctive about McDowell's view is the way that his general anti-reductionist (§§1.2-4) and realist/idealist perspective (§1.1) allows him to justify a principled refusal to assign priority to either agent-side *or* world-side features in metaethics. I need to say more, though, about how this refusal works. It will be helpful, in doing this, to present anti-priority realism against the background of David Enoch's attack on idealisation in the context of subjectivity-based theories. To begin with, Enoch has pointed out that response-dependent or dispositionalist theories are very often closely related to constructivist views; indeed, he has suggested that most contemporary constructivists *are* dispositionalists, of some broad kind (2009, 21, fn.14). This makes some sense, because all metaethical approaches have to have a world-involving component at some point – they all have to make claims about what ethical properties are, however those are construed – and dispositions look apt to do the job that constructivisms would implicate. Consider: according to constructivism as I have described it, the core of the view is that subjective responses, constrained solely by some procedure (as opposed to directly by substantial ideas about content), give rise to the moral facts<sup>38</sup>. A disposition-based account of moral facts is then immediately available: they are dispositions to cause responses in agents whenever those agents' psychology passes some procedural test. Enoch provides an argument against views like this that trades on the procedural nature of the ideal agent that they involve. He questions whether they can provide

---

38 Street 2010 (364-6) objects to this sort of characterisation that it is better to view constructivisms as holding that normative truths arise from a particular sort of *perspective*, rather than from the application of a procedure. However, I don't think this makes any difference to my argument here or throughout the Thesis: constructivists have to provide a plausible account of the boundaries of this perspective, and this seems to play precisely the same role as a procedure in whittling down the candidates for truth.

any rationale for idealising – for restricting the set of subjective responses that are valid – consistently with the motivations with which they begin.

It is important to note immediately that Enoch does not, actually, view his argument as applicable to McDowell (or to David Wiggins' similar thesis). This is because he recognises that there is, on McDowell's account, no conceptual priority, and not necessarily any metaphysical dependence, between the level of moral concepts/properties, and the non-moral levels on which that level supervenes<sup>39</sup>. He also notes that 'no-priority' views are invulnerable to the concerns he raises in a later paper, and this time he specifically acknowledges that an accusation of 'constructivism' would be illegitimate because constructivists take a stand on the Euthyphro contrast: they hold that normative truth depends on correct human response, and not the other way around (2009, 21, fn.16). As I put it above, for constructivists the responses 'give rise to' the moral facts; this is not a claim that McDowellian metaethics needs to make. On the other hand, it will be helpful in explicating McDowell's position to go into more detail on *why*, exactly, it is not vulnerable to the sorts of concern that Enoch raises.

Enoch's argument, in summary, is that metaethical approaches which tie truth to the subjective responses of agents are forced to engage in some idealisation in order to prevent their theories giving wildly implausible results. Thus, rather than my responses *now* determining the ethical truth (or some part thereof), it would be the response of an ideal version of me – someone with full information about the non-normative facts, for instance; or someone who is not motivated by partiality (procedurally construed)<sup>40</sup>. The problem for these idealising projects is that there is no clear rationale within their starting-points for such

---

<sup>39</sup> He claims that the argument could be modified so as to apply to 'no-priority' views too, but gives no indication of how this could be done. (Enoch 2005, 765, fn.18)

<sup>40</sup> Compare my commentary on Baber's response to the problem of adaptive preferences in Chapter 3 (§1.2).

idealisation: it rapidly begins to look *ad hoc*. To put it slightly polemically, it looks as though the proponents of these views realise that the place they begin is utterly inhospitable to any decent ethical perspective, and desperately aim to cover up this fact by splashing some stolen paint on the walls. The most immediate problem is that people who hold that there is no truth *apart* from subjective response cannot claim that some responses are better than others in virtue of tracking this independent truth (Enoch 2005, 761-5). In addition to that, which Enoch calls the 'natural' answer, there are a number of other possibilities for defending idealisation here. Enoch goes through these in turn, and shows that they cannot be made to work for the thinkers he targets; I don't need to follow the rest of his discussion here.

It is worth reflecting again on why, exactly, this challenge cannot work against McDowell. One of Enoch's claims provides a good clue. He says that “[a]s even the geography of their texts shows, idealizers start off with actual responses, then patching [*sic*] up extensional inadequacies by idealizing.” (2005, 769). But, for McDowell's case at least, this would not be a charitable claim. As it happens, McDowell never actually does the sort of work in which this tendency might be observable – he would likely regard it as too first-order: a matter for constructive engagement rather than quietist therapy. But there is a good reason for that, quite independent of his metaphilosophical stance: from his perspective, such projects make very little sense. Ethics is *sui generis*, and the only way to say anything about it is to *do* it; nothing can be said from outside. It seems more fair, then, to see McDowell as holding that the primary, orienting concept of ethical/political epistemology should be that of *the person who gets things right* (otherwise known as the virtuous person)<sup>41</sup>. This concept, in contrast to those

---

41 Cf. Zangwill 2003, which argues against many dispositionalist views on the grounds that they radically distort moral phenomenology and practice. He notes that this argument doesn't apply to normatively-laden accounts of the relevant dispositions – such as McDowell's – but argues that such accounts are vacuous. Hopefully I say enough in this Chapter to draw the sting of this criticism. Kwall 2004 provides an extensive response to Zangwill here, although it too does not apply to norm-laden accounts.

of the half-hearted idealisers that are Enoch's main targets, is a concept of a correct-agent, rather than of an agent who is arbitrarily stipulated to be correct, or a metaphysically independent correctness that just happens to be realised by an agent<sup>42</sup>. The concept of correctness plays an essential role in the individuation of the psychological form in question. As a result, the concept of moral correctness should be understood relationally, with a constitutive connection to agential capacities; likewise, the concept of the virtuous agent should be understood as constitutively linked to an idea of getting things right. The idealising projects that Enoch criticises begin, as he says, with an idea of subjective responses and underlying psychological capacities that does not *necessarily* implicate any concept of correctness at all. For McDowell, however, understanding the relevant psychological capacities as the basis of ethics, and understanding one norm-involving shape that human psychology can take – the shape of the virtuous agent – go together from the very start.<sup>43</sup> The link to McDowell's take on the rule-following considerations should be clear: here, as there, the way out of the fly-bottle is shown by refusing a theoretical demand to take up a perspective independent of that which is to be explained. Behaviour governed by ethical rules, as with linguistic ones, cannot be accurately thought about from outside the form of life in which those rules figure.

---

42 This is similar to the 'justificatory practices' rationale that Enoch rejects when offered on behalf of his targets. McDowell's account of our implicit justificatory practices is, at the very least, much closer to the sort of account that Enoch argues for, however, which suggests that McDowell could use this rationale for idealisation where others could not. Cf. Enoch 2005, 777-8, especially fn.38.

43 A second line of response to arguments like Enoch's would note that the connection between beginning on a process of discursive thought about an area, and beginning to have in view a normative end-point that can be identified with the possession of some more-or-less specific set of agential capacities, may be fully *general*, in line with the sort of considerations that I mentioned in connection with Thornton above. Thus, it is not merely ethical engagement that presupposes a constitutive link between making contact with the world, and approximating some sort of epistemic agent, it is all worldly engagement *per se*. If Enoch's arguments were to tell against *that*, it looks like they would prove far too much: they would amount to an attack on McDowell-style idealism/realism syntheses themselves. In line with my earlier reticence about Thornton's suggestion, however, I don't want to commit to this sort of reply.

## §2 Virtue as a Shapeless Unity

I concluded the last section with a recognition of the thoroughgoing anti-reductive approach that McDowell takes to the individuation of the perspective with which objectivity and right relationship to ethical reality are associated: the perspective of virtue. Before I move on to consider other central elements of McDowellian metaethics in some detail, I need to place what I've said in the context of a helpful piece of terminology. One way of expressing the thoroughly *sui generis* nature of ethical terms, concepts, properties, and anything which involves them, is to say that such elements of the world are *shapeless*<sup>44</sup>, when viewed from anywhere outside of ethics itself.

Since I'll occasionally deploy this terminology throughout the rest of the Thesis, it is worth dwelling on it briefly, to clarify what I do – and do not – mean by it. This is especially important because most of the work that has been done on this concept, and its proper role in metaethics (if it is thought to have one), has taken it to have a different, and often more radical<sup>45</sup>, significance. For example, Pekka Väyrynen has recently argued (2014) that shapelessness does not entail anything particularly special about ethical terms, concepts, or properties *per se*, apart from the untenability of non-revisionary forms of reductionism<sup>46</sup>, because the kinds of reductive theoretical move which proponents of shapelessness have argued are impossible seem to be impossible in very many non-ethical cases too. Väyrynen

44 The term originates from Blackburn 1981, 167.

45 Roberts 2011, for example, argues that the shapelessness thesis about evaluative concepts, in conjunction with the view that thick concepts are irreducible to thin ones (which I am not taking a stand on in this thesis), entails a very radical form of particularism; one far more radical than any of McDowell's own work would suggest. Cf. the very weak version of particularism described in the next subsection.

46 Strictly speaking, the shapelessness of ordinary/pre-theoretical ethical concepts is compatible with revisionary reductionisms like that of Railton (e.g. 1986), as Väyrynen notes (18, fn.32). This is because revisionists can, in principle, simply revise the concept until it *becomes* identical with some pre-existing independent concept. Clearly, doing this would violate the motivations behind this entire Thesis: if we allow metaethicists to revise ethical concepts without regard to extensional adequacy, in principle *any* first-order ethical view (i.e. about justice) could end up being true!

thus argues that the failure of such reductionist programmes does not support many significant positive metaethical proposals (11-19). I can accept this argument without difficulty, because as a quietist I am unconcerned with justifying the more constructive or revisionary theories that Väyrynen notes are unsupported by the fact of shapelessness (cf. 19-21). To give another example, Gerald Lang has argued<sup>47</sup> that considerations about shapelessness do not effectively support objectivism about ethics over non-cognitivist or error-theoretic rivals, because these views need not suppose that shapelessness is false (2001, 202-4).<sup>48</sup> This is similarly irrelevant to my concerns here, because I am not concerned to argue against proposals which do not deploy a notion of ethical truth (cf. fn.1 above). For my purposes in this Chapter, and in the wider Thesis, I do not need to endorse a version of the shapelessness hypothesis which extends beyond the generalised anti-reductionism on which I've focussed to this point.

## §2.1 Particularism, Cognition, and Affect

As McDowell says, rather than it being the case that we should grasp virtue in terms of a conceptually-external relation to a conception of right action – understanding the shape of virtue 'from the outside in' – we can only really understand what it is to be virtuous by *being* (however partially) virtuous, and thereby grasping it 'from the inside out' (1998h, 50)<sup>49</sup>. In this

---

47 Alex Miller has endorsed the view that McDowell's argument against non-cognitivism is mistargeted on similar grounds (2003, 244-56).

48 The clarification of shapelessness as a moderate thesis with exclusively negative (anti-reductive) entailments should be viewed as of a piece with the tentative concession, made in fn.36 above, that the postulated links between the perspective of virtue and the perspective of objectivity can be viewed as *de re*, concerning properties rather than concepts. As before, this is perfectly sufficient for McDowellian metaethics to play the supporting role in the capability approach that I have urged is needed.

49 The requirement that the state of virtue be a fully unified and comprehensive one – which is necessary in order for virtue to play the multiple theoretical roles that it does in a view like McDowell's – potentially has significant consequences for ethical practice. As Daniel Jacobson (2005) has noted, views which place such strong idealising constraints on what can count as an instance of virtue make it unlike most ordinary skills in most ways, and ensure that individuals will only extremely rarely possess it. In fact, Jacobson is so amazed by the difficulty of achieving the sort of skill of virtue that the McDowellian way of thinking involves that he expresses doubt that virtue – so conceived – is so much as possible for human beings. It remains unclear

section, I shall explore some of the aspects of this orienting idea, dealing at once with the conjuncture of affective and cognitive elements that virtue involves, and the conceptual particularism which McDowell defends in relation to it. Thus, the subsection engages in two tasks, both of which are linked to McDowell's general anti-reductionist position. Firstly, I note that McDowell's view involves conceptualising virtue as an irreducible combination of cognitive and affective psychological elements. This has the major advantage of incorporating motivation-relevant elements, reason-sensitive elements, and states (such as emotional reactions) which intuitively have a role in good ethical thought and action, all within one unified state<sup>50</sup> which is individuated with reference to a norm of correctness. Secondly, I bring out McDowell's particularism, as the claim that exceptionless principles (or even hedged principles<sup>51</sup>) are not required for ethical thought or practice to be governed by reason.

McDowell's view of virtue is sentimentalist, in the sense that he thinks of virtue as involving a shaping of affective capacities, including those that, like the psychological states that are conventionally labelled 'desires', involve tendencies toward action. However, he also thinks of the virtuous agent as being disposed towards the formation of appropriate beliefs, and as – at any given time – possessing many such beliefs. This may seem to be a problematic

---

to me exactly why he thinks this; clearly full virtue is going to be an extraordinarily rare achievement, perhaps it is even the case that it has never once been realised in the history of our species. But so long as it is not impossible, it still makes sense to place it as a regulative ideal whose interconnections provide a framework within which less idealised ethical thinking can take place. Cf. Ch.6, §3. Cf. also Stichter 2007.

50 One of McDowell's most distinctive contributions to Aristotelian moral psychology is his 'silencing' thesis: the view that the virtuous person does not face a dilemma between conflicting reasons, but (at the moment of decision/judgement) experiences all reasons but one to be 'silenced'. This thesis is problematic, and I do not rely on it here. See (1998h, 67-72); (1998i, 90-94); and (1998j, 46-9) for McDowell's development of the concept; Christensen 2009 and Stark 2001 for critique.

51 Cf. Lance and Little 2006, 2007a and 2007b; Väyrynen 2007. There is now an enormous and enormously complicated literature on particularism, but, if I am right, almost all of it is irrelevant to my Thesis, because the essential part is the denial of 'subsumptivism' alone, and that is weak enough to be compatible with almost all more general versions of both particularism and its negation ('generalism').

combination of claims, because it violates a central tenet of the 'Humean' theory of moral psychology, which is that desires (/desiderative states more generally) and beliefs (/cognitive states more generally) are metaphysically 'distinct existences' and can always be pulled apart from one another<sup>52</sup> (Smith 1994, 117-119). McDowell does not need to claim, against this 'Humean' theory, that there are no person-level states like 'belief' and 'desire' that hermetically isolate cognitive/representational and affective/motivational elements of mind exclusively. He does not even need to claim that such states play no role in the psychology of ethics. All that is required is that the psychology of *virtue* does not (distinctively) involve the possession of such isolated elements, but rather a unified person-level disposition<sup>53</sup>. This disposition happens to have functional similarities with both general 'Humean' types, but is irreducible to any sum of tokens of those types. Thus, although it is true that the virtuous have ethical beliefs (and in many cases it may also make sense to attribute distinctive desires to them too), there is no reason to expect such beliefs to conform to a psychological model that was generated independently of ethics<sup>54</sup>.

Take a particular judgement, the judgement that S is suffering unjustly, and that this suffering can readily be alleviated by resources at hand. On McDowell's view, the virtuous agent who comes to form such a judgement will thereby be motivated to alleviate the suffering, without our needing to postulate an independently intelligible<sup>55</sup> desire to do so on their part (1998i, 84-5; cf. Also 1998h, 53-4). For the virtuous, perception that a person is suffering in such

---

52 This, in turn, is central to the 'Humean theory of motivation'; As David McNaughton puts it: 'The combination of belief and desire is required to motivate [an] agent to act. Desires without beliefs are blind; beliefs without desires are inert' (1988, 21).

53 Smith concedes (119) that it might be (rationally) a requirement for virtue that an agent have both cognitive and desiderative states, so whether or not he can agree with this sentence will depend on what is involved in the disposition's being 'unified'.

54 This is McDowell's 'anti-psychologism about psychology'; cf. fn.80 below.

55 That the desire must be *independently* intelligible is an entailment of the Humean theory: how can X and Y be postulated as distinct existences if Y is not even intelligible apart from the understanding of X?

circumstances is sufficient to explain (and, not coincidentally, justify) their motivation to alleviate it<sup>56</sup>. We do not need to (and, plausibly, cannot) identify two independent states that the virtuous agent happens to have, but which can in principle be possessed separately by non-virtuous agents. It is this feature which violates the Humean theory. For the Humean, in this case there must be A) a cognitive state involving the recognition of the suffering (or, for example, recognising that it would be right to relieve suffering in this case), and B) a separate desiderative state involving the motivation to alleviate suffering (or, perhaps, to do what is right).

Smith's 'Humean' attack on McDowell's view ultimately<sup>57</sup> involves the claim that it is implausible that any unified motivationally-relevant belief-state could exist, because it requires that there be beliefs such that (e.g.) an individual can have them at one time, when their desiderative constitution is virtuous, but must be thought to 'lose' them whenever they fail to be adequately disposed to act well (1994, 122-4). On the face of it, this is a good objection; it does seem outlandish to think that people's abilities to entertain propositional contents (and so form beliefs about them) can depend in this strict way on whether or not they have certain motivations. However, this interpretation reflects an uncharitable understanding of the McDowellian view. Margaret Little provides a nice account of how this is so (1997, 72-

---

56 This anti-Humean aspect of McDowell's ethical psychology coheres very well with the particularism described below. One reason to be sceptical of the Humean postulation of independent desires is that it is unclear that there are any independently describable desires of the non-virtuous that can exactly match the patterns of motivation that the virtuous have (1998i, 84-5). The affective/motivational responses of the virtuous may be uncodifiable, just as the content of ethics itself is.

57 There is a preliminary attack involving the claim that there is a contradiction in the very idea that a unitary state could be both belief- and desire-like. Smith claims that cognitive and desiderative states have different 'directions of fit'; different dispositional profiles. On Smith's favoured dispositional theory of desires, a desire is simply a disposition to be motivated towards certain actions in certain circumstances. This would not work if the existence of the desire depended on not recognising that its content (what it is a desire *for*) obtains – if so, no-one could desire things that are not the case. However, part of the dispositional profile of a belief *is* a tendency to vanish when presented with evidence that its content does not obtain. (1994, 113-6). Smith wants to argue that this involves a contradiction. However, this argument is invalid because the anti-Humean does not have to believe that the functional dispositions are all directed at the same contents. In the end, Smith appears to concede this (118). Cf. Little 1997, 70-71.

6). It is not charitable to view the anti-Humean as holding that the understanding of the virtuous is reducible to belief in coarse-grained propositions<sup>58</sup> like 'x is suffering unjustly, and I can help', but then supposing that whether or not such a proposition is cognitively available to an agent depends upon their concurrent motivations. Rather, they simply deny that the cognitive availability (or not) of such propositions is sufficient to capture virtuous cognition<sup>59</sup>. Little urges a comparison between virtuous perception and the phenomenon of 'gestalt shifts', here. If a person who has previously been able to see the image of Marilyn Monroe in a 'Magic Eye' painting later fails to be able to do so, it would not make sense to say that they ceased to judge (or believe) that the image was there, but it would make equally little sense to insist that no cognitive change had occurred<sup>60</sup> (73). Unless the Humean can provide a unified account of cognition which refers solely to coarse-grained propositional content, arguments like Smith's are not cogent.

The most important advantage of the anti-Humean view is that it is extremely flexible in principle – any element of human psychology could, if it became plausible, be implicated as part of virtue<sup>61</sup>. In particular, this allows affective change (or difference) to have an intrinsic role in epistemic improvement (or superiority)<sup>62</sup>. As McDowell puts it: 'one might ask why a

---

58 Little's formulation may be too strong at times. She holds that McDowell-type views think of the cognition of the virtuous as (distinctively) not susceptible to capture in propositional terms *at all*. This seems unwarranted – all that is required is that it not be susceptible to complete capture using only those coarse-grained contents that are, intuitively, cognitively accessible by most agents most of the time.

59 Compare, here, the disjunctivist claim that the content of experience differs between veridical and non-veridical cases. See §1.1 above.

60 Cf. Wittgenstein 2001, Part II, §xi.

61 There are two other important advantages. Firstly, as (e.g.) Schroeder 2005 notes, 'Humean' theories of motivation may provide support for Humean theories of reasons, which are species of proceduralism, with all its attendant tendencies towards extensional inadequacy (i.e. revisionism). McDowell-style anti-psychologism avoids such threats. A second threat that psychologism generates is a pressure towards non-cognitivism. The 'Moral Problem' of Smith (1994) can be resolved by denying cognitivism (12), and an assertion of the Humean account is part of its constituent inconsistent triad.

62 Cf. McDowell 2009b, 50-1, which is especially clear about the ineradicable role of affective development in the formation of virtue. For McDowell's Aristotle, 'the content of the intellectual state [which is united by cognizance of the good] is formed by moulding the *orektikon* [the desiderative part of the soul]' (51). Cf. also Nussbaum 1986, 307-17, esp.307-9; Nussbaum 1990b, 75-82.

training of the feelings (as long as the notion of feeling is comprehensive enough) cannot *be* the cultivation of an ability...to spot...the fitnesses of things' (1998e, 147)<sup>63</sup>. After I've described McDowell's second major contribution to moral psychology, I'll return to place these benefits in the wider context of feminist moral epistemology.

The second area of moral psychology where McDowell has made a relevant contribution concerns ethical particularism. The core of McDowell's version of particularism is the denial of a thesis that has since (e.g. Audi 2008) been called 'subsumptivism'<sup>64</sup>. This is the thesis that, in order for an ethical judgement to be genuinely governed by reason, it must involve the subsumption of a particular case under a general ethical law. For example, the judgement that a policy is unjust because it violates dignity (by, say, precluding the capability for reproductive health<sup>65</sup>) will only be fully governed by reason if 'dignity' and 'the capability for reproductive health', and the bundle of attributions of ethical valence denoted by 'unjust', can all be spelled out in a way that avoids reference to the particularities of the situation that the judgement is about. The primary interest of such a thesis, for my purposes at least, is that if it is true there must necessarily be something wrong, epistemically speaking<sup>66</sup>, with any

---

63 As Jeremy Randell-Koons notes in his 2003, his arguments against response-dependence theories – which argue that they lead to a problematic relativism, and which take aim especially at those with an affective component – cannot be straightforwardly aimed at McDowell, because his view is non-reductive. He does suggest that because there is no readily-identifiable moral emotion (or set thereof), it is difficult to see how to arrive at an account of the affect-laden response implicated by the 'basic equation' for ethical goodness (see §1.4 above). However, this is not particularly convincing, because we can moralise the relevant emotions – rather than speaking of anger, we can speak of 'justified anger', and so on. This will make the relevant components of the theory circular (if it wasn't that already), but, as elsewhere, it isn't clear why quietists should be very concerned about that. Compare D'Arms and Jacobson 2000, 729-32 on this point. Cf. also fn.27 above.

64 On my proposal, the denial of subsumptivism is compatible with there being, metaphysically speaking, many worldly features that always have the same ethical import no matter the context in which they occur. More radical particularisms, such as that of Jonathan Dancy (2004, *inter alia*) have denied that there are any such features.

65 Again, see Appendix A.

66 Possibly, a subsumptivist could hold that non-subsuming judgements are acceptable because they have heuristic value in everyday cases, even though they are necessarily imperfect. (Likewise the anti-subsumptivist could agree that, even where a judgement involving subsumption is actually incorrect (because there happens to be no such law), making judgements with reference to laws might have heuristic

judgement that does not have this structure. This would be unfortunate, because the lack of such a structure is, arguably at least, typical of most of the judgements that feminists and (especially) Aristotle-inspired philosophers have identified as being essential to good ethical practice. In particular, it may threaten the epistemological stance I've defended, on which a person may be (relevantly) virtuous and (relevantly) possess objectivity in making a judgement even when there is no more general (more abstract, more 'independent') law that they are relying on in doing so. I have advertised this sort of flexibility as an advantage! So, the claim that McDowell defends is this: the concepts deployed in the judgements of the virtuous need not be susceptible to capture in statements of universal law<sup>67</sup>. Virtue can be a kind of rationality even if subsumptivism is not true of it.

At the most basic level, these sort of claims should be read simply as reiterations of the message that McDowell identifies as flowing from the rule-following considerations: we cannot understand either the worldly extension generated by a rule, or the personal disposition associated with a grasp of its intension, without understanding the rule itself from the inside. So, if there is no procedure for correct ethical judgement which could be given and understood independently of virtue (if the relevant rules are not rails), it follows that there is no *need* for ethical generalities; agents *already* have to be engaging creatively with the world for ethics to exist (as a *sui generis* domain) at all (1998h, 57-65). The result is that the truth about ethics need not be *codifiable*; it need not be susceptible to being written down, such that the rules expressed in such writing are understandable (and, so, applicable) by anyone other than the virtuous themselves. In turn, this means that subsumptivism must be false, as a

---

value.) Although subsumptivism, as I understand it, is a claim about ethical concepts and their role in judgement (as is, therefore, its denial) it has epistemic consequences; if a judgement is not conceptually well-formed, it can hardly be the bearer of truth or justification.

<sup>67</sup> Cf. Nussbaum 1990b, 66-75

universal thesis about the nature of ethical judgement: there may be some knowledge-involving, reason-governed judgements which do not involve the subsumption of a particular case under a law that outstrips it in generality. (66-9). This message is technically compatible with the truth about ethics being extraordinarily simple, along the lines of total act utilitarianism, for example<sup>68</sup>. On one hand, if virtue really were that simple or easy to instantiate, emphasising its *sui generis* would have relatively little point; it would be very hard to make a practical mistake on the basis of a reductionist proposal in such a world, because the truth could be communicated very easily. However, it becomes far more important to emphasise the truth in particularism when – as seems to be the case – the truth of ethics is enormously complicated, and the epistemic (and, connectedly, practical) value of simple, abstract ethical theorising is relatively low.

The largest benefit of both of these elements – particularism and the anti-Humean proposal, apart from their flexibility, is that they provide nice links to central feminist work on moral psychology. I'll highlight two, as described by Jaggar (1989), and Little (1995) respectively. Firstly, Jaggar (amongst other feminists) has identified numerous ways in which the devaluation of emotional responses have served to disadvantage women as contributors to ethical thought and debate, given the conventional association of femininity with the dominance of emotion over (supposedly dispassionate) reason (Jaggar 1989, 163-5). In turn, the illusion of dispassionate reason occludes the possibilities for social change by disparaging the 'outlaw emotions' that might otherwise draw attention to serious injustices such as sexual harassment; encouraging the identification and articulation of outlaw emotions will then be an effective way to increase knowledge of such problems (165-9). Secondly, as well as

---

68 Cf. McKeever and Ridge 2005, 88-89, on the implications of 'anti-transcendental particularism'.

adding detail to many of Jaggar's emphases (1995, 117-121), Little supplies reason to think that affect-laden awareness, especially that paradigmatic of care, is especially reliable as a path to ethical knowledge. Caring about a person as an individual focuses critical attention on them, noting complex similarities and differences between their situation and those that the carer has previously encountered, and being more willing to relax general assumptions that have been operative. Little argues that an impersonal, dispassionate investigation of their situation is unlikely to be as epistemically reliable (122-5)<sup>69</sup>. As well as providing support for the McDowellian approach I've outlined, these links help to illuminate further the deep congruence between that approach and the sort of epistemological orientation that even strong critics of Nussbaum's justification for her capability theory favour.

I conclude that McDowell's version of particularism is *compatible* with even very simple and general principles being true, so long as they are judged so from the perspective of virtue, and not pretended to have a source independent of it. If there is a sound ethical argument for the existence of a generality, that can be accommodated. Such arguments communicate existing ethical understanding, which – although it cannot be *reduced* to any independent law – may be expressible using nomological language. Here, as well as with regard to the question of the cognitive/desiderative make-up of ethical psychology, McDowell's view turns out to be precisely as flexible as we require.

## §2.2 Reasons Internalism as a form of Proceduralism

In this section I turn to a possible argument against the McDowellian approach. The universalism to which I am committed involves the claim that a wide range of agents (all

---

<sup>69</sup> Cf. Khader 2011, which reapplies these ideas to the topic of a just, accurate approach to the identification of needs in global development practice.

humans, in fact) have certain practical reasons (the reasons of global justice), without giving any argument to the effect that this universal scope follows from anything about purely-descriptive human psychology. The tradition of 'reasons internalism', represented most importantly by Bernard Williams, opposes this sort of proposal. This is of especial relevance to my argument because the kinds of consideration that are given in support of reasons internalism also appear to favour a proceduralist approach to ethical epistemology, of precisely the kind I've rejected since Chapter 2. However, rather than explicitly discussing the possible ways in which internalist intuitions might motivate proceduralism, my strategy here will be to remove the threat by giving an argument for a reasons externalist stance. In this section, then, I explicate a McDowell-inspired response to reasons internalism, and bring out some respects in which it coheres with claims I make elsewhere, especially the claim of feminist social epistemology that differently-situated individuals may need to rely on transformative engagement with others to inform them about the right thing for them to do.

Reasons internalism<sup>70</sup> (following Setiya 2012) is the view that

The fact that  $p$  is a reason for  $A$  to  $\Phi$  only if  $A$  is capable of being moved to  $\Phi$  by the belief that  $p$ . (4)

There are potentially many different reasons for a McDowell-type metaethicist to be suspicious of internalism. But the one I'll concentrate on and aim to preserve in the rest of this subsection can be called the 'eudaimonic motivation'. The basic thought here is that there is a strong connection between the good for someone and what they have reason to do<sup>71</sup>. Some

---

70 It is important to note that this model of internalism, which is inspired by Bernard Williams' founding paper 'Internal and External Reasons' (2012), differs significantly from certain others. In particular, as Setiya notes (2012, 14-15), it differs from Michael Smith's comparably influential interpretation (in his 1995 and elsewhere). Cf. also McDowell 2006. On a view like Smith's, the theses about reasons I eventually endorse in this section might be internalistic after all. I can't explore this in detail here.

71 Arguably, this corresponds to thinking of reasons in the way Setiya identifies in this passage: "Finally, normative or 'good' reasons might be thought of as grounds on which it would be good to act: good things to have among one's reasons for acting; reasons that conform to relevant norms." (11).

possible reductive views, such as the vaguely Humean assimilation of goodness-for to what satisfies desires, can preserve this connection. A development from a constructivist reading of Kant, according to which there are universal and purely formal constraints on the human will that determine both what reasons there are and what the good is, might also do so<sup>72</sup>. But, if I've been correct in concluding that no such procedural/reductive proposal can be made to work, these will not be satisfying ways of maintaining the thought. The natural alternative, given the Aristotelian resonance of most of my claims so far, would be a view which associates what individuals have reason to do with the deliverances of the perspective of virtue – which connects to the good in being essential both to its recognition and to its realisation in a life. The problem, simply, for reasons internalism, is then that people frequently do not seem to be in a position to be motivated by the facts about their good; virtue is relatively remote from them, and so, if internalism were true, the attractions of the eudaimonic motivation would be lost.

It is worth noting explicitly that my argument at this point, in defending a sort of reasons externalism, is not dissimilar to what I've said against other varieties of proceduralism earlier in the Thesis<sup>73</sup>. The thought is that just as there are structural similarities between proceduralism about well-being (i.e. preferentism), and proceduralism about ethical-epistemic justification (i.e. the constructivism that Jaggard implicates), there are similar links between the latter (justificatory) proceduralism and a sort of proceduralism about reasons. In other

---

72 I am not concerned with realist interpretations of Kant here; these may well be immune to (many of) the criticisms I've raised of proceduralist approaches. See e.g. Langton (2007) for one defence of the realist reading.

73 I could potentially put this more strongly: if a generally-applicable buck-passing view is true (about the good, about the right, or whatever), then forms of reasons internalism will actually be *equivalent* to forms of proceduralism about the areas for which we should pass the buck (well-being, right action, etc.). So, for example, if the simple Humean theory of reasons were correct, and all reasons were entirely grounded in desires, and we should pass the buck when thinking about well-being, it would follow that things will only be good for a person to the extent that they desire them – which is a classic, if wildly implausible, form of proceduralism about well-being. See Scanlon 1998, chapters 2 and 3, for the canonical buck-passing story.

words, the necessity of incorporating a species of externalism in ethical epistemology – through recognising that virtue cannot be individuated without consideration of external-world relations – is mirrored in the necessity of incorporating a form of externalism into our theory of reasons. Recognition of these necessities requires rejecting the relevant varieties of proceduralism. The lesson is much the same throughout: no adequate normative approach – to well-being or the human good, to right action, to methodology, or to reasons – can be picked out using internalistic (or, which may be the same thing, reductively 'psychologistic') materials alone.

However, the most straightforward proposal about reasons that we can generate from the McDowellian position is fatally flawed. The eudaimonic motivation points us to the thought that the reasons a person has are generated by goods intrinsic to the situation in which she acts virtuously. However, this has implausible implications for most of the quotidian, non-ideal cases with which we should be concerned. This is because, *prima facie*, the reasons someone has if they are 'well-informed and well-disposed' will differ from those they will have if they are massively ignorant, lacking in an essential skill, or suffering from a deep-seated vice or irrationality (Williams 1995b, 189-91). A different way of putting this is to say that the straightforward proposal – which would say that the reasons someone has in a case are those they would be motivated by if they were virtuous – commits the conditional fallacy. Although it is true that a subject S would have reason to act virtuously *if* she were virtuous, it does not follow that S has reason to act virtuously *given that she is not* virtuous. This may appear to have the extremely unpalatable consequence of precluding the satisfaction of the eudaimonic motivation.

I think that the solution to this – avoiding the conditional fallacy while maintaining the eudaimonic motivation – is to identify a division between two largely-separate sorts of reason. Rather than try to transplant the reasons of the virtuous into the lives of the non-virtuous, we should accept that, although – because of the connection between virtue and the human good – all humans have reason in a broad and holistic sense to act virtuously, there is a different, equally important sense of 'reason' on which non-virtuous people usually have reason to act apart from virtue<sup>74</sup>. What a person has reason to do in a way that relates to the immediate context of deliberation may differ from what they have reason to do in general, in virtue of their good; if the situation is dire enough, the best someone can do is act badly. According to this model:

S has 'D-reason' to  $\Phi$  in C iff S\* *would want S to*  $\Phi$  in C (where S\* is S's closest fully virtuous counterpart).<sup>75</sup>

and also

S has 'V-reason' to  $\Phi$  in C iff S\* *would*  $\Phi$  in C.

The 'D-reason' formulation is inspired by the 'advice model' approach pioneered by Michael Smith (1995) and taken up by many others in the years since<sup>76</sup>. This two-part account does not commit the conditional fallacy, since the D-reasons that someone has are not proposed to

---

74 At this point, and throughout this section, I am talking about reasons as overall, all-things-considered reasons, rather than *pro tanto* or otherwise partial considerations. This helps to highlight the difficulties that I deal with here, but much of what I say would apply (*mutatis mutandis*) to e.g. *pro tanto* reasons too.

75 Robert Johnson argues that advice models like this fail to preserve the motivations for reasons internalism, precisely because they posit a connection to the desires of S' idealised counterpart *about* S's actions, which are not first-personal desires *to* act at all. Thus, the advice-theoretic connection cannot guarantee that S's reasons will be capable of explaining her actions. (1997, 621-5). Of course, this is not a problem for me, because I am not committed to internalism. (As a result, it is likely that my proposal will not be able to incorporate all its alleged benefits.)

76 Cf. Kawall 2014 (136-40), and the proposals discussed therein, for an alternative way of including both a virtuous-advice element and a direct virtue-replication element in thinking about different parts of practical rationality. Cf. also Svensson 2010, which poses important questions about the connection between virtue and right action, to which Kawall's view responds (2014, 140). My view is immune to Svensson's main argument, because the distinctions I've drawn generate two conceptions of 'right action' ('D-rightness' and 'V-rightness'), and these can be used to accommodate his worries. Unfortunately, I do not have space to examine the details here.

be identical to their V-reasons, except in the exceptional case where they are virtuous<sup>77</sup>.

V-reasons, as the eudaimonic motivation suggests, are reasons to *act well*; to realise the goods of virtue in one's own life. Contrastingly, D-reasons can be better understood as reasons to *do the best I can*<sup>78</sup>; these will, indeed, be identical to the reasons that my virtuous counterpart would want me to act on, since my virtuous counterpart would take my non-ideal constitution and circumstances (the things which prevent me from achieving my full good) into account. However, everyone has the reasons of virtue, regardless of the difficulties that their non-ideal situation throws up. In this way, a variant of Smith's 'advice model' can be made compatible with the Eudaimonic motivation, and thus with the idea that everybody has reason to live a good life, in some holistic sense, regardless of how distant from virtue they might be.

One option here is to build *all* motivational inadequacies into the set of barriers to virtue that generates the gap between V- and D-reasons. This would mean that no-one ever has D-reason to do something that they are not inclined to do because of their vicious motivations, and would immediately make the relevant conception of D-reasons internalistic in Setiya's sense. This does not seem plausible, however. To see this, it is necessary to look a little more deeply at McDowell's specific rationales for disagreeing with Williams' original (2012) version of internalism. This can be seen best in the following passage:

[T]o believe an external reason statement [on McDowell's construal] is... to believe that if the agent deliberated correctly, he would be motivated... in the direction in which the reason points. But there is no implication, as in Williams' argument, that there must be a deliberative or rational procedure that would lead anyone from not

---

77 Thus, the situation of virtue can be characterised – in addition to the characteristic features I've already noted – as that situation in which D-reasons and V-reasons are identical.

78 In Ch.6, I bring out a further connection here between the Eudaimonic motivation for an independent, largely non-action-guiding conception of V-reasons, and the controversy within political and social theory about the relevance of ideal-theoretic norms to non-ideal circumstances, which I discuss there in relation to the work of Lisa Tessman. See Ch.6 §3.

being so motivated to being so motivated. On the contrary, the transition to being so motivated is a transition *to* deliberating correctly, not one effected *by* deliberating correctly; effecting the transition may need some non-rational alteration such as conversion. (McDowell 1998k, 106-7)

McDowell's point here partly has to do with the recognition of a distinction between procedural rationality (conformity to the canons of logical inference, for example) and substantive correctness in practical reasoning (which involves the application of standards which are irreducible to any independent procedure). This is essentially a restatement of the anti-reductionist stance which has been my emphasis throughout this Chapter. If becoming virtuous (which everyone has reason to be) involves internalising irreducible rules, there will be no independently-specifiable set of patterns against which candidates could be compared which would provide an adequate test for the status of a consideration as a genuine reason. And, if virtue involves the formation of affective capacities, then the affective constitution of the virtuous will be equally shapeless<sup>79</sup>. It is thus implausible that there are any 'neutral' psychological states – individuated apart from the deployment of ethical concepts – which would be of essential ethical relevance as generators of reasons (even once narrowed down by procedural constraints).<sup>80</sup> However, it may seem that the erection of the distinction between D- and V-reasons does enough work to alleviate any worries here. This would leave it open for us to place all failures of motivational (e.g. affective) adequacy into the set of barriers that generates a gap between D- and V-reasons. Thus, in every case where an individual (in a given context of deliberation) had a motivational inadequacy (was viciously selfish, say, in failing to care properly about the needs of others), this would ensure that the individual did not have D-reason to do the virtuous (e.g. non-selfish) thing.

---

79 Another passage from 'Might There be External Reasons?' is similarly illuminating: 'from certain starting-points there [may be] no [procedurally] rational route...that would take someone to being as if he had been properly brought up. (Being properly brought up is not itself a [procedurally] rational route into being that way.)' (1998k, 102).

80 In this way, the rejection of (for example) the restrictive Humean conception of desires as reason-giving states is founded in McDowell's general 'anti-psychologism about psychology'; cf. 1998f (331).

This is not so, however. McDowell's response also includes the thought that Williams' conclusion 'depends crucially on the premise that a transition to a correct view of the reasons for acting that apply to an agent...must be capable of being effected, for the agent, by reasoning' (111). The motivation for the division between D- and V- reasons is not that there is some special privilege that motivation-preserving ('procedurally rational') reasoning has, which we should preserve by restricting D-reasons in line with it. Rather, the motivation for the distinction is to avoid conditional fallacies by ensuring that we do not assume that there are no genuine barriers to virtuous action in the relevant agent's context of deliberation. This would only amount to an incorporation of the internalist's fetish for procedural rationality if all of the motivational inadequacies which make virtue impossible (in the relevant context) were unalterable *tout court*, rather than merely unalterable by coming to understand a piece of procedurally rational argument<sup>81</sup>. As McDowell implies, this is implausible: sometimes, a person can be brought to recognise the salience of a consideration by *altering* their (e.g.) affective constitution in a way that (*ex hypothesi*) cannot fail to be procedurally arational. Given this, the difference between the McDowell-inspired approach to reasons I've sketched here and Setiya's 'reasons internalism' does not merely lie in the postulation of V-reasons – almost all of which are likely to be 'external', at least relative to some agents – but also in that many D-reasons will fail to fit within the internalistic mould as well.

This suggests a connection to some of the central points of the previous Chapter. There (§3) I noted that there are good reasons, flowing from work in feminist social epistemology, to think that ethical objectivity is likely to require transformative engagement with others, particularly

---

81 Cf. esp. McDowell 2009b.

in highly non-ideal circumstances. For example, against a background of serious epistemic injustice it is especially likely to be necessary to listen sympathetically to the vivid narratives of those in other social/cultural locations, in order to enable emotional engagement to overflow the barriers to understanding that unjust marginalisation has erected. There are likely to be cases of this type in which the affective-dispositional distance between agents is not so great as to preclude better action *per se*, but is great enough that communication of the grounds for better action will not be able to draw on affective similarity to a sufficient extent to count as procedurally rational. Dealing adequately with this sort of case will require externalism about D-reasons, in addition to that about V-reasons.

### **§2.2.1 Excursus: On the Barriers to Virtue**

I have just argued that the set of barriers to virtue<sup>82</sup> – that which causes D-reasons to diverge from the reasons of virtue proper – does not include all inadequate motivational dispositions. In this subsection, I shall give some examples of the sort of thing that probably *should* be included, and note some further connections to feminist epistemological ideas along the way. Of course, I do not pretend that this is a comprehensive list, only an indicative one.

Firstly, some individuals may lack skills that would be required to act well in a context. Familiarly, for example, someone might lack the ability to recognise the contours of a relatively alien cultural practice, with the result that they are not capable of treating people respectfully within that practice. Because such skills take significant amounts of time and

---

82 What I say about the distance between D- and V- reasons in non-ideal circumstances does not entail that one's conception of virtue and its constituent goods never has any direct effect on what people should do when virtue itself is unavailable to them. It is at least likely that reflection on the goods of the virtuous life should inform one's general sense of one's own life – regretting and expressing regret about some things, for example, and imaginatively discussing with others how our situations or ourselves could become more amenable to fully virtuous agency. See Ch.6, §3 on this point.

effort to develop, the lack of them may render virtue impossible within the immediate context of deliberation and action. Secondly, there are persistent psychological flaws such as phobias, delusions, and biases. Especially when opaque to introspection, or immune to interruption by conscious effort, these are likely to function as barriers to virtuous agency which cannot be (fully) overcome either by individual reasoning or by transformative engagement by others within the deliberative moment. Finally, there are traditional vices, such as selfishness, greed, envy, and resentment. If these are sufficiently well-established, neither procedurally-rational argument nor proselytical intervention may be adequate to disrupt them, and enable action from virtue instead.

These examples help to illustrate the category of barriers to virtue that I think important for clarifying an acceptable theory of reasons, and preserving both the eudaimonic motivation and the demand for guidance in non-ideal contexts. I shall now show how they shed light on the broader theoretical context as well. Building on my discussion in the previous Chapter (§3), there is a dilemma for virtue-centred feminist epistemologies between treating virtue as a socially *dependent* trait, in the sense that no-one can be virtuous unless their social context is constructed such as to enable excellent action, and treating virtue as a fully *social* trait, with its locus at some level higher than the individual agent. These reflections on reasons and the barriers to virtue shed more light on these alternative options, in a way that I believe ultimately favours the former.

Firstly, take the absence of skill. In some cases, it might be possible for the social context to 'compensate' for this by providing information (relevant to the pursuit of the good) which the skill would have made available if properly developed. On the individual-centred approach,

the fact that the information is provided need not be a sign that virtue is realised in the situation, because although the individual may do what they should have done<sup>83</sup>, they did not do it from virtue. On the group-based approach, it is less obvious that there is anything bad about this case. Here, which approach seems better justified will depend on whether it is truly the *skill* that is necessary for excellent action, or only the information that the skill is needed to impart. If the latter, the group-based approach seems perfectly accurate. But if the original description was correct, and the skill itself is a constituent of good action, the individual-centred viewpoint is superior. Secondly, take a standard implicit racist bias. Here, other individuals may be able to overrule biased decisions, or alter the context of deliberation by, for example, anonymising candidates for a position. In this case it seems clearer that, even if the alterations prevent seriously harmful action by the agent, virtue is still not present, and the agent does not truly act well; clearly, this favours the individual-centred approach. The final sort of case, on the other hand, seems inconclusive, because – given that we are ruling out the possibility of affectively transformative interventions – the social context is unlikely to be able to compensate for serious vice in a way that might even appear to maintain virtue at the group level. Neither approach to the issue of location would argue that virtue is present in this sort of case, then.

The differences between these approaches to virtue are significant in two ways. Firstly, because the move up to the social level might have weakened the case for the division between D- and V-reasons, by making room for argument that excellent agency is not out of reach after all, so long as such agency is located at a level which includes multiple individual agents. However, I hope I have done enough to suggest that such a move is not very plausible

---

83 That is, perform the right sort of act, where act-types are individuated independently of their etiology (insofar as that is possible).

– social context is important for the identification (and cultivation) of virtue, but the locus of virtue is still the individual agent. Secondly, because the usefulness of looking to the social context illuminates the worth of the virtuous advice model. In many non-ideal cases, there is an extent to which other individuals can function as (probably partial) proxies for the virtuous version of the agent, and thereby help them to act better than they otherwise might. The ultimate advantage of this bifurcated perspective on reasons is that it makes proper sense of improvements such as these without pretending that they could amount to genuinely good agency.

### **§2.2.2 Defending the Divide: Alienation and Blame**

If what I've said is correct, then Setiya's 'reasons internalism' is probably false, even when formulated narrowly about D-reasons. In the remainder of this subsection, I shall return to consider how this thesis can be defended against Williams' later criticisms<sup>84</sup>. In so doing, I shall also provide a more detailed sense of the different roles of D- and V- reasons in ethical thought and practice.

I've postulated two kinds of reasons: D-reasons, whose function is to enjoin someone to do the best they can, and V-reasons, whose function is to enjoin people to act well. Technically, whether or not this model amounts to a denial of Setiya-style internalism turns on whether there is always a path from the actual situation of S to the situation of S's virtuous counterpart that is sufficiently<sup>85</sup> motivation-preserving. As I've argued, it seems likely – although it is an

84 Of course, there are many other, often more subtle criticisms which would apply to my claims but which I do not have space to consider here. I have in mind especially the non-Humean defence of reasons internalism given in Markovits 2011, responding to which would require a quite different argument than the one I've given here.

85 There is significant room for argument about (e.g.) how long temporally or in number of steps such a path can be while remaining consistent with the motivations for internalism, and this 'sufficiently' qualifier stands in for a solution to this area of dispute. Clearly, I cannot provide such a solution here.

empirical question whose answer is not entirely obvious – that there are very many V-reasons, and even some D-reasons, which would be 'external' in this sense. If so, it appears that I am resolutely opposed to Williams and his many defenders regarding this theoretical point. It is thus worth reflecting again, in concluding my discussion of this issue, on Williams' original response to McDowell, only part of which turned on the problem of the conditional fallacy. Williams' other counterarguments are twofold. The first concerns the prospect of alienation, the second concerns blameworthiness.

My discussion of the alienation question will be relatively brief. As I have already said, in discussing the accusation of alienation that may be an undercurrent in Jagger's normative critique of Nussbaum-style methodology (Ch.3 §2.1.1), it is both clear that I have little space to discuss first-order ethical issues here, and equally clear that the issue of alienation is a difficult and substantive one; it cannot be taken for granted that alienation is unproblematic, but it equally cannot be taken for granted that causing alienation will always be unjust (or impermissible, or inappropriate in other ways). Perhaps some alienation is a price which must be paid for a more just world; perhaps inducing alienation is sometimes no harm at all. On the other hand, it may seem more damaging to raise the spectre of serious alienation in the context of claims about *reasons*. Williams' version of this point (1995b, 191-2), updated for my purposes, would turn on the idea that if the link between the D-situated individual and her would-be virtuous counterpart is too attenuated, suggesting that she ought (in any sense, however distanced from her immediate non-ideal choice situation) to act virtuously amounts to suggesting that she become someone else. In the most extreme case, this might be suggested to amount to a kind of psychic suicide. However, this argument would involve reading quite a lot into Williams' text. All Williams really says to suggest a sense of alienation

is the following:

the fact that [ $\Phi$ ing] can be a worthwhile activity surely does not give everyone a reason to engage in it... from both an ethical and a psychological point of view, it is important that [a reasons claim] should say something special about [its target], and not merely invoke in connection with him some general normative judgement. (1995b, 191-2)

There are two elements here. Firstly, Williams implies that there cannot be any reasons which would hold, not due to anything 'special' about their target, because they are simply universal in scope, being anchored in the good rather than the vagaries of desire<sup>86</sup>. Clearly, this is brutally question-begging, and I need not respond to it seriously. However, the claim that vague truths about general 'worthiness' do not immediately entail a reason for every person is plausible. Nevertheless, if any particular universalistic account (e.g. of dignity) is compelling in its own right, this will provide for more than this ambiguous idea of worthiness. I suggest, then, that we can sensibly pass the buck for a defence against this sort of internalistic attack to whatever can be said in favour of the idea that some reasons (e.g. of justice) are genuinely reasons for all humans, which I shall say more about in the next Chapter (and finally in Chapter 6).

The second issue is more involved, but is also more suggestive of ways to develop the D-/V-reasons contrast I've identified. It centres on Williams' claim that the role of blame as a reactive attitude (Strawson 1962) cannot be properly made sense of by non-internalist theories of reasons. For Williams, there are three types of cases, all of which involve what he calls 'focussed blame', which is blame directed at another with reference to something they ought to have done (or not done); as he notes, just as there is a connection between 'ought'

---

86 Williams immediately goes on to consider an Aristotelian naturalist basis for linking general considerations with individuals, but the version of such a view that I propose in the next Chapter would be unsuitable by his lights. It does not (unlike the view he considers (1995b, 192)) involve any claim that humans (*qua* human) are capable of the good life in a way which would entail the presence of the necessary 'internal' motivational materials.

and 'can', there seems to be a connection between 'ought to have' and 'could have'. He also thinks that 'ought to have done' is connected to 'had reason to, but did not' (1995a, 40-41). In the first, 'easy' case, an agent is blamed for  $\Phi$ ing, where they had internal reason to not- $\Phi$ . Here, the propriety of blame is supposed to be clear, since blame may bring to the agents' awareness the considerations which would have lead them not to  $\Phi$ , had they deliberated (procedurally) correctly (41). In the second sort of case, an agent is blamed for  $\Phi$ ing, but they did not have an internal reason<sup>87</sup> to not- $\Phi$ . However, their motivations happen to be such that the very fact of having been blamed (e.g. by a person who they respect) will transform their motivational constitution so that they gain an internal reason that they did not have before. Thus, proleptically, the presence of an internal reason to not- $\Phi$  is ensured, and the propriety of blame is, again, clear. (41-3). In the third sort of case, however, an agent is blamed where there is no internal reason to not- $\Phi$ , and also no prospect of proleptically bringing about such a reason. In a 'hard' case like this, Williams states that blame is 'inappropriate', because it would make no sense to blame someone for something which they could not be brought to recognise was not well-supported by reasons (43).

In responding to Williams<sup>88</sup>, I should immediately concede that V-reasons should probably be disconnected from blame. Williams is probably right that blame is essentially aimed at influencing those who have done something that they should have not (and thus could have not, within the relevant deliberative context) done. If a deliberative context is such that the

---

87 At least, not an 'all-things-considered' internal reason.

88 Miranda Fricker has recently proposed a structurally similar – if significantly improved in detail – account of blame (2014). Her most important innovation for my purposes is the suggestion that the induction of 'remorse', as a self-directed reactive attitude to one's wrongdoing, is the constitutive aim of blaming. (On her account, blame is essentially a communicative practice). This is consonant with what I shall suggest in Chapter 6: that claims about V-reasons are aimed at the different, and somewhat (phenomenologically) weaker, attitude of *regret*. I believe that her account is a good one to capture one practical role of D-reasons, at least as long as it manages to avoid the first critique below. I suspect, however, that it will be vulnerable to the second, fetishism-based critique, and would need to be revised accordingly.

virtuous action is unavailable to the agent – or (technically) available to them, but such that they are exculpated from doing it (typically, because it would involve overwhelming costs) – then they cannot meaningfully be blamed for their failure of virtue. This makes it easy to meet his challenge of explaining how external reasons are connected to the appropriateness (or inappropriateness) of blame. This is because the inappropriateness of blame in 'hard' cases will, on my view, line up with a lack of (D-)reason to  $\Phi$ , just as it will for Williams. What is more, the appropriateness of blame in both 'easy' and proleptic cases can be explained in the same way: in both cases, it is appropriate to blame because blame stands a chance of inducing someone to recognise the reasons they had to act other than they did. However, building a positive case for externalism about D-reasons may now seem more difficult. I have resisted internalism about D-reasons on the ground that transformative engagement can alter motivational dispositions in a clearly reason-promoting way, and internalism neglects this. However, the proleptic mechanism that Williams identifies might seem to do the same job as such transformative engagement. There are two critical differences, however.

Firstly, consider proleptic advice, rather than proleptic blame (since blame is backward-looking, and interventions within the deliberative context must take place before the action does). Williams can justify the claim that proleptic advice can be an appropriate kind of action for a third party (42-3). However, explanations of such appropriateness cannot refer to the reasons that the advisee has, since, at the time the intervention is mooted, Williams does not think that such reasons exist. It seems perverse, though, to insist that the justification for advising someone is not based in their reasons even though it aims to induce recognition of such. The adviser would be aiming to induce recognition of a reason that the advisee did not have! An externalist understanding of D-reasons has a simple, and more plausible, story –

transformative advice aims at inducing recognition of reasons which already exist.

Secondly, consider the nature of the desires which secure the success of proleptic interventions such as those that Williams relies upon. On the face of it, these desires have to be unguided by the broader set of reasons, because they are supposed to alter that set from the outside. Apart from the fact that this would make the process of proleptic transformation less reliable, it also appears to be fetishistic: it imputes a significant desire for the wrong sort of thing. In order to see this, it is necessary to bring out the notion of 'motivational' (or 'judgement') internalism, since it is in terms of this concept that debates about fetishism have been framed. Schematically, motivational internalism is the view that ethical commitment (belief/judgement/approval, etc.) is strongly linked (in some specified way) to practical motivation. More specifically, some (e.g. Michael Smith) have argued that there is a general *rational* connection between ethical commitment and motivation. However, this general proposal turns out to be problematic. Hallvard Lillehammer (1997) provides a salutary reminder of the bad consequences that would result if we were to incorporate such a general rational requirement for motivational internalism<sup>89</sup>. Lillehammer notes (194-5) that it seems perverse to insist that just *anyone* would be irrational not to be motivated in accordance with their ethical judgements. For, if someone makes horrific moral judgements, it seems both possible *and rationally welcome* for them to fail to be motivated in accordance with them. As Lillehammer says (195), having a content-neutral motivation to act in accordance with one's ethical judgements is a kind of fetishism. Likewise, it is rationally preferable that an agent *not* be moved by (e.g.) general respect for an intervener to value things differently (and thence to act differently), if the intervener is *not*, in fact, relevantly virtuous. Having a

---

89 Cf. also Johnson 1999, 69.

content-neutral disposition to be transformed by interventions by others is, thus, equivalently fetishistic. The solution is to regard the desire<sup>90</sup> (on the part of S) as a desire to be transformed by those in a *better epistemic position*<sup>91</sup>; not merely a desire to conform to the expectations of others. Here, as elsewhere, we have to build substantive presuppositions into the motivational basis for proleptic transformation, if we are to preserve the possibility that transformation might be a reliable way to come to act on genuine reasons. Here, as elsewhere, we should not fetishise *any* merely procedurally rational process, proleptic or not.

It is important, now that fetishism is in the picture, to make explicit why my account is not vulnerable to Smith's original (1994; cf. especially Smith 1996, 180-83) accusation of fetishism against externalism. Smith argues that motivational externalists must appeal to a merely *de dicto* desire to do what is right. He thinks that this goes for 'good people' as well as thoroughly imperfect agents. This seems fetishistic because (at least) the virtuous should be motivated by (e.g.) the prospect of alleviating suffering (*de re*), and not the rightness of alleviating suffering *de dicto*. However, the foregoing discussion should make clear that this criticism could not apply to the virtuous themselves, on a McDowellian account, since for them motivational internalism is effectively true: they will be motivated by their reasons directly. In addition, it also does not apply to a non-virtuous target of a D-reason claim, at least if my argument in the previous paragraph is correct. This is because we need not regard S as being transformed by S\* (or, rather, some proxy for S\*) on the basis of a desire to be respected (*simpliciter*), or to 'do the right thing' *de dicto*. Instead, they will be transformed – if they are transformed in accordance with substantive ethical reason – on the basis of the

---

90 I am using 'desire' here in a broad, non-Humean sense, according to which desires can have cognitive elements. Cf. §2.1 above.

91 Of course, this is not to say that this desire should not (or even could not) have affective elements too – a desire for something like respect (or love) is perfectly appropriate, but will only have no whiff of fetishism when it is guided by substantive reason.

thought that<sup>92</sup> S\*'s intervention will bring them to see things aright. It is true that, before S\*'s intervention, S is incapable of being moved by the ethically-salient features of things *de re* (her genuine D-reasons), but at every stage of the deliberative process she is being moved by something with a genuine (*de re*) ethical significance. Before the transformation, she is capable of being moved by virtuous proleptic advice (*de re*), and after it she will be capable of being moved by her D-reasons themselves (*de re*). *Contra* Smith, it is not fetishistic to desire a genuine improvement in one's (ethically-relevant) epistemic position, even though it would be fetishistic if one's desire did not build in a sensitivity to the way things are, epistemically speaking<sup>93</sup>.

These two advantages help to motivate my division between the two sorts of reason. More generally, I have argued throughout §2.2 that a more complex account – of the conditions in which reasons exist – is needed than has generally been provided. In Chapter 6, I shall draw on this account to argue that the sufficientarian capability approach to global justice should be interpreted as (primarily) a theory about the V-reasons of justice that all human beings possess. This McDowell-inspired account of reasons thus forms another important part of the framework I am proposing for the capability approach.

### §2.3 Anthropocentricity and Universalism

The bulk of the argument of this Chapter is now complete. In this final subsection, I revisit the anthropocentrism of the approach I've defended. McDowell's secondary-quality realism takes seriously the thought that ethical facts may be species relative, both in the sense that it

---

92 Where the motivational basis for the intervention is non-conscious, they will not be motivated by this *thought*, but by a some complex personal disposition whose content is functionally equivalent; I do not think this makes any significant difference to the argument. Cf. McDowell 2009c, 33-40.

93 Compare McDowell 1998l on the argument of this paragraph.

is only beings with human-like faculties that are capable of comprehending them, and in the sense that the kinds of practice that they are facts *about* may also be uniquely human. In the next Chapter, I explore one way of making sense of universalism, which develops precisely that thought: that the modes of relationship that justice regulates, and so the need for justice itself, are characteristic of our animal species. In this way, it may be a significant advantage of a framework for an approach to global justice if it can coherently include some restriction of this kind – this will help to rationalise the central commitment to human-universalism about justice that is a standard presupposition of cosmopolitanism<sup>94</sup>. One possible difficulty here is that, as I have noted, McDowell's position conjoins two kinds of species-relativity that could, in principle, come apart: the capacity for knowledge of the norms of a practical domain might be separable from the capacity to engage in the functionings that are distinctive of the domain itself. However, unless there is some particular reason to suppose that these joint restrictions are not effectively aligned with one another, this does not seem to threaten the plausibility of the view. Likewise, it appears the overall position, too, cannot be a *problematic* sort of relativism. If a different life-form were discovered that lived in much the same way as us, having (e.g.) a similar emotional constitution, and a capacity for ethical debate and practical co-ordination, this element could be expanded to include them. The species-relativity here is (in both respects) of a defeasible, flexible sort, because the thesis of anthropocentricity was originally motivated precisely by the hypothesis that our faculties *are* unique in this way<sup>95</sup>. For this reason, it could not do real violence to the approach to alter this assumption, if this became reasonable in its own right.

---

94 See Chapter 6, §4.1 for more argument in this vein.

95 See, for instance, McDowell 1998n, 118. Elsewhere (1998m, 169-74, esp. 171) McDowell acknowledges that parts of human nature (as well as the nature of other possible animals such as his 'rational wolf') play a role in enabling the acquisition of reason (and, presumably, thereby all of the many activities which reason is required to regulate).

The role of this Chapter has been to make clear the deep congruence between A) the kind of epistemology that expansive and demanding approaches to global justice like capability theory require and, B) McDowell-style quietistic metaethics. In essence, the attraction of this combination can now be stated relatively simply. Any plausible capability approach (one which meets the desiderata I gave in the first two Chapters) requires us to make sense of the integration of substantive values as constraints on methodology, when engaging in inquiry about (e.g.) the content of dignity. In particular, we must leave open the possibility that, in making decisions about how to formulate our ongoing methodology, we must rely on judgements which have been generated by that methodology itself. We are likely to have to do this even when the judgements are distinctive of particular cultural locations, in a way that might seem exclusionary. As this Chapter has described, McDowell's view is capable of making sense of these necessities in a compelling way. Along the way, I have also noted many other points at which McDowell's anti-reductionist position can impart a valuable flexibility to ethical practice, opening up a broad space of possibilities for moral psychology as well as epistemology. Occasionally, this flexibility has also been shown to help in avoiding certain first-order errors, e.g. with a restrictive psychologistic theory of reasons. Finally, I have just begun to bring out the possible connection of this metaethical framework with neo-Aristotelian naturalism, a view which has great promise for making sense of the universalism implicit in approaches to global justice. Developing a version of such naturalism which fits with everything I've argued for thus far then forms the task of the next Chapter.

## Chapter 5:

### A Circle: Humanity, Well-Being, Justice, Virtue

In the last Chapter, I characterised a metaethical position, that of John McDowell, that I hold is a plausible take on the central questions of metaethics, and that I contend can, as part of a wider collection of views, provide the framework for a plausible capability approach to questions about justice (*inter alia*...). In this Chapter, I address the other main element of that collection: neo-Aristotelian naturalism. In brief, neo-Aristotelian naturalists believe that the fact of *humanity* – the fact that we are the particular kind of animal life that we are – is of major significance for normative thought, and that Aristotelian ideas are needed to explain how this is so. This is a position that has undergone a significant revival in recent decades, in parallel with (if lagging somewhat behind) the general revival in Aristotelian normative theory. However, the interpretation and evaluation of such claims has been fraught, with most critical attention from major non-Aristotelian philosophers being dismissive, where it has occurred at all<sup>1</sup>. Especially troubling for my purposes is an apparent tension between neo-Aristotelian naturalism and the McDowellian theses I endorsed above. This tension emerges in two places, one quite general, one more specific. Generally, there seems to be an incompatibility between the quietism of McDowellian metaethics and the apparently substantive nature of naturalist metaethical claims<sup>2</sup>. Specifically, neo-Aristotelian naturalism seems to threaten the thoroughgoing *sui-generis* or conceptual autonomy of normative, and

---

1 Examples, which I discuss throughout the text below, include FitzPatrick (2000); Lewens (2010, 2011); Copp and Sobel (2004), Skorupski (2012).

2 This general issue is of much less significance for my project in this Thesis, but I engage with it a little in Appendix B.

in particular ethical, discourse and practice. My project here, then, is to elucidate and defend a version of neo-Aristotelian naturalism such that these tensions (and other difficulties) can be seen to be illusory, thus preserving the benefits of the framework I've been building to this point.

I begin §1 by reintroducing Nussbaum's version of neo-Aristotelian naturalism. Then I build from these basic ideas by giving a brief sketch of the role that neo-Aristotelian naturalism plays in fleshing out and justifying a capability approach to global justice. §2 then describes the core thesis of neo-Aristotelian naturalism itself, as it figures in the work of Michael Thompson, Philippa Foot, and Rosalind Hursthouse. After an outline of these proposals, §2.2 describes an important challenge to them, which centres on the idea that they are problematically reductive. §2.3 assesses the extent to which such challenges are likely to be successful, concluding that the threat posed is significant, and warrants a definitive turn away from reductive treatments. §3 is then composed of the development of a non-reductive alternative, which coheres tightly with the McDowellian metaethical view that was my concern in the previous Chapter. Issues covered in the course of this include the ultimate relationship between the human good and biological humanity (§3.1), and the right (or virtue) and the good (or well-being) (§3.2).

## §1 Introducing neo-Aristotelian Naturalism

### §1.1 Naturalism in Nussbaum

The principal inspiration for the neo-Aristotelian naturalist account of the foundations of capability theory that I shall offer is, unsurprisingly, the work of Nussbaum. Naturalism crops up in two major ways in Nussbaum's work. Firstly, as I made clear in Chapter 2 (§2.3.4), the

conclusions of the clearly constructivistic Rawlsian elements of Nussbaum's approach rely – if they are to cohere with her substantive good methodology – on a thesis of convergence. It is difficult to explain this thesis, or even to make it out in detail, unless human nature is such that *homo sapiens* is intrinsically oriented towards the mind-independent good. In this way, she apparently implicates a non-normative, extra-ethical account of human nature. I do not have much to say about this here, but it is worth noting that, since the version of neo-Aristotelian naturalism that I propose below requires only an ethically-laden concept of human nature, my version of the grounding for capability theory may have fewer commitments than hers<sup>3</sup>.

Secondly, and more importantly, Nussbaum also makes use of a normative conception of human nature<sup>4</sup>. This is explicitly invoked in order to provide a framework within which to fit theoretical approaches (i.e. the capability approach). Humans have a good, and we arrive at a conception of this good by reflecting on what is significant about being human. When we do this, we provide support for a capability approach, in several ways. Firstly, in that we can form a conception of the virtue of justice (one of the things that makes lives fully human), which should then be construed according to the capability approach (see her (1993), and (1995)). Secondly, in that we can form a conception of the general ingredients of the good life; these will then be the things that people should be enabled to pursue (especially in her (1992), and most others). Thirdly, in that general reflection on the most distinctively *human*

---

3 This is not to say that I think no invariant core of human similarity exists, apart from the ethically-delineated one that I rely upon here. My reticence to involve such a conception has more metaethical motivations than scientific ones – if one claims that a common nature is leading us towards the ethical truth, but this nature is not itself dependent on ethical evaluations, there will A) be little or no room for a culture-laden conception of objectivity, which I have urged is necessary, and B) be little explanation for why this common nature has this truth-conducive effect, when it is entirely unrelated to the truths it supposedly leads us toward.

4 It may be helpful to compare my initial treatment of Nussbaum's invocation of humanity as an element in her theory of justice, in Ch.2 §1.1.

elements of our lives ('practical rationality', and (sometimes) 'sociability') helps to support a capability approach over alternatives, including one based solely on functioning. ((1998), amongst others). In addition to these roles, the conception of humanity also clearly plays some role in delimiting the space of beings whose goods can make claims of justice, at least until her (2006, 325-407), where non-human animals are incorporated into her capability approach.

## §1.2 Sketching a neo-Aristotelian Naturalist Capability Theory

Given this background, I can begin to elaborate the actual content of a neo-Aristotelian naturalism, suitable to support a capability approach to global justice. The detail of many of the claims that I make in this section will not be fully clear until later in the Chapter, or in the next Chapter (esp. §1), but I hope they, and their various interconnections, will be transparent enough to begin with.

We should start, as Nussbaum does, with a conception of a distinctively *human* life. This is a life characterised by a certain range of functionings. These extend from classic 'doing' functionings such as interpersonal communication, to 'being' functionings like having success in personal projects; from functionings as mundane as having appropriate calorific intake, to those as sophisticated as the development of an aesthetic sensibility<sup>5</sup>. As I shall note, many recent neo-Aristotelians, drawing on work by Michael Thompson, link functionings very strongly with humanity itself. For Thompson, concepts like 'humanity' – 'life-form concepts' – are normative or quasi-normative concepts which unify sets of functionings together. What being human (or, *mutatis mutandis*, any other organism) actually *amounts* to is having a

---

5 See Appendix A: 'Bodily Health', and 'Senses, Imagination, and Thought'.

property which sets expectations of functioning, and thereby both plays a role in explanation when functionings match the expectation, and generates judgements of defect when they do not (1995; 2004a; 2008). Theoretically, this implicates a limit case – e.g. the perfectly human being<sup>6</sup> – that would instantiate each characteristic functioning fully. Being human *simpliciter*, however, need not involve the possession of most (or even many) of the 'normal' functionings. Rather, I suggest that humanity can be identified (as a concrete, worldly property rather than a concept) with a disposition whose *end* is species-typical functioning, but which may be very easy to frustrate if conditions are not apt<sup>7</sup>. I will ultimately suggest that the attribution of this disposition is a fundamentally ethical matter. However, there are familiar examples of minimal species-typical functionings (e.g., perhaps, a certain range of genomic information, recognisably human morphology, involvement in human relations, predictably human cognitive function) which will inevitably have to be used as evidence for the presence of the disposition in particular cases. (It is unsurprising that they would be evidence for it, since it is, in part, a disposition to produce exactly those features.) As well as providing evidence of its presence, some of these statistically-normal features (when placed within a sufficiently complete and coherent framework) can be thought of as providing the supervenience-base of the disposition.

---

6 This, like similar concepts I refer to elsewhere, such as 'the virtuous agent', might seem indicative of a narrow, perhaps elitist, privileging of one highly-specific kind of person. The connotation is misleading. There is, on any plausible view, more than one (indeed probably an extremely large range) of relevantly 'perfect' lives that are possibilities. Likewise, although there might be only one virtuous course of action in some given single situation, there will in reality be many very different ways to live an entirely virtuous life. And, needless to say, there is no reason whatsoever to think that the perfectly human being (or the virtuous agent) would resemble the powerful more than the powerless.

7 This may be an idiosyncratic interpretation – it is generally unclear what kind of worldly thing species-natures are supposed to be on most neo-Aristotelian accounts. Thompson, for instance, is more interested in giving a logical account of species concepts. Foot and Hursthouse do not appear to discuss the issue at all. Nussbaum, even in her later work (e.g. 2006, 181-2), in which she drops the concept of 'basic capability', seems to think of humanity as a range concept, with actual functioning within the range being necessary for human status; Hope 2013 (162-3) provides a compelling criticism of this element.

Turning back to the explicitly ethical realm, such features can also be equated (in terms of their role) with Nussbaum's 'basic capabilities' (1987, esp. 25-30) – the minimal human capabilities which, when present, generate demands for the improvement of the life that involves them<sup>8</sup>. Appropriately, the same basic conceptual structure will then apply here as there, for the purpose of interpersonal comparisons of well-being: being human, just as such, ensures the applicability of certain judgements about relative goods and the absolute quality of lives. Specifically, every human is benefited by something to the extent that it brings them closer to the full attainment of the life-form concept, and harmed to the extent that they are distanced from it<sup>9</sup>. And a human has a good – i.e. evaluatively *adequate* – life to the extent that it meets *enough* of the life-form concept.<sup>10</sup>

With this structure in place, I can turn to consider the place of justice in the lives of humans. Nussbaum, as I have said, treats justice – in her similar development in *Non-Relative Virtues* (1993) – as a good concerned with the proper ordering of the public sphere, and especially with distribution. This is an option for me here, and if I followed her in this it might imply little substantive change. However, because of the problems with public/private distinctions (see Chapter 2, §1.3), and the thin conceptions of institutions and the Rawlsian 'basic

---

8 Note that, strictly speaking, capabilities *are* functionings, at least according to the definition I gave in Ch.1 §4 – they are 'being' functionings, consisting in an ability to effectively choose valuable options. Capability thus turns out to be a subset of functioning, which makes the theoretical exegesis simpler here. Compare MacIntyre 1999, 65 here.

9 At this point, this may sound like an extremely outlandish claim. I hope that, by this Chapter's end, it will not seem very radical at all.

10 It is important to remember that here, as always, I am engaged in a Neurathian method. The prospect of finding a way of making different sorts of judgement – about humanity *simpliciter*, about well-being, about justice, about virtue – cohere with one another is motivation enough for identifying, in this instance, a number of different levels within the species norm; one for human dignity, one for adequate exemplification, one for perfect exemplification. At no point is it necessary to claim that the existence of such levels would be plausible in abstraction from the ethical context. For example, judgements about 'adequately doggish' lives might not make sense in the way that adequately human ('good') lives do, if we did not engage in evaluation of the lives of dogs. Or there might be just two levels for some life-form – one of minimal exemplification, and one of perfection, with no levels of adequacy in between. But the inapplicability of such details in non-human cases doesn't damage the ethical motivation for proposing them, or threaten the universality of the basic structure of life-form concepts more generally.

structure' on which they encourage a focus, I shall use a different concept to do this work. Rather than treating justice as a virtue dealing with institutional orders, I shall treat it as a virtue concerned with interpersonal relationships – all of which will, of course, be mediated by such orders (to varying extents). The complex sociality of human beings is surely one of our most important traits, and so it should not be surprising that a functioning dealing with interpersonal connections will be correspondingly central. Complete justice, *vis-a-vis* a given individual at a given time, will require all of that individual's relationships to be adequately good. If we presume that every human being (*per se*) is disposed to form relationships, of some appropriate (perhaps very minimal) kind, with every other human being<sup>11</sup>, given the right sort of circumstances, it will follow that no human being will be just, and thus that no human being will be able to live a good human life, unless they are living in a just world – a world in which relationships are apt to adequately flourish<sup>12</sup>. It follows from this that every human is harmed by global injustice, because many, if not all, of *their* relationships, and many more of the possible relationships that they are essentially disposed to acquire, are damaged by it. It follows in turn that every human being has reasons, of some sort at least, to promote justice on a global scale.

This provides an entrance point for capability theory proper. For, plausibly enough given the arguments summarised in Chapter 1, what global justice requires is that everybody be enabled to live a decent life. Not a perfect life, that is, or a life that should not, in any sense, be better than it is<sup>13</sup>, but a life that is not so bad that its condition must damage the relationships it

---

11 A more thorough defence of this claim is given in Chapter 6.

12 We could, as Nussbaum appears to in her *Frontiers of Justice* (2006, 325-407), hold that humans have minimal default relationships with non-human animals (or some sub-set thereof), as well as with humans. This would provide the resources for a critique of injustices against these beings too. Although I do not see any deep barrier to the acceptance of this claim, within the framework I set out here and in the following chapters, I will not consider it further.

13 By contrast, one way of describing the level of well-being that I am calling the *good* life is precisely as that

involves. Using the term I have deployed throughout, what justice requires, at the minimal level, is that everyone have *dignity*, because the lack of a life of dignity cannot fail to damage all the relationships that the individual is involved in. This provides the rationale for a sufficientarian conception of threshold values of the various functionings and capabilities that are constitutive of such a life. The claim, ultimately, is that no-one can have a truly satisfactory life unless the world is just – and the world cannot be just until all humans possess the threshold value of every valuable capability<sup>14</sup>.

If this view is defensible, the achievement of the principal purpose of this Thesis would seem to be at hand. A basis for a capability theory of justice, similar to that of Nussbaum, would be available that would cohere with a plausible approach to meta-level issues. Justice would be connected to human well-being, both because bad relationships are a clear detriment to those who have them, and because the content of justice is partly determined by what people need to live decent human lives<sup>15</sup>. And universalism and robust objectivity of judgement would be assured, at least in principle, as long as neo-Aristotelian naturalism can be made to cohere with the basic McDowellian metaethic I've presented. There are number of challenges remaining however, which will be dealt with in the remainder of the Thesis. The rest of this Chapter will involve fleshing out the sketch of a neo-Aristotelian naturalism that I have just

---

level of functioning such that, although it possibly *could* be made better, there is no sense in which it *should* be. This marks the point at which relative comparisons of well-being, and claims about how people should live differently, start to lose their practical force.

- 14 This is quite a strong claim – it implies, given that the world has (plausibly) never been adequately just, that no human being has ever had a good life. (Although they may have had very good lives in any number of particular respects *apart* from justice, and may have had many good particular relationships *inter alia*). As with many others in this Chapter, this is a significantly less controversial claim than it appears to be. It is a commonplace, at least amongst those actively concerned with global justice, that the contemporary world is highly unjust. If you add in the view that injustice precludes the possibility of proper relationship, it will follow that everyone in the contemporary world is harmed significantly, and it is not a large step from there to the view that no-one has a good life: a life about which one cannot say that it should be considerably better.
- 15 This sort of account, then, has good prospects of incorporating attention to the sociality and dependency which are MacIntyre's emphases in his 1999 – relation-dependent goods will clearly be relevant to dignity when justice itself is conceived of as a relational matter.

given. Finally, in the last Chapter I fill in this sketch, providing a more thorough understanding of justice in the neo-Aristotelian frame, and then finally drawing the various threads of the Thesis together, showing how they amount to a strong supporting background for a capability approach to global justice.

## §2 Reductionism?

I have offered an outline of a framework for capability theory. But I have said little about the neo-Aristotelian naturalism that plays such a central role in it. In this section I remedy that, offering a description of the core claims of recent Aristotelians in the area, and assessing their vulnerability to an important sort of challenge, which threatens to derail the entire project.

### §2.1 The Core View

To put it simply, neo-Aristotelian ethical naturalism is the view that what humans *are* as natural, animal beings bears on what is good for them, and on the way they ought to live. Unlike the view of Nussbaum that I have described, however, most neo-Aristotelian naturalists do not view the conception of humans as animal beings as being a straightforwardly ethical conception. How, then, does human nature come to have ethical import, on these views?

#### §2.1.1 Life-Form and Normativity

The basis for the neo-Aristotelian answers to this question that I shall consider<sup>16</sup> has been

---

<sup>16</sup> Probably foremost of the views that I shall *not* consider here is that of MacIntyre in his *Dependent Rational Animals* (1999). In some ways, his proposal there mirrors the structure of my own cocktail of McDowell and Nussbaum. In particular, the role that is played in human good by a conception of reason that is nevertheless embedded in, and dependent on, particular communities and worldly circumstances in his thesis (see especially 155-62) seems functionally similar to the role of virtue – *qua* responsiveness to the good – in my own. However, some threat of reductionism is still present: it is not clear, as it is with

made explicit by Michael Thompson, most comprehensively in his 2008 book *Life and Action*. Since both Foot and Hursthouse rely on (an earlier, less detailed, version of) Thompson's analysis, and since it is in any case the best account of the basic concepts underlying neo-Aristotelian naturalism available, I'll begin by describing this view, before saying more about how Foot and Hursthouse make theoretical use of it.

Thompson's central claim is that there is no way of characterising a large number of concepts relating to life and the living without involving an Aristotelian concept of *life-form*. Attempts to do so in straightforwardly scientific terms are never satisfactory: one might attempt to do so using a range of empirical concepts – 'homeostasis', 'reproduction', 'organisation', 'growth and development', 'adaption', but when the details of such concepts themselves begin to be filled in, it emerges that they implicitly rely on an idea of *typicality* or *normality* that makes the analysis largely circular. For example, 'organisation' requires some concept of living functioning to avoid the implication that newly-dead corpses are alive (2008, 36), and 'growth and development' similarly cannot be reduced to mere thermodynamic *change* – it is only when changes of some specific kind takes place that they could count as *growth*, or *development*, and this specific kind seems to be precisely that which characterises the species in question (43). Thompson insists that this does not amount to the claim that such concepts

---

McDowell, that the good that ethical inquiry aims toward and partially constitutes is itself ethical. This is suggested, at least, by MacIntyre's tacit acceptance of the idea that the individual good excludes the goods of others, at least in principle. If he did not accept this, there would be no need to argue (as he does e.g. at 105-110) for the importance of a given (virtuous) action as something that serves the individual's good, *independently* of its rightness – if the human good is thoroughly ethical, the presence of ethical value is reason enough for any human, regardless of whether or not they might benefit from a non-ethical point of view.

It is also important to note that I do not oppose MacIntyre's view on other (supposedly) McDowellian grounds, in particular the ground that concepts like 'belief', 'desire', 'reason', etc. cannot possibly find application in non-human animal life *at all*, because of their essentially linguistic nature. If true this would dramatically limit the scope of the analogy between human and animal flourishing that MacIntyre proposes. McDowell has moderated his position on these issues (2008b, 234-8), and even if he had not, I would not need to go along with them in the relevant details – they are not essential for the metaethical stance I have endorsed. See Lovibond 2008, for an overview of the issue.

are *entirely* circular, or involve 'egregious organicist metaphysics', but rather that there is a certain logically distinctive kind of *thought* that all propositions about the organic involve, which is encapsulated in concepts of life-form (48). To think, or say, of something that it is alive is to think of it as instantiating a life-form. Life-form concepts, in turn, are given content by the various functionings that exemplify the beings that they place together – a conception of normal developmental paths, of typical behavioural responses to different stimuli, of things that the exemplary life-form-bearer *does* and *is*. 'Rabbits eat grass', 'Wolves communicate using howls', 'Spiders have eight legs', 'Swans pair-bond for life', etc.

There are a few things that are important to note about these concepts, on Thompson's development. Two of these follow from the fact that they are not names for statistical regularities among phenomena, or collections thereof. We do not produce life-form concepts by isolating particular functionings, in abstraction from the beings that manifest them<sup>17</sup>, and then inventing a life-form concept for each point at which similar functionings frequently coincide, like the points of overlap on a functioning Venn diagram. Rather, we begin with a version of a life-form concept, and alter it only when it makes sense to do so according to the whole concept; new information can make a difference only when it would cohere with the existing conception. The first consequence is that it may be the case that a functioning is characteristic of a life-form even if only a minority of the bearers of that life-form actually manifest it. The classic<sup>18</sup> example is 'having 32 teeth', a functioning apparently characteristic of human animals, but which is not instantiated by many or most humans at any given time – since some teeth will never have developed, others may have fallen out, etc. (68). The second

---

17 It is questionable whether we can even think of something as *being* a functioning, unless a life-form concept is involved somehow. Beings and doings are literally the stuff of life. So, arguably, this abstraction couldn't even make sense.

18 The origin is Anscombe 1958, 14. Compare (McDowell 1998m, 171), which figures in the argument of the next section.

is that the epistemology of life-form concepts cannot be straightforwardly empirical or scientific – merely noting that lots of animals of kind *x* instantiate a functioning *f* does not (of itself) give a reason to include the functioning in the life-form concept *x* (68-73). In the end, the space of life-form concepts is *sui generis* relative to other concepts, including the concepts of lawlike empirical patterns that are typical of the sciences.

The most important feature of life-form concepts for my purposes, however, is the fact that they produce *standards*, relative to which organisms can be evaluated. The simplest explanation for why this might be so would turn on a claim that life-form concepts *are* norms themselves. Thompson resists this conclusion, saying that it is unilluminating, at best, to analyse life-form concepts, or their elements, as norms (73-6), but I cannot go into this in detail<sup>19</sup>. In any case, it is relatively unimportant, because even if life-form concepts aren't normative *themselves*, they do *imply* norms:

If, though, we want to apply 'normative' categories to sub-rational nature, and apart from any relationship to 'our interests', then the question inevitably arises...: Where does the standard come from?...The system of natural-historical propositions [predications of a life-form] with a given [life-form]...as subject supplies such a standard for members of that kind. We may implicitly define a certain very abstract category of 'natural defect' with the following...principle of inference: *from*: "The S is F", *and*: "This S is not F," *to infer*: "This S is defective in that it is not F." (80).

In this way, the contents of the life-form concept *human* have a bearing on evaluations of individual humans, in the light of functionings – behaviours, relationships, character-traits, etc. – that they do or do not possess.<sup>20</sup>

---

19 It ultimately appears to come down to his claim that life-form judgements are 'transparently "factual" or "positive"', and analysing life-form judgements as normative does 'violence' to that (2008, 75). This seems like a very thin basis for such a significant claim, especially since if life-form concepts were normative themselves this would make it clear why they have normative consequences, rather than merely entailing norms for some reason that remains obscure. The dichotomy between the normative and the 'factual' or 'positive' that Thompson relies on here is obviously questionable, and a plausible metaethical realism like McDowell's should dissolve it. But I shall not insist on treating life-form as inherently normative, for simplicity's sake.

20 Thompson provides little more in the way of development of a grounding for ethics, although he plainly endorses the spirit, if not the letter, of Foot's project in *Natural Goodness* (2008, 81-2).

### §2.1.2 Foot's Gloss, and Hursthouse's

This sort of account might provide a grounding for ethical evaluations on its own – if the functioning 'being virtuous' were an element in the human life-form, for example, this would amount to such a grounding. The two most significant versions of neo-Aristotelian naturalism to date – of Foot and of Hursthouse – however, have pursued a grounding less directly, by providing a gloss on a general idea of 'natural goodness', and so influencing the shape of the human life-form concept. In Thompson's basic schema there is no explicit general concept of 'natural goodness' – there is only failure or success at exemplifying the particular life-form of one's kind. But Foot insists that such a concept is necessary, *apart* from the specific norms implied by life-forms, because it is not clear why *all* the elements of life-form should have significant normative weight<sup>21</sup>. The examples she gives are 'The blue-tit has a round blue patch on its head', and 'The male peacock has a brightly coloured tail', which, she thinks, are both genuine elements of their respective life-forms<sup>22</sup>. However, only the latter could reasonably be thought to give rise to evaluations of good or bad in organisms that manifest it, or fail to. The peacock's tail is essential to attract a mate. But (seemingly), the colour of the blue-tit's head 'plays no part in [its] life...[But] what counts as “its life” in this context? And what is “playing a part”?' (2001, 30).

Foot offers an answer to this question, and it is through this answer that the most distinctive –

---

21 Brewer 2009 also supports the addition of a 'teleological' requirement on evaluation-relevant elements of the life-form, for similar reasons (200-202). He claims that, without such an addition, it will not be clear why not living up to one's life-form is a bad thing. As with Foot's reasoning, this argument does not apply where the life-form is itself a subject of ethical evaluation, as it is in my proposal below.

22 There is room for scepticism about this: perhaps, if the round blue patch is the only thing that distinguishes blue-tits from otherwise identical organisms, there *is* no distinctively 'blue-tittish' life-form, and so no bird could be judged defective for not instantiating that functioning. This thought would depend, however, on there being a system of evaluations – i.e. ethics – *already at work* in deciding what counts as a life-form and what does not, and the approach Foot is taking seems to preclude that.

and, as we shall see, most problematic – part of her approach comes into view:

In plants and non-human animals<sup>23</sup> [evaluation-relevant elements of the life-form] all have to do, directly or indirectly, with self-maintenance...or with the reproduction of the individual... This is 'the life' characteristic of the kind of animal with which the [relevant functionings] here have to do. What 'plays a part' in this life is that which is causally and teleologically related to it, as putting out roots is related to obtaining nourishment, and attracting insects is related to reproduction in plants... The way an individual *should be* is determined by what is needed for development, self-maintenance, and reproduction. (31-3).

So, an element of the life-form is relevant to evaluation iff it contributes to – by being a means to – the continuation of the life of the individual and its species<sup>24</sup>. It is not entirely clear how this restriction works – after all, if having a blue patch on one's head is *genuinely* what typifies the life of one's species, *any* continuation of that life that does not involve having a blue-patched head will not be species-typical, and so will not be *good*. So I am skeptical whether this 'teleological' restriction on evaluation can be made to work. We can either take seriously the idea of life-form-typical functioning, in which case *all* elements of the life-form are teleologically related to its continuation, since without them whatever 'life' would continue would not be exemplary of the kind. Or we can invest in this notion of 'life' that does not presuppose the elements of the life-form – or, at least, not all of them – in which case we have abandoned Thompson's conceptual framework entirely. It is not clear what we might replace it with while staying true to the basic Aristotelian thought: that something's species determines what is good for it. This is a serious problem for Foot, but I'll leave it aside, both because my purpose here is not to criticise all other versions of neo-Aristotelian naturalism, and because there is a more general problem looming, a species of (what I shall call) *anti-reductionist* challenge, which I shall introduce in the next sub-section, and return to consider Foot's vulnerability to it after that.

---

23 As I shall note in §2.3.1, Foot endorses a (rather *ad hoc* seeming) exemption from this gloss on natural goodness for human beings. This is not of significance for the moment.

24 This opens the way to an account of the *human* good, and a grounding of virtue in that account, the details of which are not important for my purposes at this point.

Before I move on to the critiques, however, I need to describe the form that Hursthouse's grounding of ethical evaluation takes, since it differs somewhat from that of Foot. Similarly, it begins with a gloss on the ends with which natural goodness is concerned:

[A] good social animal<sup>25</sup>...is one that is well fitted or endowed with respect to (i) its parts, (ii) its operations, (iii) its actions, (iv) its desires and emotions; whether it is thus fitted or endowed is determined by whether these four aspects well serve (1) its individual survival, (2) the continuance of its species, (3) its characteristic freedom from pain and characteristic enjoyment, and (4) the good functioning of its social group – in the ways characteristic of the species. (1999, 202).

Unlike Foot's account, however, this account of natural ends is not presented as restricting the elements of a life-form concept that are relevant to evaluation. Rather, these ends seem to be a gloss on the natural good of *all* living things, directly. Thompson's life-form concepts – which Hursthouse invokes on the next page (203) – must then play a more minimal role, perhaps fleshing out the references to 'species' and 'characteristic' functionings. Moving on to evaluations of human individuals, Hursthouse insists that these four ends are all still apt to provide an account of the human good (206-16)<sup>26</sup>, although she claims that rather than the four aspects that are to be evaluated above, humans have only one – character traits<sup>27</sup>. Such character traits can then be assessed according to whether they serve the four ends, conceived of in such a way as to reflect humanity's distinctive 'ways of going on' (238). Doing so, Hursthouse thinks, will play a role in justifying the development of some character traits – as virtues – over others.

---

25 Non-social species are supposed to have a simpler good, involving no reference to group functioning.

26 Notably, however, she argues that each must be reinterpreted in the light of human rational capacities, which is arguably most characteristic of humanity (217-26). This does not go as far as is necessary towards making the natural good congruent with plausible ethical conceptions however – why should even modified versions of these ends be considered of major intrinsic worth?

27 Her reasoning for this is apparently just that we would not usually call someone a 'good human being' for being e.g. *healthy*. (206-7). This seems a little superficial, but in context it makes sense: Hursthouse, unlike myself (and Nussbaum, and, probably, Foot), seems not to be attempting to produce an account of 'good human' *qua* 'human who has a good life' or *qua* 'good example of humanity', but just *qua* 'admirable human', or something similar.

## §2.2 Anti-Reductionist Critique

I have just described two versions of neo-Aristotelian naturalist grounding for ethics (or, at least, parts thereof). Each, as well as (potentially) the basic neo-Aristotelian concept of *life-form* on which they rely, faces a number of serious challenges<sup>28</sup>. Many of these take a general anti-reductionist form, and it is one of these that I'll focus on here, as an illustrative example. Ultimately, I shall suggest that most such problems<sup>29</sup> have the same root: it is problematic (to put it mildly) to use a conception of natural goodness individuated apart from ethical concerns in building to ethical conclusions.

The strand of anti-reductionist argument that I'll focus on centres on the observation that reductionist forms of naturalism are apt to rely on a sort of *fatalism*<sup>30</sup>. The fatalistic reasoning might look like this: humans are animals, with animal natures; it follows that there are ways that they are biologically determined to behave; and so A) since 'ought' implies 'can' (and 'cannot' thus defeats 'ought to'), no-one can say that humans ought to do anything against this nature, and/or B) if Aristotelian teleology is true, and things ought to behave according to their kind, humans positively *ought* to act in accordance with this nature<sup>31</sup>. Notably, this is a

---

28 The most important challenge that I'll *not* address at length is an empirical one. Several commentators have suggested that the best contemporary biological science provides no room for the concepts of natural, species-specific essences, on which the neo-Aristotelians draw. Important examples include FitzPatrick (2000), and Lewens (2010; 2011). It should be clear, however, that none of the thinkers I discuss here rely on such biological science – at least not in any straightforward way. The absence of teleology (it is even less obvious that biology is empirically *incompatible* with teleology) from contemporary biology does not constitute a major threat. Of course, if someone accepts a *general* 'naturalist' view, on which philosophy cannot substantially go beyond the natural sciences, this will be unconvincing. Then, every philosophical approach would have to closely relate itself to the natural sciences. But this is not of special significance here, since the general McDowellian metaethics I have endorsed would already be untenable were this metaphilosophical stance warranted. Cf. the discussion of metaphilosophical naturalism in Appendix B.

29 Cf. for example, the different argument of Andreou 2006, which argues that it is likely to be unclear that (*prima facie*) immoral dispositions are not untypical of humanity, without mistakenly attributing a statistical epistemology to the Aristotelian naturalists. Her argument could be avoided if ethical concepts were allowed to do explicit work.

30 Some of the examples adduced in Antony 2000 are especially powerful.

31 Newman 2014, for example, provides an appalling example of this kind of reasoning in the context of sexual ethics. He argues (validly) from two Aristotelian naturalist (or very similar) premises to the impermissibility of homosexuality, masturbation, and *in vitro* fertilisation. This abhorrent fatalist attitude is,

very strange and (*prima facie*) implausible thesis as stated. Firstly, it is not clear why anyone would expect it to succeed as an argument – implicitly, anyone who is proposing that humanity should live one way rather than another almost certainly believes that they *can* live that way, and probably has evidence to back this up, in the form of examples of people who *have* lived that way. Secondly, the B) conclusion relies on a very strange form of Aristotelian teleology indeed. One of the ways that it is strange is exegetical: the notion of determination seems to ignore the distinctions between *final* and *efficient* causes, etc., which Aristotelian naturalists use; good action, for typical Aristotelians, is action with the fulfilment of our nature as its *final* cause – it is *the end for which it is done* – but this is perfectly consistent with our nature not being in any sense its *efficient* cause. It is clear how the idea of *final* causation is connected to justification, because both causation and justification are similarly teleological, and both can have the same end.<sup>32</sup> But why would something's *efficient* cause play a role in justifying it, in the way that B) presumes?

It may be helpful to examine an example of Hursthouse's reasoning in this context. The following passage may seem fatalistic, in the way I have described<sup>33</sup>:

Could impersonal benevolence, as a character trait of human beings, foster these...ends? The question is, admittedly, wildly speculative, but on the face of it, it rather looks as though the species and familial bonding that are part of our biological, animal nature, and make us 'partial' to our own species and children, play an essential role in sustaining these two ends. (225).

One element here, at least, is the thought that partiality might be a *necessary* part of human life; that human life *could not be any other way*. This may appear problematically fatalistic.

---

thus, very much a live concern.

32 This distinction also appears to play a significant, if implicit, role in Foot's comments at (30, fn.1). There she distinguishes the notion of 'function' from that of Ruth Millikan (among others) – Millikan's is a 'historical' conception, which I take to be a reference to the fact that function for Millikan is an etiological concept, referring simply to the properties that have played a causal role in the evolutionary selection of a trait (e.g. Millikan 1993). Cf. McDowell 2009d, esp. 267-75.

33 The context, for clarity, is a discussion of Singer's view that benevolence should be directed equally towards human and non-human animals, on pain of 'speciesism'.

However, despite the appearances, it is not so. This is already visible, in the references to 'sustaining ends', rather than living in accordance with nature. But it is immediately clarified further:

This is *not* to lapse back into resting content with our nature as we find it, *not* to deny that we could reshape our nature in such a way that we no longer cared particularly about our own species or children. Maybe we could. It is to draw attention to the fact that, with respect to the continuance of the species and to the good functioning of the social group [two ethical ends, for Hursthouse], our natural tendency to bond to other human beings and to our children seems to be serving us rather well. (225-6).

This illustrates nicely both the threat of reductionistic fatalism, and one means by which it can be avoided.

It may remain obscure how the argument of the previous two paragraphs relates to anti-reductionism. To clarify, let me note that the sense in which fatalism is *false* is also the sense in which ethical judgement can be effective<sup>34</sup>. Fatalism is repellent because we know that there are a wide range of ways that we<sup>35</sup> could be; a range that includes but outstrips all of the ways that humans have been, or have even been *conceived* to be. Even if it is materially impossible for *homo sapiens* to instantiate some functionings, as it is currently constituted, this is not *necessarily* significant, because we might be able to transform ourselves such that this would not be so. Obviously, we are limited by the physical boundaries of the universe, and what our technology allows us to change within it. And, equally obviously, the fact that we *can* change ourselves to enable us to do something does not establish that we *should*.<sup>36</sup>

---

34 I take it that this thought is a major part of McDowell's argument concerning the 'rational wolf' example in his 'Two Sorts of Naturalism' (1998m, 171-3), especially this: 'Having acquired reason, [the rational wolf] can step back from the natural impulse and direct critical scrutiny at it.' (171).

35 Note that this 'we' can be defined first-personally using direct self-reference – 'we' refers to whatever life-form it is that I am a member of. In principle, this is then compatible with the thought that, given an intransigent and ethically worthless conception of humanity, I can form the judgement that I am not human (and neither are any of the other individuals who I regard as significantly similar to me). See Thompson 2004a.

36 Poetically, our limits are something like the starry heavens above, and the moral law within, to paraphrase Kant inappropriately (Guyer 1992, 1). The boundary of human possibility and the boundary of technological possibility (itself necessarily bounded by physical law) only come apart if ethics demands it (as, in many cases, it probably will).

But humanity could remake itself, and arbitrarily restricting our form (as opposed to restricting it on an ethical basis) to that which we currently have is unjustifiable.

There is room to generate a wide, and apparently disparate array of anti-reductionist arguments, focussing on quite different aspects of naturalist accounts. There is something that unites them, however. The uniting thought is that it just isn't plausible that an ethics-neutral conception of natural goodness could ground ethical claims. In turn this is because the *sui generis* nature of ethical reasoning ensures that any non-ethically-specified feature of the world is of questionable ethical relevance. If we attempt to look at human life as such *sub specie aeternatis*, from outside any ethical perspective, we may well find that there is nothing there. And if we begin from a notion of human nature that *constrains* us – rather than being a precondition of everything we could (even conceivably) do – we will neglect the ethically-evaluable possibilities that we *do* have. If all this is correct, the project of neo-Aristotelian naturalism, if it is reductive – if it aims to ground ethics in something with which it is significantly discontinuous – must fail. As McDowell says:

Reason does not just open our eyes to our nature, as members of the animal species we belong to; it also enables and even obliges us to step back from it, in a way that puts its bearing on our practical problems into question. (1998m, 172).

It is only within the pre-existing context of ethical reasoning that a conception of human nature can be established as relevant to how we should live; it is only if that conception is itself the product of ethical judgement that this relevance can be guaranteed.

### §2.3 Difficulties of Interpretation

So far, I have introduced the basic neo-Aristotelian naturalist framework, and I've given some reasons to be suspicious of its cogency, at least given some ways in which it might be

developed. However, I need to do more to concretely link these wrong turns to existing neo-Aristotelian approaches (or, alternatively, to absolve them of such mistakes). In this subsection, then, I move through each of the three main thinkers in turn, assessing their vulnerability to these problems, and clarifying, in the course of this discussion, the issues at stake.

### §2.3.1 Foot

Foot's teleological requirement for elements of a life-form to have evaluative relevance, as noted in §2.1.2 above, informs her ultimate account of natural goodness to a significant extent. However, it is dubious why the three ends she identifies – 'development', 'self-maintenance' and 'reproduction' – should be thought to have ethical relevance for humans. In fact, Foot appears to admit that this account of the ends of natural good is not suitable to ground ethics. Unfortunately, this takes the form of an *ad hoc*-seeming exception for human beings. For Foot, '[h]uman good is *sui generis*' (2001, 51). While she thinks that 'lack of capacity to reproduce is a defect in a human being', she also holds that '[t]he bearing and raising of children is not an ultimate good in human life, because other elements of good such as the demands of work to be done may give a man or woman reason to renounce family life.' (42). Although she might attempt to accommodate this within her teleological account by arguing that such work is also for the sake of the three ends, she does not appear to do so<sup>37</sup>.

---

37 Obviously, this violence to the claim that human good stands to human life in just the same way that (non-human) animal good does to animal life (e.g. 2001, 44). Instead, she weakens the claim to being only that 'a common conceptual structure remains' (51). It would fit better with the core aim of neo-Aristotelian naturalism to drop the 'development, self-maintenance, and reproduction' gloss entirely, except if it is to be understood, as I later advocate, as an explication of what is involved in preserving life-forms, rather than something independent to which life-forms are related. But this would conflict with Foot's explicit motivation for providing the gloss in the first place: her dissatisfaction with the thought that *all* elements of the life-form, *per se*, are of normative relevance.

Instead, Foot eventually arrives at the proposal that the central locus for evaluations of human beings is the quality of their rational wills (66-7). She then proceeds (66-80) to give a sketch of an account of various factors that can contribute to the goodness or badness of a person's will. However, she still maintains that there is a sense of human 'good' that is not identical to the 'goodness' which is constituted by the quality of the will (92). And, although there is a connection between the human good and happiness (81-90), and between virtue (which is connected to the quality of the will) and happiness, this connection is left unclear (94-8). In the end, Foot appears to endorse this disunity, in a way that may be of comfort to anti-reductionists: she does not believe that she has provided a method for settling all disputes about substantial moral questions, but that

in a way, nothing is settled, but everything is left as it was. The account...merely gives a framework within which disputes are said to take place, and tries to get rid of some intruding philosophical theories...that tend to trip us up.' (116).

However, the overall unclarity of her perspective remains. Despite this non-reductionist-seeming conclusion, her early claims that the meaning of 'good' is the same in its applications across the animal world, and that life-form concepts need to be combined with teleology to be evaluatively relevant, coupled with her failure to clarify that evaluation of humans is a matter for ethics, justify some continued anti-reductionist suspicion of her proposals.

### §2.3.2 Hursthouse

In one respect at least<sup>38</sup>, Hursthouse's position is clearly less vulnerable to the anti-

---

38 It is unclear whether Hursthouse is more or less vulnerable than Foot with respect to the structure of her grounding. On the one hand, she relies less on Thompson's pure life-form concepts, which I take to provide the best prospect for a truly non-reductive Aristotelian naturalism. On the other, at least her view then only has *one* conception of natural goodness, against which things can be evaluated. If this is a thoroughly ethical conception, open to revision based on ethical judgement, her view stands a reasonable chance of being non-reductive, the concerns I mention immediately below notwithstanding. (As it happens, the content of that single conception of natural good is, ethically speaking, quite implausible: are individual survival and the continuation of the species really of significant intrinsic value? But the reductive/non-reductive distinction is about the *structure* of a grounding, not its specific normative *content*, so this lack of cogency is not

reductionist critique than Foot's. This is because Hursthouse explicitly views natural-goodness evaluation as a part of ethics, and her discussion appears also to endorse a kind of Neurathian feedback between straightforwardly ethical judgements and 'naturalist' ones (1999, 207-8). For example, she notes that there will be a "subtle interplay between the possible validation of a given character trait as a virtue [i.e. in terms of its contribution to the human natural good] and a modification in one's detailed conception of that virtue" (227). These are precisely the sorts of interconnections that one would expect to see in a fully non-reductive naturalism. However, there is at least one element in Hursthouse's treatment that looks problematic.

Both her non-reductionist presentation, and the worrying factor that remains, can be seen in the following passages:

The pretensions of Aristotelian naturalism are not, in any ordinary understanding of the terms, either 'scientific' or 'foundational'. It does not seek to establish its conclusions from 'a neutral point of view'. (193)

This is, plainly, intended to avert the suspicion of reductionism. However, this is supposed to be coherent with the view that:

Neither side [of an imagined disagreement] believes what they believe about how life works on the basis of...observation or statistical analysis. The beliefs are part and parcel of their ethical...outlook, and the...disagreements should surely count as ethical disagreements. But they are far from *obviously* part of an ethical outlook, and far from being obvious candidates for 'evaluative beliefs'.... I suppose we could classify them as 'ethical but non-evaluative beliefs about human nature'... But if we do classify them thus, the parameters of the contemporary debate about the rationality or objectivity of morality shift... the extensions of 'value-laden' and 'from within an ethical outlook' do not coincide. (189-90).

It is not clear what work, exactly, the notion of the ethical-but-not-evaluative is supposed to do here. The thought appears to be that recognising the *prima facie* factuality of judgements employing concepts like 'survival' or 'the continuance of the species' tells against their being

---

relevant for my current concerns.)

value-laden, in the way that other more obviously ethical concepts seem to be. The claim that some only-ethically-accessible facts are nevertheless not 'value-laden' is worrying; it suggests, Hursthouse's pleas of innocence notwithstanding, that there is a significant gulf between concepts like 'virtue' and 'good', and the concepts she wants to ground them in. I take it that 'value-ladenness' signifies, for Hursthouse, and Foot<sup>39</sup>, the involvement of elements of psychology which had – for much of the 20<sup>th</sup> Century – been the near-exclusive domain of non-cognitivists: emotions, pragmatic commitments within social groupings, and so on. The insistence that ideas of natural good are not 'evaluative' thus suggests that they do not involve these emotional and social capacities, and do not carry the 'non-rational' and 'non-objective' taint that those capacities have traditionally<sup>40</sup> been supposed to carry<sup>41</sup>. This is problematic. Firstly, because the implication that not being thus 'evaluative' is an (epistemically) good thing reinforces the (gender-prejudicial and false) idea that emotional responses and social positions are at best epistemically irrelevant, and at worst essentially biasing. Secondly, because it is hard to make the view that natural goodness concepts are non-evaluative consistent with the thesis that they are part of ethics and depend upon virtue for their recognition. Surely, if the top-level ethical concepts are 'evaluative', and those concepts are in justificatory contact – through both being parts of ethics – with these simpler, more 'natural' concepts, the connection between them will have to be content-preserving if it is going to be justificatory. If the 'continuance of the species' is going to justify virtuous action<sup>42</sup>, the

---

39 Foot's discussion in Chapter 1 of *Natural Goodness* – which is aimed squarely against the non-cognitivist, and which she treats as a foundation for the work that follows – would appear to confirm this suspicion in her case.

40 As should now be clear, McDowellian metaethics draws the sting from such contrasts, by insisting that ethical awareness *irreducibly* (or 'undisentanglably') involves things from both sides of the dichotomy. Ethical awareness *just is* worldly engagement which involves a particular, *sui generis* cocktail of human capacities. But if Hursthouse's metaethics was similarly oriented, what could 'non-evaluative' mean, other than 'non-ethical'?

41 Perhaps not coincidentally, these are often gendered concepts, being 'soft' and 'feminine' rather than 'hard' and 'masculine' like the natural sciences. See Code 1991, amongst others, for development of this sort of line of thought.

42 Or, for example, if judgements about virtue are going to play a role in forming a conception of the

concepts 'continuance of the species' and 'virtuous action' will have to share content. And if (as I am presuming, based upon the argument of the previous Chapter) the McDowellian conception of ethical thought as irreducibly cognitive *and* affective/motivational is correct, this will be impossible unless there is 'evaluative' content on both sides.<sup>43</sup>

The problem, somewhat ironically given the important work that neo-Aristotelians have done in *undermining* oppositions between fact and value<sup>44</sup>, is that Hursthouse, at least, seems to be relying on that very dichotomy. If, alternatively, there is *no* strict distinction between fact and value – if propositions with value-laden content can nevertheless be factual – then the motivation for what she says here evaporates. Rather than arguing that facts about health or disease, or natural goodness or defect, are so *obviously* factual (but not 'evaluative') that ethics can be factually-based too, the way is open to use the *prima facie* factuality of these claims as counterexamples to the distinction – they are clearly factual *and evaluative too!* Once the rigid dichotomy is finally laid to rest, there will be no need to stress differences in kind between traditionally ethical judgements (about virtue, about justice) and 'natural' ones (about humanity or health). And then no suspicion of reductionism will arise: ethical judgement will have no in principle boundaries, it will be *evaluative facts* all the way down – from 'virtue' to 'humanity' – and every possible element of an ethical perspective can thus be placed into Neurathian contact with every other.

---

'continuance of the species', *mutatis mutandis*.

43 One way that this might not be so would be if the relationship between the simple, 'natural' concepts – 'the continuance of the species', 'characteristic freedom from pain, and characteristic enjoyment', etc. – and the ethical ones 'the good life', 'virtue', were *merely* instrumental. Presumably, one event can be justified relative to another if it makes that other *more likely* (and the latter is itself of value) no matter how different those events are – one requiring 'evaluative' capacities for its conceptualisation, the other merely requiring 'factual' ones. But, I take it that neo-Aristotelians are after more than merely an instrumental link between the human natural good and the good life, virtue, etc. If not, it would be hard to see how their view differs from (perhaps two-level) consequentialism. At any rate, an instrumental link would not be sufficient to play the required role in a Nussbaum-style capability approach – possessing the capabilities, as an essential element of the good, would no longer be related to our identities as human, except perhaps by accident.

44 Hursthouse (108-120) has, indeed, contributed to this project. See esp. Putnam 1993, Walsh 2003.

### §2.3.3 Thompson

Like the others, Thompson is aware of the threat from anti-reductionism:

The objection I...want to take seriously starts from the thought that in employing such notions as *life* and *organism* and *life-form* or *species* we introduce something *foreign*, in particular something 'biological', or crudely empirical, into the elements of ethical theory. Any such view, one thinks, must involve either a scientific dissolution of the ethical, tending maybe towards an 'evolutionary ethics', or else the covert substitution of an outdated metaphysics for what we know to be empirical. Each path leads to its own absurdities. Together they may be thought to betray a yearning to view our practices 'from outside' or 'from sideways on' in hope perhaps with providing them with a foundation or external grounding... But suppose that [my positive, exegetical project concerning concepts of 'life-form', etc., is successful]. Then the employment of such concepts within ethical theory would merely make articulate something already implicit in *pensée sauvage pratique* – and it might seem that a...neo-Aristotelian is looking at things head on, not sideways on. (2008, 31-2).

Clearly, Thompson means this as a riposte precisely to the McDowellian worries that I have described – 'sideways-on' is classic McDowell. What is more, as I shall show, Thompson's relatively pure concept of life-form, without the potentially reductive glosses that Foot and Hursthouse offer, stands the best chance of supporting a plausible neo-Aristotelian naturalist ethic; perhaps this relative purity is partly the result of greater attention to anti-reductionist concerns. There is also the grain of a dismissal of those concerns here, however, that I should respond to. Thompson, in effect, notes that accusing a thesis of reductionism can be little more than a rhetorical flourish, overlaid upon a more straightforward attack on the plausibility of the specific claims the thesis makes. As he says, *if* a version of neo-Aristotelian naturalism is plausible as an exegesis of our discourse and practice surrounding ideas like 'good', 'virtue', etc., it will hardly be an objection to say that it is reducing ethics to something else. In that case, the naturalist would be revealing what ethics *is*. This is not an effective rebuttal of McDowellian worries, however (if that is what Thompson intends), because the anti-reductionist evidently *doesn't* think that the exegesis is plausible, and points to features

of ethics in evidence<sup>45</sup>. Accusations of reductionism about ethics are valid just in case the theories they aim at are implausible revisions of ethical thought, as opposed to exegeses of it; and whether or not this is the case here is, of course, precisely the point at issue.

It appears, then, that neo-Aristotelian naturalism's vulnerability to anti-reductionist argumentation is somewhat mixed<sup>46</sup>. Each of the views I have sketched has some degree of defence against such arguments, but each remains vulnerable, not least because of continuing unclarity. It is, therefore, hard to definitively establish whether or not Hursthouse's or Foot's versions of naturalism are problematically reductive. It seems possible, for all I have said, that what Foot and Hursthouse have offered is not a theory that tries to shortcut ethical judgement at all, but merely an (occasionally quite implausible) ethical account of the good life<sup>47</sup>. This possibility notwithstanding, the source of the problems that threaten them is clear: both Hursthouse and Foot seem to seek a basis for ethical judgement that lies, to some extent, beyond ethics itself, and conceive of natural goodness within that frame. The direction to turn, in looking for a neo-Aristotelian theory that will be immune to such worries, is then equally clear: the life-form, and the judgements of natural goodness that it supports, must be located within the sphere of ethical thought. Explaining how this can be done effectively will

---

45 Perhaps the reductionist could admit that what they offer is not a plausible exegesis of *ethical* thought, or anything connected with it, but insist that there is some entirely separate kind of practical thought that is implicated by concepts like 'natural defect'. It is sufficient to note that it is not clear what this alternative could be. Firstly, such a system of evaluation would inevitably have to *compete* with ethics, if it did not aspire to ground it – sometimes the evaluations would diverge – and that looks like a competition it couldn't fail to lose. Secondly, people (or most of them anyway) already engage in and care about ethics; I see no reason to think that anybody cares about this *other* sort of evaluation. Of course, people do sometimes engage in Aristotelian-sounding evaluative talk about humanity – some behaviours are described as 'inhuman', others as 'unnatural'. But I do not know how to hear such things except as engagements in ethics. If someone were genuinely to evaluate me as (e.g.) not being a very good human on the grounds that I have 31 healthy teeth rather than 32, I do not know what to think except that they are doing ethics (badly).

46 John Skorupski has attributed a form of reductionist Aristotelian naturalism to Terrance Irwin (2012; 319-21); if his exegesis is correct, Irwin's version is more clearly reductionistic than Foot's, Hursthouse's, or Thompson's.

47 This would be the case iff their theories can be reconstructed so as to meet the two conditions I give in §2.3.2 below.

occupy the rest of this Chapter.

I'll conclude this section by offering a first attempt – at a way of expressing the link between humanity as a life-form concept and the concept of the good life – that is not reductionistic in the way that Foot and Hursthouse's links threaten to be. Before, in my summary of the neo-Aristotelian naturalist grounding of the capability approach, I said that an individual has a good life to the extent that they adequately instantiate the species norm; a standard that fundamentally expresses what it is to be a member of humanity. It is perhaps unclear, assuming that the argument of this section is correct, and no reductionistic account of how ethical good relates to biological humanity is plausible, why there should be any connection between these concepts at all. On one level, the intuitive plausibility of the basic connection is not entirely relevant, because one of my purposes throughout the Thesis has been to provide an account of humanity which involves a clear link to ethics (and particularly justice). Even if neo-Aristotelian naturalism had no inherent plausibility, this would be a benefit, because it would help to clarify the basis for universalism. There may be limits, beyond this motivation, to what I can say; if people genuinely do not feel the pull of the thought that better exemplifying humanity could be the same sort of thing as living a better (human) life, I may not be able to make that thought attractive. However, I think it *is* deeply attractive – what the good person is is *a good example of humanity*. Once we get clearer on the nature of this standard of humanity and its complex interconnections with ethical thought – which I hope the next section will achieve – the plausibility of this basic link should be clearer, too.

### §3 An Alternative View

I have claimed that neo-Aristotelian naturalism, as developed at least by Foot and Hursthouse, contains a tension between an implicit impulse to reduce ethics to something more primal, and (occasionally) the explicit endorsement of the view that no such reduction is either necessary or possible. In this section, I describe and defend an insistently non-reductive version of neo-Aristotelian naturalism.

### §3.1 The Human Life-Form and the Virtuous Life

#### §3.1.1 Life-Form and the Good Life

As I have said, the problem with many forms of neo-Aristotelian naturalism is the extent to which they attempt to reduce the human good, and the virtuous perspective from which that good is uniquely visible, to a basis that is ill-suited to account for the particular rational character of human life, and that is not clearly of ethical relevance. It is not clear that 'development, self-maintenance, and reproduction' (Foot), or 'individual survival,...the continuance of [the relevant] species,... characteristic freedom from pain and characteristic enjoyment,...and...the good functioning of [the relevant] social group' (Hursthouse) are significant enough to ground virtue, right action, or ethical goodness. Indeed, it is not really clear that such things are intrinsic goods *at all*. This seems to leave neo-Aristotelian naturalism in a very bad position. However, it does not mean we have to abandon all of the connections that recent Aristotelians have sought to defend; nor does it mean that we have to discard all of the *ways* in which they have defended them. We can rehabilitate something of the ambitions of Foot and Hursthouse (*inter alia*), by seeing the human life-form as a thoroughly ethical idea – not something extra-ethical to which ethical concepts might be reduced – and by reconstructing the neo-Aristotelian naturalist story in that light.

Consider Thompson's idea of 'life-form', as described in §2.1.1 above. The core thesis is that life-form concepts supply normative standards, of a kind; they specify a set of functionings that organisms<sup>48</sup> must have if they are to count as exemplifying that form well. Since it is extremely rare for any individual organism to perfectly instantiate such a standard, it is necessary to distinguish between *minimal* and *full* exemplification. Once an organism is recognised as minimally instantiating a life-form F, the standard that F gives applies to it, and it will be evaluable in terms of its distance from full exemplification. As Thompson says:

If someone...asks, "But what does 'what *most of them* do' have to do with what [some individual] does?" the answer will have to be "Not much, really." But if, in the other case, someone asks, "What bearing does 'what *they* do' have on what [that individual] does or is doing?" the answer will have to be "Everything." For, again, every thought of an individual organism as alive is mediated by thought of the life-form it bears. A true judgement of natural defect thus supplies an 'immanent critique' of its subject.<sup>49</sup> (2008, 81) [My italics.]

Being human means being disposed to go on in a certain way – to exhibit certain functionings; a certain *form*. But, given our actual concepts of humanity, it also brings in a broader conception of functioning that most human organisms might never actually instantiate<sup>50</sup>. In thinking of a being as human one brings in this more-rarely-realised standard of humanity as much as the everyday, minimal one.<sup>51</sup>

Notice, however, that there is nothing in this concept of 'life-form' that specifies that all such life-forms have to be characterisable without using *sui generis* concepts, such as those which belong to ethics. Nor is there any suggestion that all life-forms must have much in common –

---

48 Arguably, what it is *to be* an organism is to be evaluable in terms of one life-form concept or another, as I suggested in §2.1.2 above. Compare Thompson 2008, 56-62.

49 Compare Nussbaum 1995, 118-9.

50 Remember that there is no logical barrier to this – life-form concepts are not based upon statistical regularities.

51 It might not be the case that this rare standard *has* to be implicated along with the common, minimal one. Perhaps the concept could work differently, specifying only some very basic and clearly ethically irrelevant characteristics. If it did, this might give rise to suspicion that there is not, actually, anything worldly that we are making contact with in using life-form concepts as evaluative standards. But I take it as sufficient for my purposes to claim that this is *in fact* how life-form concepts, or at least the life-form concept 'human', works. Compare Thompson's contrasts between 'logical', 'local', and 'substantive' 'Footianism' in his 2003.

that they all count as life-forms is a meta-level, *logical*<sup>52</sup> fact about them, and is compatible with a wide range of other differences. Perhaps there could be living beings composed of silicon materials, rather than carbon ones, for example. It is not clear, then, that all life-forms have to be glossed in terms of (e.g.) reproductive advantage, or individual survival, or typical bodily development. Nor is it clear why the life-form must *exclusively* involve things of that 'natural'<sup>53</sup>, seemingly culturally-invariant sort. The reason why such features are exclusively emphasised by Foot and Hursthouse is, perhaps, that they aim to be characterising 'humans', *qua* '*homo sapiens*', and it does not seem very plausible that members of prehistoric *homo sapiens* exhibited much behavioural continuity with contemporary humans (unless cultural specificity is ruled out). But, firstly, another notable feature of life-form concepts is that they *don't* have to match up to scientific concepts of species (Thompson 2008, 53-62), so there is room to claim that pre-historic *homo sapiens* weren't really 'humans', in the ethically relevant sense, at all. And, secondly, even if we need to preserve a close extensional similarity between the concepts *homo sapiens* and 'human', it is not really clear why culturally-specific features couldn't be included. After all, all members of *homo sapiens* probably had the biological *potential* to instantiate many culture-specific modern functionings, if only the right process of enculturation had been present in their environment. And, given life-form-membership is a matter of disposition towards characteristic functioning, evidence of the relevant biological potentials potentially *just is* evidence of life-form-membership<sup>54</sup>.

Once this is established, it is open to us to formulate a definitively non-reductive version of

---

52 Cf. especially Thompson's endorsement of Frege's project of logical analysis (2008, 13-22).

53 McDowell would say 'first natural' here. Compare Appendix B, §4.

54 Compare §1.3 above. Arguably, what I am here referring to as 'biological potential' will include many of the features which form the set of human 'basic capabilities'. On the other hand, perhaps some basic capabilities will not reduce to anything straightforwardly biological, standing for a complex interaction between biological materials and (perhaps cultural) features of the environment.

neo-Aristotelian naturalism, using the same basic analysis of life-form concepts that Thompson has provided, but insisting that it is a matter of *ethical* evaluation what 'humans' are, and what is good for them as such. Many of the specific normative theses that neo-Aristotelian naturalists emphasise, like the importance of human sociality and habituation, or the ubiquity of vulnerability and dependence on others<sup>55</sup>, can still be (and, plausibly, are) correct, on this interpretation. But they are correct because they are excellent ethical assessments of human lives, accurately described, and not because they are able to reach down below the level of ethical evaluations of features of life to a supposedly deeper one. Rather than thinking of such evaluations as registering that certain ways of living rather than others serve the continued existence of particular *homo sapiens* bodies and their statistically-regular organic functionings (which they often will, irrelevant as this may be) we can think of them as directly serving the continued existence of good human lives.

I can now reconsider the language of Foot and Hursthouse, and evaluate whether or not it has continuing relevance. I shall show that it does, given certain revisions. For we can reinterpret talk of (e.g.) 'development, self-maintenance, and reproduction', not as processes that occur (or fail to occur) in organisms – conceived of independently of any ethical frame – but as processes that operate within the human *life-form* itself<sup>56</sup>. That is, what is preserved by the processes is the *formed* life, not merely life *simpliciter*. In the human case, what is preserved is the good life (ethically considered), where this is a life of virtue, and where it is the continuous inculcation and practice of virtue itself that does the preserving<sup>57</sup>. This

---

55 See MacIntyre 1999, which provides a detailed exposition of various ways in which Aristotelian thinking can help us to recognise the significance of our animal fragility, including helpful links to similar work by feminists such as Eva Feder Kittay and Virginia Held (e.g. 3).

56 If so reinterpreted, there would be far less significant conflict between Foot's thesis and my proposal at this point. Despite e.g. what Thompson says in his 2003, which would support this *rapprochement*, I still have my doubts as to whether this would be a fair reading of Foot, but I cannot go into further detail here.

57 As should be clear from Ch.4 (esp. §§1.3-4), this is tautological within the McDowellian framework: virtue

preservation of value will be of two kinds. Firstly, the practice of virtue is itself a major part of the human good, and so working to make people more virtuous (most straightforwardly by becoming more virtuous oneself) will constitutively realise the goods the virtuous life involves. There are several reasons to make this claim. A) it is essentially the same as Foot's emphasis on the distinctive centrality of practical rationality – the 'good will' – to human life, and preserves the cogency of that point<sup>58</sup>. B) it seems to be necessary to accommodate the *epistemic* significance of virtue, as the disposition to see things aright. Plausibly, having this sort of knowledge is of value even apart from the good actions that it causes – especially when the affective components of such knowledge are kept in mind. And finally, C) the life of virtue has many other elements whose intrinsic value seems clear, *prima facie*. To give some specific examples, the virtuous would exhibit proper dispositions of care for others, and would respond with appreciation to the goodness of others and of external objects, which cannot fail to make their lives better – these are valuable functionings if any are. Secondly, virtuous dispositions will have an instrumental link to goods that they cannot realise alone<sup>59</sup>. Although it might not be possible to guarantee the existence of valuable friendships, for example, by individual virtue alone, individual virtue will make them more likely to come about. Looking at it this way makes it easy to see why the telos of the processes of virtue-creation – of its 'development, self-maintenance, and reproduction' – can ground ethical

---

is a name for the disposition to judge and act rightly; in accordance with the good. The virtuous then cannot fail to preserve the good, if anything can. As I also stress later (arguably, *contra* McDowell) it is not plausible that virtue is sufficient for the good life, and so there is no *guarantee* that the successful practice of virtue will lead to the preservation of the good over time. But it is the only thing that could be expected to do that. (Consider: a supernova could wipe out all life on Earth tomorrow, even if everyone was perfectly virtuous. This would bring the human good to an abrupt end. But it would not follow that someone must have been unvirtuous despite appearances. There are no guarantees.)

58 See §2.3.1 above.

59 This follows, again, from the McDowellian conception of virtue with which I am working. The virtuous will act, in response to existing value, or the prospect of greater value, in such a way as to contribute to its preservation and further realisation. The virtuous are able to promote the good so effectively because they are excellent *epistemic* agents where it is concerned – they see things aright, where ethics is concerned; and they are disposed to act in accordance with their normative judgements.

value<sup>60</sup>; for what these processes aim at *just is* ethical value: the particular shaping of *homo sapiens* materials that is the good human life.

### §3.1.2 Cultural and non-Cultural Formation

The claims of the last paragraph may remain unclear, or seem wildly revisionary of neo-Aristotelian orthodoxy, or both. To try to allay these worries, I shall offer a diagnosis of the difference between my view of life-form and the Foot-Hursthouse view I reject. In summary, a thoroughly non-reductive naturalism must meet two conditions. Firstly, the contents of the human life-form concept must be a matter of ethical evaluation – evaluation, that is, from the perspective of virtue. Secondly (as a consequence), theorists must allow *cultural* processes of formation into the content of the human life-form; in particular, they must include precisely those cultural processes that result in virtuous agency<sup>61</sup>. If virtue is of constitutive value, and through its link to the general promotion of the good is of overwhelming instrumental value too, virtue should be our focus. The only way to know what the good life – the life of true humanity – looks like is to cultivate virtue; the only possibly reliable way to *live* that life is also to cultivate virtue<sup>62</sup>. Why then would we pretend that virtue itself is justified only by some extrinsic, prior good that it has no standing to critique?

Reductive views, it now appears, are distinguished by their reticence to endorse this sort of

---

60 It is a good question whether or not it really makes sense to say that what is going on here is 'grounding'. Ultimately, what I am saying here is that ethical value grounds itself (we have a diverse family of normative concepts, but they are all intertwined). Cf. McDowell 2009e. I discuss these issues somewhat further in Appendix B.

61 There are likely to be other cultural conditions on a good life, apart from those that enable virtuous agency, and these must be included in the life-form too, but it is especially clear that the human good depends on virtue, whatever else it may depend on; in any case, the virtuous would promote these virtue-external goods insofar as they can be promoted *anyway*.

62 Of *course* the virtuous will endorse the processes that lead to their becoming so – whatever quality they have been able to add to their lives will have depended on those processes. This is insubstantial, and of itself provides no practical guidance. Only an actual, first-order image of virtue can do that.

theory, most likely because their 'naturalist' aspirations lead them to view it as a trivialising evasion; not really a substantive theoretical contribution at all. This suggests two potential problems for the view I have offered. Firstly, it threatens its putatively 'naturalistic' status. Secondly, it implies that it is a total break from the tradition it claims to preserve. I shall say little about the first problem here. However, it may help to note briefly that two of the ways that 'natural' is used, in ordinary contexts and in philosophical discussion, *would* preclude the anti-reductive theory I have offered from counting as naturalistic. The first is 'natural' as opposed to 'cultural' or 'learned'; this is the contrast implied by the (thankfully now infamous) dichotomy between 'nature' and 'nurture'<sup>63</sup>. If a 'naturalism' were a theory that does not invoke cultural elements by definition, then of course this non-reductive form of Aristotelianism will not be naturalistic. But it is not clear why such naturalism would be desirable: enculturation is a worldly process, and learned behaviours are no less real than innate ones. The second is a usage idiosyncratic to recent metaethics. According to this, naturalistic views are just those that do not treat the ethical as *sui generis*, attempting to provide accounts of ethical concepts (or properties, psychological states, methodologies, etc.) in non-ethical terms.<sup>64</sup> Given that my non-reductive version of neo-Aristotelian naturalism precisely *does* treat the ethical as *sui generis*, following McDowell, it cannot be naturalistic in this sense either. Despite this, I think there are significant senses in which this view *is* naturalistic, and this is enough to retain the description, but I cannot defend this claim fully here<sup>65</sup>. Ultimately, the nomenclature for the approach is of no deep importance.

---

63 Cf. Mill 1904, which offers a similar distinction. On the latter sort of conception, it is obvious that ethics is non-natural, because ethical practices are clearly (if to a disputable extent) an artefact of human social construction. Disputes about 'naturalism', so construed, would be of little philosophical interest.

64 See Ridge 2014 for an overview of the vast number of ways these distinctions can be made. Not all naturalistic theories (so thought of) are naturalistic through-and-through: 'non-reductive naturalism', such as the Cornell Realism of Sturgeon and Boyd, is against *sui-generality* at the level of methodology, but content with it at the level of concepts.

65 Appendix B supplies the beginnings of a defence.

The second problem is a worry that in constructing my view I have done great violence to the tradition to which it claims to belong. It is easy to see why this might be the case: if the point of neo-Aristotelian naturalism was to relate ethics to the notion of *life, simpliciter*, I will appear to be entirely outside the tradition. But this issue is not as clear-cut as it seems to be. Against this implication, it is significant that Aristotelians like Foot are *already* not talking about the preservation of life *just simpliciter* – for it is specifically the life that belongs to an individual's *kind* that is supposed to be developed, preserved and reproduced by that individual. There is, in fact, no notion of *life simpliciter* available within the neo-Aristotelian framework, at least as formulated by Thompson; to be alive is just to bear a life-form, and so anything that is alive will bear some specific life-form or other. For contrast, consider a gene-selectionist variant of ethical naturalism, according to which individual organisms are vehicles for genetic material, and individual success is a mere proxy for the success of that material<sup>66</sup>. On such a model, there would be something characteristic of life (the presence of genetic material) that would be more basic than species-identity (because species-identity is a property of organisms, not individual genes). Then, there would be something *apart* from any species-specific life that might (or might not) be continued by the success or failure of the individual. It would then make sense to say, on this view, that the preservation of life (*qua* genetic material) was served or vitiated by the success or failure of individual organisms. But, clearly, this is very far from the view that Foot *et al.* propose<sup>67</sup>. Once we see this, it should be easier to place culturally *formed* life at the heart of neo-Aristotelian naturalism.

---

66 What I have in mind is a normative transformation of a caricature of Richard Dawkins' thesis in *The Selfish Gene* (1976) and later works. I do not mean to suggest that anyone really holds a view as implausible as this.

67 Cf. Lott 2012a, esp. (422), here. Lott 2012b provides an interesting attempt to deal with recent objections to neo-Aristotelian naturalism (especially those of FitzPatrick (2000) while preserving the independence of life-form judgement from ethics. I believe that the cost of this is an even greater than usual obscurity of the grounds to view life-form judgements as ethically relevant, but I cannot defend this here.

### §3.2 The Right and the Good

I have endorsed the view that the image of the life of the virtuous agent, and the image of the exemplary human life, are closely related concepts. In this subsection I explicate further the proposal that the life of the virtuous agent and the good life are connected. Specifically, I shall claim that the virtues are necessary (but not sufficient) ingredients of the good life. Firstly, I'll explain this proposition, and its place in the tradition, in a little more detail. Then I'll address the question of what consequences this has for the abstract relationship between the right and the good.

Given what I have said so far, the availability of a strong claim about the relationship between the right and the good should be clear. After all, I began by explaining that the human good is characterised (for neo-Aristotelian naturalists) in terms of life-form evaluation. In turn, life-form evaluation, when applied to humankind, picks out the quality of rational will as centrally important. When this is then glossed in terms of the disposition to see things aright and act accordingly, the circle is complete. The connection between the right and the good here is twofold. Firstly, it is (constitutively) good to act rightly, in virtue of the centrality of the good will. Secondly, right acts will be just those that are properly responsive to the values that things have<sup>68</sup>. Human goodness requires right action; right action is responsiveness to the good.

---

<sup>68</sup> In Chapter 6, I introduce a complication in this schema, with reference to the virtue of justice, specifically. The upshot is that we are likely to need two notions of right action; one which connects to D-reasons, and the other to V-reasons (see Ch.4 §2.2, and especially Ch.6 fn.72). Only the conception of right action which connects to V-reasons will connect to virtue, as something both necessary and (partially) sufficient for the human good. To clarify, when I refer to 'rational will' and 'the good will' above, I am still not talking about a faculty whose deliverances can be individuated apart from consideration of what they are connected to in the external world: the rational will which connects to virtue (and is a constitutive part of the good) is a will *which responds to the reasons that there are*.

That said, it is not yet entirely clear either how strong these connections are, nor which elements of the tradition I am being faithful to in asserting them. As to the first question, there are four options. Firstly, virtue (and the right actions it enables) could be necessary *and* sufficient for the good life. Working within the naturalist framework I have described, this would amount to the claim that goodness of the will *exhausts* the human natural good. An organism would then be a good example of the life-form 'human' – and, thus, a 'good human' – to the extent that, and only to the extent that, she possesses a good will. Arguably, Foot comes close to endorsing this position when she emphasises the overwhelming importance of the good will for human life (2001, 66-80)<sup>69</sup>. Secondly, the virtues could be sufficient for a good human life, but not necessary for it. On this view, virtue merely opens up one of a myriad of different equally good ways of living, the rest of which do not require a person to act rightly. Thirdly, the virtues could be necessary but not sufficient for a good life. As I have said, this is the alternative I endorse. Finally, they could be neither necessary nor sufficient. This last seems to be Hursthouse's preferred view. She claims that “[Virtue] is not necessary, since it is generally acknowledged that the wicked may flourish like the green bay tree. And it is not sufficient, because...[a]s soon as it is admitted that exercising virtue on a particular occasion may lead to my life's being cut short, or to its ruin, [that] claim is undercut” (1999, 172). Rather, 'for the most part, by and large' (1999, 185), virtues benefit their possessor.

It is not clear to me, especially if the other McDowellian alterations in the structure of neo-Aristotelian naturalism are made, why virtue should not be considered necessary for the good life. Recall that, on the view I have urged, the connection between virtue and the good is not

<sup>69</sup> She certainly seems to endorse the necessity claim: “virtues play a necessary part in the lives of human beings as do stings in the life of bees” (2001, 35). See fn.70 below, however, for remarks that suggest that she doesn't after all think that virtue is sufficient.

instrumental but (at least partially) constitutive. Virtue *just is* goodness of the will, and goodness of the will *just is* part of what makes a human life good. A life lived with a bad (evaluatively inadequate) will, then, cannot fail to be a bad (evaluatively inadequate) life. On the other hand, there seems to be little reason to hold that virtue, so conceived, is *sufficient* for a perfect – or even a minimally good – life. If someone is impeccably and invariably virtuous, *they* will do no wrong, and so a major element of their good will be fully present. But goodness of the will is not plausibly the whole of the human good, and one individual's being perfectly virtuous cannot guarantee that they will not starve and die, or be in constant, abject misery, if only worldly conditions beyond their control are sufficiently harsh. The world does not mysteriously conspire to reward virtue with extrinsic goods; it is its own reward<sup>70 71</sup>.

Even given this, though, it is not entirely clear what sort of relationship is implicated here. In particular, it may seem that what neo-Aristotelian naturalism proposes is a reduction of the right to the good. Some recent Aristotelians have seemed to endorse such a reduction. Hursthouse, for example, seems to do so in her initial statement of the core premiss of virtue

---

70 Cf. McDowell 1998p; 2009c. Both Foot and Hursthouse, notably, hold that McDowell's account of how virtue intrinsically benefits its possessor is too strong. However, I think both read McDowell wrongly here. Foot claims that McDowell *identifies* the good life with the life of virtue (Foot 2001, 97-8). Hursthouse claims that McDowell cannot make sense of tragic losses that virtue might entail (Hursthouse 1999, 179-87). Both rely heavily on McDowell's claim that "no sacrifice necessitated by the [virtuous life]...can count as a genuine loss" (1998p, 17). On the face of it, this does seem to depend on a tacit thesis that virtue is sufficient for the good life (which I have already dismissed as implausible). However, two things need to be noted. Firstly, McDowell's thesis quoted above seems to depend on his thesis of silencing (18), which I already reject (see fn.50 in Ch.4). With silencing removed, there is no need to say anything as strong as the above. I can concede that *some* features of the nevertheless-virtuous act are tragic losses, which deny the agent the full realisation of their good, without denying that the rightness of their action is of constitutive benefit to them; a benefit that is not excised by the tragic context. Secondly, in any case, McDowell offers a gloss on the species of good that virtue *is* sufficient for that makes his claim much more plausible. "[T]he life [of virtue] is...the most satisfying life possible for its subject, circumstanced at each point as [she] is." (19). Even if tragic circumstances prevent the virtuous person from living a perfect life, their virtue allows them to live *as good a life as they can*. If the controversial claim is only supposed to be true in this restricted sense, it is not clear why Foot or Hursthouse (or myself) should object to it.

71 An obvious connection here is to Nussbaum's work on luck and vulnerability in the good life. See (2001b, 318-72).

ethics: 'An action is right iff it is what a virtuous agent would characteristically (i.e. acting in character) do in the circumstances' (1999, 28). If this statement were trivial, this would be unproblematic (and it would fit well with the McDowellian view I am proposing to treat it precisely *as* trivially true). However, Hursthouse specifically rejects any interpretation of this statement according to which it *is* trivial (30-1). Interpreting it as a substantive claim, it amounts to a reduction of one thing: right action, to another: action that flows from virtue (substantively construed). When the connection between virtue and the human good – whereby states of character are virtues just when they are appropriately oriented towards that good – is considered too, it follows for Hursthouse that the rightness of actions is reducible to their contribution to the human good. Later I shall endorse an argument, suggested by David Copp and David Sobel, to the effect that this reductive stance is implausible. Before that, I want to clarify that not all neo-Aristotelian thinkers think that a reduction is available here. Foot, for example, seems to hold that we can assess kinds of act independently of assessing any states of character that they are (contingently) associated with (e.g. at 2001, 113-5). McDowell too seems to deny this reduction's validity, but for quite different reasons. Recall that the virtuous agent, for McDowell, just *is* the person who sees things aright and acts accordingly. This leaves no room for any distinction between the image of the virtuous agent and the image of right action: the connection between them is (mutually) constitutive, and they are required to cohere. Given the further necessary connection between virtue and the human good that I mentioned above, it will follow that those who act rightly thereby contribute to their good. But this does not follow because one separate thing – right action – is reducible to another – virtue – but because right action and virtue should not be understood independently of one another in the first place.<sup>72</sup>

<sup>72</sup> There are echoes of McDowell's quietistic rejection of the Euthyphronic choice here, as discussed in the last Chapter. Similarly, the Copp-Sobel argument below is reminiscent especially of his 'siblings' metaphor (1998o).

It is fortunate that Hursthouse's reductive stance here isn't mandatory for neo-Aristotelians, because it does not seem particularly compelling. Consider the following:

For Hursthouse...the alternatives we face are morally neutral in themselves and they only gain a moral color by reflecting the moral color of the states of character...that we can turn toward them. This view does not make sense of our admiration for the virtues... We admire the fact that a virtuous person is attuned to and energized by the morally relevant facts that she detects in the world. We do not see her virtue as a searchlight penetrating a world that would otherwise have no moral features at all; we see it as a telescope and source of energy that detects and strives and is admirable for these reasons. (Copp and Sobel 2004, 552).

In line with the McDowellian, quietistic account of metaethics that I advanced in the previous Chapter, I think there is something right about this, but that it does not tell the whole story. Hursthouse's picture here plays the role that the constructivist played there, while Copp and Sobel play the role of the 'robust' realist. Their claim is that Hursthouse cannot account for the sense in which there are norms of right action that do not depend on the subjective character states that individuals just happen to have (however appropriate those states may be in their own terms). On the other hand, it may seem mysterious *why* those norms of right action are *norms for* people if we can't make sense of the idea that acting in accordance with them would be good for them. From any even vaguely Aristotelian point of view, it is hard to see how something can be good for an agent if it has nothing to do with their particular identity<sup>73</sup>. The McDowellian version of neo-Aristotelian naturalism, however, does not have either of these problems. The relation of the virtuous to facts about right action is not one of *creation*, but it is not much like astronomical discovery either. We cannot make sense of virtue (and its admirability) without reflecting on the appropriateness of what the virtuous do, but we cannot

---

<sup>73</sup> Compare Williams on Kantian theory in *Ethics and the Limits of Philosophy* (2011, 60-78). More generally, the attack on the 'peculiar institution' of 'morality', which Williams typically contrasts with a broader, less inhumane 'ethics', also resonates to some extent with my claims here. A properly Neurathian methodology stands a good chance of avoiding the alienating features he identifies so well, without falling into the quagmire of relativism and subjectivism that his own account responds skirts so near to (e.g. 156-173, on the rejection of universalism, and 192-3, on subjectivism about reasons).

make sense of that very appropriateness without reflecting on the quality of a life of virtue<sup>74</sup>. The right cannot be strictly reduced to the good using the concept of virtue, but it cannot reasonably be thought of as entirely independent of it either.

In this Chapter, I have aimed to present a revised version of neo-Aristotelian naturalism which preserves the first-order potential of that view – as a motivation to introduce a broader range of anthropocentric considerations into ethics, and as a way of making sense of the significance of humanity for global justice – without introducing the difficulties that it has often been thought to pose. My conclusion is that, although there are serious problems within the tradition, as I've interpreted it, it can be made compatible, both with the McDowellian metaethical framework which I laid out in the previous Chapter, and with plausible first-order judgements about humanity, the good life, and the role of virtue. In the last Chapter, I build on this conclusion with the ultimate aim of producing a framework for a capability approach to global justice which makes sense as a whole.

---

74 The Neurathian theme here has been an emphasis throughout, but it is worth noting a difference in the way it is useful at this point. Before (Chapter 3, esp. §§2.3-4; Chapter 4, esp. §1.4), Neurathian epistemology's primary significance lay in a feedback loop between first-order normative judgements and judgements about the epistemic significance of different agents' contributions; an 'epistemological naturalism' that allowed us to reason about biases, marginalisation, etc. in a way that cohered with a plausible metaethics. In this Chapter, however, it has been of significance in establishing a feedback loop between various concepts, some clearly first-order norms, some more metaethical in origin. Rather than attempt to isolate one particular concept (justice, say), and proceed to make decisions about its construal and significance independently – perhaps relating it back to other concepts at the end of this process – the framework I propose encourages attempts to find a version of the concept that coheres with others, working on each with the others in mind. Cf. esp. McDowell 1994i, 78-82, and Hursthouse 1999, 153-7 and 165-7, for some different Aristotelian endorsements of the project of, so to speak, 'unifying from within'.

## Chapter 6:

### Humanity as the *Polis*

In this Chapter, I return to the focal topic of the Thesis, focussing on a clarification of the connection between the individual-focused Aristotelian naturalist account of the previous Chapter, and global justice proper. Throughout the Thesis I have maintained that plausible accounts of justice require plausible accounts of the human good. But so far, the latter notion has received greater attention than the former. I have said that justice can be construed as requiring that every individual instantiate some diverse bundle of functionings, including many capabilities, suitably distributed over the course of their lives. Further, in the last Chapter I claimed that we can and should think of right action as itself of constitutive benefit to the agent, of appropriate emotional response as of constitutive benefit to the responder, etc. This implies that the connection between justice and humanity is (at least) twofold: firstly, the greater presence of justice is of constitutive benefit to all those (all humans) who are implicated in that fact, and secondly, the presence of justice will be of instrumental benefit to all subjects of justice, because justice involves their having the wide range of goods-in-themselves that I have been calling 'dignity'.

In Chapter 5, §1.2, I sketched the neo-Aristotelian view that I believe to be suitable for providing a basis for a capability theory concerning global justice. The core of this view is that every human being possesses (*eo ipso*) a relational connection with every other, and that there are certain minimal norms which hold of all such connections, whatever additional

layers of relationship may exist in particular cases. A sufficientarian version of the capability approach, like that of Nussbaum, is then an excellent candidate to give an account of those norms, since no relationship can flourish if one party to it does not have dignity. The central thought, then, is that there is a strong link between the contribution to the good life of individual human beings of the ways they relate to one another, and the justice or injustice of even large-scale political states of affairs, including (e.g.) the institutions that comprise the global economy.

There are a number of significant barriers to the acceptance of this link, which it will be the primary aim of this Chapter to reduce or avoid. Firstly, it may seem that a basis for justice in the lives of individual beings and their dyadic relationships is problematically individualistic, and will not have the resources to account for many social dynamics which are relevant to justice. Secondly, it may appear that linking justice with the idea of human relationship vitiates the potential benefits of neo-Aristotelian naturalism, in a way suggested by a paper by Candace Vogler (2006). Thirdly, it may seem that linking justice with what is necessary for all people to live a decent human life will be overly demanding, and that this severs a necessary connection between justice and responsibility. Fourthly, it remains to be seen what practical roles I propose that a capability theory can fill, especially given some of the difficulties which my discussion in this Chapter will uncover. Fifthly, more needs to be said to situate this approach relative to existing proposals about global justice, in particular with reference to articulating and defending the idea that each human is related to every other. And finally, apart from the work of diminishing these threats, the original aim of the Thesis – to establish a capability approach more solidly, with special attention to the role of metaethics – needs to be re-appraised, and final conclusions drawn.

The Chapter begins, in §1.1, by recapitulating the central features of the approach I favour, describing some of its benefits, and then arguing that, despite appearances, it need not be vulnerable to at least the most straightforward accusations of individualism. §1.2 then uses some of Michael Thompson's work to give more detail to the central idea of justice as a personal disposition pertaining to relations. §1.3 responds to the second potential barrier above, giving an exegesis of Vogler's 'promulgation problem', which threatens the coherence of a second-natural approach to human form with a neo-Aristotelian approach to justice, and then arguing that it poses no real threat. §2 responds to the first and third barriers noted above, by distinguishing the notion of personal-relational justice that I think primary, from a structural notion that is overwhelmingly useful in non-ideal – radically unjust – contexts like most of those that make up the contemporary world. After drawing on Iris Marion Young's work to elucidate the notion of structural injustice and the differences it makes, I present an interpretation of it that coheres with my broader claims, and defend this interpretation from two alternatives. §3 discusses the implications of these arguments for the practical role of a capability approach, which I link to recent disputes about 'ideal' and 'non-ideal' political theories, and address the fourth barrier. After this, §4.1 deals with the fifth barrier, by arguing that the idea that every human is necessarily related to every other can be plausibly understood as, or as equivalent to, a basic commitment of a variety of ethical cosmopolitanism that is widespread in theories of global justice today. Finally, §4.2 concludes the Thesis with a summary of my argument.

## §1 Justice as a Property of Individuals and Relations

### §1.1 What is essential to an Aristotelian approach?

As I've suggested, the most important element of an Aristotelian approach to justice is the idea that justice is concerned with one component of the good life of every individual human being. This respects the thought that there could be nothing unjust about a pattern of social interactions, or an individual interaction<sup>1</sup>, if no wrong were done to some affected individual<sup>2</sup>. Instead, it is precisely when wrongs are done to individuals that relationships are marred, and it is when relationships are marred that a central element in the human good is diminished. Within this tradition, justice, quite generally, can then be thought of as pertaining to right relation to others<sup>3</sup>. Thus, a person is (fully) just iff she is disposed to be rightly related to all wrongable others<sup>4</sup>. There are a wide range of possible contributing factors to a relationship's right- or wrong- constitution, and it is a further question in turn how precisely these factors relate to dignity. For example, it might be that any relationship can only be rightly constituted if it involves certain successful dispositions of care<sup>5</sup>, and genuine care cannot fail to recognise lack of dignity as a negatively salient feature of its possessor; thus, a just person would be committed to experiencing the lack of dignity of the related other as a serious and wrongful failure of their care-disposition to achieve its end<sup>6</sup>. However, I do not need to take a position on the details here, over and above the claim that lack of dignity, in the capability-theoretic sense, is incompatible with right-relationship. Regardless of those details, if justice is something that pertains to relationships, and universalistic, *global* justice must be something that relates to every human being<sup>7</sup>, then it must be the case that every human being is

---

1 On my view, injustices – as occurrent *events* – only happen when other-affecting actions do; what I later call 'structural injustice' is not an event but a worldly state of affairs, and what I call 'personal justice' is not an event but a virtue – a personal disposition to act justly, thereby occasioning just events.

2 As Thompson notes (2004b, 339-40), this is not to say that someone cannot 'do wrong' in some broader sense without wronging someone, it is just to say that nothing can be *unjust* unless someone is wronged.

3 As usual, I do not pretend to non-circularity here – the relevant kind of 'rightness' may not be specifiable independently from the notion of justice itself.

4 This is not incompatible with the thought that there are beings to whom humans can relate, but which are not subjects of justice, since in such a case *any* way of relating to such a being would be a right way.

5 It is easy to imagine close alternatives centring on notions of respect, or even agapeic love, for example.

6 Cf. Held 2006 here.

7 I have stressed before that neo-Aristotelian naturalism needn't build in the idea that only humans can be

(essentially) disposed to have some minimal kind of relationship with every other<sup>8</sup>. The distinctively neo-Aristotelian naturalist ethical claim is that, given opportunities to interact, humans *qua* human are disposed to interact in right, if minimal, ways.

This basic way of thinking about justice has at least two clear benefits. Firstly, it recognises that all human lives are necessarily bound together, with a significant amount of the good of each depending on the good of others not merely instrumentally but constitutively, and provides a way of accounting for that interconnection and its practical import that transcends the vicissitudes of family, culture, nationality, or even acquaintance. It seems possible for human beings with nothing else in common to recognise something in one another that immediately provides a basis for beneficial mutual engagement, and no more parochial account of our interconnection can make straightforward sense of this<sup>9</sup>. Secondly, if it is workable, this kind of account promises to generate a conception of justice that is continuous, recognising no private/public dichotomy that would exclude many interpersonal or social considerations from being assessed as just or unjust<sup>10</sup>. On this view, 'justice' names the first

---

subjects of justice, although my focus throughout the whole Thesis is on humans and what justice might require *vis-a-vis* them. This approach does suggest, however, that the extent to which non-human animals can be subjects of justice will be the extent to which humans and non-human animals are constitutively capable of entering into relationships of equal dignity with one another, or (correlatively) the extent to which humans can *wrong* non-humans (and/or *vice versa*).

8 I intend this as an abductive inference. Although it is logically possible that each human being could be only indirectly related to certain others, and so acquire reasons of justice involving them similarly indirectly, through complicated webs of (some kind or kinds of) relationships which just so happen to connect each individual to every other, this is not an especially plausible explanation of the facts in question. Most inessential kinds of relationship are A) extremely fragile and contingent, B) not genuinely human-universal, and C) do not seem adequate to ground global justice *per se*. For example, familial and nation-state relations are clearly not human universal. Also, economic and political-institutional connections are mostly not human-universal, are generally fairly fragile, and are not obviously ethically relevant in a general enough way – clearly, capability-theoretic justice extends far beyond economic or political-egalitarian factors; these institutional connections also relate many non-human beings, which might not be thought to be proper subjects of justice. For these reasons: because it will at least be the case that an Aristotelian account of global human relations will be clearly and robustly universal, and able to accommodate the right range of subjects, the universal-relationship account gives the best explanation of the phenomenon. Cf. §4.1 below.

9 Cf. Nussbaum (1993, esp. 292-6) on this point.

10 See the discussion in Chapter 2 §1.3 here.

virtue<sup>11</sup> of interpersonal relations and interactions *per se*, not merely of abstractly-conceived 'institutions' or the 'basic structure'<sup>12</sup>. This is an advantage not merely because of its inclusivity and broad scope, but because it presents the hope that we can make sense of social interactions at very different scales in fundamentally the same way.

However, there is an obvious potential threat, which I noted above, and which should be addressed immediately. This is that the view I'm proposing will be individualistic, in a way that is problematic either because it will miss details that are essential to a proper assessment of justice, or because it is implausible in its own right e.g. because it misrepresents the workings of individual agency or the self. At the most basic level, this criticism is relatively easy to rebut. Although this approach is 'individualistic' in a sense – it does require that anything that makes a difference in terms of justice make a difference to an individual's life – it does not restrict the evidence base for enquiries about justice to only those features of individuals that can be considered in abstraction from the social context<sup>13</sup>. It allows that injustice-constituting features can affect people by affecting those to whom they are related, or by affecting groups of which they are a constituent. Since I also accept a potentially very broad-ranging externalism about mental content (see Chapter 4, esp. §1.1 and §1.3), it is also likely to be the case that many individual mental states depend for their very existence and identity on features of other people or of social groups. All that appears to be ruled out by the kind of individualism that this Aristotelian view involves, then, is the thesis that a 'group' could be a subject of justice, or have a good more generally, where that 'group' is itself considered in abstraction from the individuals that are supposed to compose it. But even those

---

11 It is extremely likely that justice, construed (as it is here) as a relational compatibility with dignity, does not exhaust the goods (or the corresponding virtues) that can be realised by human relationships. But this is unproblematic: it remains the case that justice is of primary importance in the relational parts of our lives.

12 Cf. Rawls 1999, 3.

13 In fact, it may not require that there even *be* any such features.

who rail most fervently against methodological individualism in political philosophy do not usually propose anything along these lines. Doing so would constitute an equally problematic inversion of individualistic (usually, liberal) atomism – a kind of group atomism – rather than a correction to it<sup>14</sup>. On the face of it, then, this individualism is unlikely to be problematic in the way that it might seem. However, I'll return to these issues below, in the context of the notion of social structure.

## §1.2 Thompson's Dikaiological Matrix

As I've said, the virtue of justice, on this view, is a disposition to be rightly related to wrongable others. In this section I'll explain further what this means, by giving an exegesis of some of Michael Thompson's relevant work, and arguing that his basic framework is a good fit for my project here. Thompson also raises two worries about his own proposal, however, which I shall argue are not well-founded, and should not pose major trouble for my account.

According to my sketch, relations of justice exist in precisely those cases where there are two beings that can have pairwise relations with one another of one particular kind – a kind that humans clearly can have, and are apt to become conscious of when certain familiar conditions are met. When I realise that justice is a factor in my relationships<sup>15</sup>, I realise that those to whom I relate can be *wronged by me*, and that our relationship can be marred in this way. Plausibly, if I realise that I have failed to be a good friend to S, what I think is not merely that

---

<sup>14</sup> Compare Ch.2, §1.3.3 here.

<sup>15</sup> I recognise that most of the instances of human interrelation that are typical of i.e. global citizenship are not, on an everyday sort of understanding, really 'relationships' at all. What is more, we should not move too cavalierly from the goods embedded in – or, more strongly, embodied by – our more obvious kind of relationships (familial, erotic, friendly, neighbourly, etc.) to this radically more expansive kind. However, this would only provide materials for a solid argument against this relational conception of justice as a contribution to the human good if it were plausible that the way we relate to relatively distant others has *nothing* in common with more local sorts of relationship; I do not need to claim that every good that can be realised in any relationship can be present in the relationships that implicate norms of global justice.

I have behaved wrongly in a way that involves S in the background, but that I have *wronged* S. Thompson's main claim in his paper 'What is it to Wrong Someone? A Puzzle about Justice' (2004b) is that what is involved in the normativity of justice and injustice is a special, essentially relational, logical form. Where this form is involved, normative judgements make reference not merely to one subject, but to two; they say, of one, that it bears a relation to the other. For example, Thompson renders the Hohfeldian notion of a claim-right as 'X has a right against Y that [s]he do A'. The thought is that the argument-places for X and for Y, as well as the content of the relation between them, are unalterable in this kind of normative proposition without the meaning being lost. One illustrative example of a point at which this form of thought seems to operate, which Thompson notes (2004b, 343-5), is in the distinction between civil and criminal law. Traditionally, criminal law is concerned with legal wrongs *simpliciter*, while civil law solely concerns the legal wrongings of (some specific) one against (some specific) other. Whereas in civil law a finding against someone will be a finding that they related in some wrong<sup>16</sup> way to someone else, in criminal law a guilty verdict need not signify that. Even if another person is harmed by the crime, according to the law they will be 'the occasion, and not the victim, of my fall' (2004b, 340).

The overwhelming advantage of using this sort of relational norm – in addition to its clear plausibility as an exegesis of what is involved in thinking about the role of justice in people's lives – is that it can be used to do normative work that many have aimed to accomplish, but whose achievement has often seemed difficult or obscure. In particular, treating claims about justice as essentially involving (some potentially massive number of) relations between specific people can do work in avoiding the excesses associated with consequentialism

16 Wrong, that is, perhaps only according to the standards embodied by the legal system, which might be completely corrupt – Thompson rightly notes that this form of normativity is no more immune to being badly or even appallingly deployed than any other (2004b, fn.19, 345).

without collapsing into the opposite extreme of an inflexible deontology (Thompson 2004b, 334). Closer to home, it can also be used to elucidate the notion of *tragedy* which Nussbaum deploys in her discussion of sufficientarian justice (2000a, 81). The basic point is that because this form of thought cannot be reduced to any non-relational proposition, we are forced to preserve certain things that seem correct in the domain of justice, but which a less pluralistic 'monadic' sort of theory would militate against. It might be, for example, that a decision to personally shoot one person to prevent a similar shooting of twenty people by others is, in the final analysis, the right (certainly the *best*) thing to do overall<sup>17</sup>. Similarly, the best possible strategy for a policy-maker to adopt might involve building a dam and displacing (thus harming the dignity of) the people of a small community, because the more pressing needs of other affected people that would be afforded must also be given due weight. It certainly appears, then, that in some of the appalling situations which the world generates, it is appropriate (in some broad sense) to do things which are nevertheless not *just*; in these sorts of cases, people are genuinely wronged by aspects of the world in which your agency is involved, even if not in a primary or immediate sense<sup>18</sup> by you as an individual, and this does damage to the rightness of your (essentially potential) relationship with them. In this way, some choices that any human agent must make are likely to be *tragic*, in something like Nussbaum's sense: even the best option involves wronging someone. But because, within this category of concepts, we cannot reduce the wrong of an injustice done to x, and y, and z, to some aggregate non-relational wrong (x+y+z), there is no risk that the genuine weight of these wrongings will vanish, as it so often threatens to do, within the tide of worse alternatives.

---

17 Cf. Williams 1973, 93-118.

18 I still hold that they are wronged by you *simpliciter*, however, since you are a part of the world, and it would hardly be virtuous to i.e. pursue a goal of self-isolation in order to minimize your immediate responsibility for injustice, even if pursuing that goal would indeed effectively promote that end. Cf. Young 2011, esp. 95-122.

As well as giving an account of the basic conceptual machinery at work here, Thompson also proposes a neo-Aristotelian naturalist interpretation of its basis, when its application in judgements about justice is involved. His account of this is, unsurprisingly, similar to the one I've given – justice is a 'natural' virtue belonging to humankind as such, and it is thus species-normal to develop both thoughts about justice and practices reflecting those thoughts (2004b, 359-60). The scope of justice will then be determined by the bounds of our species, and membership in our species will (in some senses if not others) be what grounds justice-claims. However, Thompson thinks that his interpretation of this background of justice has two significant problems (although he still prefers it to the Humean and Kantian alternatives he identifies). Firstly, he thinks that there are potential ethical objections to it, arising from the treatment of (probably hypothetical) non-human but rational and social animals. Secondly, he holds that this way of thinking about justice has difficult epistemological consequences, because he thinks that there must be non-empirical but substantive moral knowledge if it is to be true. (2004b, 376-9).

The first issue is relatively easily dealt with. It is perfectly compatible with the basic neo-Aristotelian picture that different life-forms can share elements of their proper form without collapsing into one another. So, if there is, or were to be, a non-human species who nevertheless appeared to be capable of engaging in the sort of social engagements and relationships that we regard as regulated by justice – if they genuinely appear to be wrongable – there is no deep barrier to accepting that justice belongs to their life-form too<sup>19</sup>. Accounts of justice should then include reference to the human life-form *and* this other life-form.

---

<sup>19</sup> Compare my point, at the end of Chapter 4 (§2.3), that it is not deeply inconsistent with the use of a thesis of anthropocentricity in articulating the secondary-quality nature of realism to relax this restriction if that appeared to be well-motivated.

Alternatively, it could be that the proper response to the recognition of wrongability is to think of this putatively alien species as really being identical in form to us, superficial differences notwithstanding. This preserves the basic neo-Aristotelian naturalist thought, and, more importantly, it preserves the species-universalism that has been a desideratum for me throughout. It is not entirely clear why Thompson doesn't consider this a possibility; perhaps, however, it is another relic of a too-narrowly biological interpretation of life-form.

The second issue is more difficult, because Thompson's grounds for thinking that life-form concepts and their normative substance must be non-empirical remain obscure to me; his position is that although which beings are bearers of a given life-form is an empirical question, the constitution of the life-form is not (2004a, 377-8; cf. 2008, 63-82). It seems to me that this is a mistake, and we should regard life-form attributions as empirical in the same way that the property of being 'good' is empirical – we gain knowledge of it by experience of worldly facts. It is true that we have to avoid the thought that life-form is just a matter of statistical regularity, so the epistemology of life-form cannot be *simply* empirical; we cannot simply be bundling up functionings into generalised types according solely to the regularity of their co-instantiation. But this hardly exhausts the possibilities for an empiricist view, especially if we recognise the continuity of life-form judgement with ethics. If this is recognised, there does not seem to be any reason to think of life-form judgements, including ones about the constraints of justice upon the shape of human relations, as any more epistemologically troubling than ethical judgements in general.

### §1.3 Against Vogler's 'Promulgation Problem'

Next, I turn to consider a prominent contribution of neo-Aristotelian thought about justice,

which I think goes badly wrong; doing so should head off some possible internecine objections to my proposal about global justice. Vogler's concern is to revisit a set of considerations that she thinks are present in Anscombe's renaissance paper 'Modern Moral Philosophy' (Anscombe 1958), and argue on that basis for some desiderata for neo-Aristotelian ethical theorising, focussing on theorising about justice. She begins by identifying a need, within any species of virtue ethics, to identify various agents as acting from a single 'source'. Examples of candidate sources are easy to come by – social practices such as sports or etiquette, or language games such as contract-making will provide such sources for all their participants (Vogler 2006, 354-5). However, the mere fact that someone is participating in a practice of these kinds does not do anything to explain why someone *should* act a certain way, if that is in serious doubt (i.e., if we question the worth of the practice itself). Thus, practices, conventions, and language games are not good candidates for the single source of *just* action (356-7). The characteristically Aristotelian idea is then supposed to be that, as a response to these considerations, it is *human nature* that provides both the single source – because humans as such are disposed towards justice – and the justification. Our human nature 'promulgates' the norms of justice to us, as a lawmaker promulgates a law to receptive citizens, and enables us all to act accordingly (357-60).

Vogler's idea seems to be that it is required, for justice to be on a sure footing, that both A) anyone who acts justly acts from a source that is common to all just agents, and B) *this very source* is what gives agents reasons of justice, explaining why they should act justly. It is not clear that neo-Aristotelian thought can meet both of these demands in the way Vogler requires. But this is hardly surprising; it is not obvious that most varieties of thought about

justice could be expected to meet both demands<sup>20</sup>. Her requirement seems to be that we can identify something that is guaranteed to be present in each and every *subject* of justice – each person to whom norms of justice apply – and which is also substantially responsible for the acts of just *agents*<sup>21</sup>. However, surely we want to claim that even non-just humans have reasons to act justly! Since it is *prima facie* unlikely that the just will share many motivational tendencies with the unjust, this suggestion is unlikely to work. This is strange for another reason, too: the obvious candidate for the immediate 'single source' of just actions is *the virtue of justice*. Like all virtues, this is *acquired*, not innate. (There would be no point in dubbing elements of our innate *first* nature as virtues: *ex hypothesi*, such elements would be universal!) Although it is obviously true that the processes of enculturation that give rise to any virtue operate upon – and must therefore be consistent with – *homo sapiens'* first nature, it is the developed ('second-natural') virtue, and not the 'first-natural' prerequisites, that is distinctively present in the just. But Vogler requires, in supposing that neo-Aristotelian naturalism can solve the promulgation problem, that the 'single source' be, precisely, first-natural<sup>22</sup>. In doing so, she seems to set Aristotelian naturalism a task that it could not possibly complete.

So, where does Vogler's conceptualisation of the issues go wrong? The diagnosis I offer is that she requires far too tight a relationship between the immediate source of just behaviour and the common human form which is both its ultimate source and its justificatory basis. It is

---

20 One potential counterexample is theological voluntarism in some Abrahamic mode; as Vogler (and, famously, Anscombe 1958) points out, e.g. the Biblical character 'God' would be capable of playing this 'promulgating' role, given that 'He' would be capable both of giving people ethical reasons, and of being causally responsible for anyone's virtue in acting in accordance with those reasons.

21 It is possible, here as elsewhere, to see certain neo-Aristotelians treading very close to the quicksand of normative (e.g. reasons) internalism, with the threat of relativism/subjectivism that it generates. The insistence that norms apply in just those cases where they already play some psychological role, or could very easily come to do so, is a classically internalist thought. Compare Chapter 4 §2.2.

22 This is clearest at (2006, 358).

true that, if justice is to be a virtue, it will have to be a singular source in the sense that it can be realised in multiple – indeed, since there are no Aristotelian 'natural slaves'<sup>23</sup>, in all – individual humans. But it is not at all clear that we have any need for universal human nature to be anything like the *immediate* – most explanatorily relevant – source of just action. Rather, we must hold that our human life-form disposes us to act justly, but accept that this disposition is extraordinarily easy to disrupt, only coming close to fruition in circumstances that are and have been (at best) very rare – it will not, then be the case that the existence of this disposition will function as a source of just behaviour in any practically important sense.

A second way that neo-Aristotelian naturalism can go wrong is evident in Vogler's discussion of what can be called 'the argument from non-recognition'. This argument begins from the observation that the norms that an Aristotelian naturalist claims to be implicit in the human life-form have not, in fact been widely recognised or put into effect throughout the history of human life<sup>24</sup>. On that basis, someone might conclude that such norms cannot really be implicated by human life as such. Thompson anticipates this form of objection, pointing out that since life-forms are not statistical regularities, there is no more inconsistency in thinking that almost all particular humans have ended up unjust, even though their form involves a tendency to justice, than there is in recognising that almost all particular mayflies die before breeding, even though mayflies *qua* mayfly breed before dying. However, Vogler notes, against this counterargument, that it is not merely isolated individuals, but entire societies (arguably, indeed, almost all historical and contemporary human beings almost all of the time) who have been systematically wrong about either the bounds or the requirements of justice, or both (360-2). She evidently thinks that even though the ubiquity of actual injustice

---

23 Cf. Aristotle 1998, 5-12.

24 Compare Hope 2013, 163-5. Also compare Elijah Millgram's similar argument against Foot, as discussed in Appendix B, §3.

is not logically inconsistent with the thesis that justice is an element in human form, it does tell against it<sup>25</sup>. However, the move to an explicitly ethical conception of the human life-form helps here. It is indeed plausible that were we to look at the history of our species without engaging our ethical sensibilities at all, we could not coherently arrive at the conclusion that justice is something that belongs to our kind<sup>26</sup>. But, as I have argued in Chapter 5, this should never have been the aspiration. When life-form recognition is correctly thought of as an ethical enterprise, it makes sense to insist that humanity as such is constituted (*inter alia*) by a disposition to be just, even though it is undeniably true that this disposition may have always and everywhere been thwarted.

Avoiding these missteps is crucial to formulating a plausible account of the relationship between our human nature and our need for justice. If we expect mutual recognition-as-human to be *easy*, rather than merely possible, or if we insist on using non-ethical tools to do ethical work, we will severely limit the normative import of our humanity<sup>27</sup>, and give succour to the view that it is something else – something less essential and so more hospitable to relativism – that provides the framework for ideas about global justice.

## §2 Personal and Structural Justice

I have argued that a personal-relational model of justice, based on that of Thompson, is highly promising as a component of a capability approach. However, there is a huge barrier to

---

25 Given that Vogler appears to go along with Thompson in thinking that life-form concepts are non-empirical (359-60), it is especially unclear why she thinks this!

26 Compare: if we were to look at the world in general without our folk-biological concepts in play, we could not recognise any species at all!

27 This is something that Vogler shows signs of recognising, but appears to rest content with (e.g. fn.30 on 362).

such an account's success. It may still be unclear how this view can accommodate paradigmatically institutional candidates for justice/injustice such as government policies or private-corporate structures. My approach thus threatens to be least successful for precisely its central target case: that of policies, institutions etc. which operate at the global level. What is more, there might be good reason to think that there are some injustices that a person/relationship-level view cannot help us think about: some injustices may be distinctively *structural*. In this section, I'll acknowledge the important and complex role of structural factors in injustice, but argue that they can be incorporated in a way that coheres with the basic personal-relational thought.

I have claimed that justice solely concerns wrongings within human relationships, and it appears absurd in any case to think of something as unjust if it is in no way contingent upon the ways that humans have related, or do relate, to one another. Given this, it appears that something can only be relevant to justice if it is contingent<sup>28</sup> upon human relational form and/or its history.<sup>29</sup> However, this does not imply that the justice or injustice of my relations and interactions depends entirely upon *my* relational form or *its* history, at least when it is narrowly construed; upon my dispositions to respond variously to related others, or the things upon which those dispositions depend. This is of great importance, because there are good reasons to think that many or even most injustices outstrip the capacity of any individual to

---

28 Although, as I'll note later in this section, there are complications that make it implausible to propose a strong necessary connection between claims about justice and deliberation-relevant ought-claims, there is also something close to an 'ought-implies-can' rationale here: it would make little sense to say that there was something wrong with someone's interactions which could not possibly be (/have been) remedied by altering those interactions.

29 Alternatively, one could note – as Nussbaum has (1995) – that claims about justice must be keyed to what can be expected of human beings, rather than gods. For example, while it might make sense to *wish* that – for example – the Earth inherently contained more resources so that people could live what would in some sense be better lives, it would not make sense to regard such a lack of resources as constituting or contributing to an *injustice* unless that lack were due to the shape of human interrelations.

remedy them, and are not, in this way, reducible to individual agency.<sup>30</sup> Perhaps the most obvious application of this thought within the dignity-centred capability approach I've recommended is the inability of any individual agent to ensure the dignity of all the global others to whom they owe it in the face of the massive inertial force of the global economic structure. It is wildly implausible that the lack of equal dignity of so many people in the world today can be wholly assigned to the vice of *any* single individual, even when we count (for example) failures to engage constructively in collective action as vicious<sup>31</sup>. As Simon Hope (2013, 159) has recently put it: “[o]ne cannot, as Aristotle’s own implausible attempts to the contrary show, infer the justice of a society from the virtues of its members, because some injustices are purely systemic.” There are many different reasons why it might be the case that global justice cannot be attained through individual agency, with the result that no individual can live a fully just life. However, all of them have in common the irreducibly social nature of some of the factors which affect human relationships. This implies the need for a distinctive notion of *structural* injustice, to which this sociality is central. Where the injustice of my relationships is something over which I currently have control, it is not structural; where I cannot bring it about that my relationships are just, the injustice is structural.

## §2.1 Structural Irreducibility, and the Difference it Makes

There are two species of structural factor, given this basic account of what it is for a justice-relevant factor to be structural. Some factors which bear on the rightness of my relationships

---

30 On the face of it, given that there are some cases like this, Thompson's relational view of justice cannot be plausible unless it is supplemented by a set of structural concepts like those I sketch below.

31 I recognise that this may only follow given a moderately high set of thresholds for dignity, and a correspondingly fairly demanding account of justice. However, I think that few would genuinely defend any threshold for dignity that would be low enough that this would not follow; at least, it seems unlikely that any contemporary theorist of cosmopolitan justice would be inclined to do so.

to others are *reducibly* structural – although they cannot be attributed to *my* actions or omissions, they can be wholly attributed to the actions or omissions of others. Thus, although they are not the results of *my* current vice, they may be the result of someone else's, or of a past vice of my own the injustice of which I alone cannot now rectify. This is sufficient for it to be the case that I cannot ensure my own personal justice, or (correlatively) the justice of all of my interrelations, and so these factors alone are enough to mandate a more complex account of the relation between agency and justice. However, there is also a more interesting and significant category of *irreducibly* structural factors, which ensure that the full virtue of justice may be unattainable for all individuals even in the absence of any individual vice at all. Most theorists are likely to agree that the first set of structural factors has many members, if they are willing to countenance a relational or otherwise social conception of justice at all; few individuals are so corrupt that absolutely everything that mars their interrelations flows from their current actions alone! The existence and prevalence of factors of the latter kind is likely to be somewhat more controversial, however. In this subsection, I investigate further the idea that many global injustices are irreducibly structural, and begin to assess the consequences of this insight.

Probably the most significant theorist of structural injustice over recent decades has been Iris Marion Young. Over a number of (mostly posthumously published) recent works (2004, 2006, 2009, 2011), Young developed powerful arguments for the conclusion that irreducibly structural injustices are a hugely important feature of our contemporary world, and that dealing with them requires considerable modification of traditional approaches to normative theory. Young's conception of structural injustice turns on the thought that these are not 'traceable' (2011, 44) or 'attributable' (45) to either specific 'interactions between individuals'

or 'the specific actions or policies of states and other powerful institutions' (45). She uses a central example, Sandy, whose insecurity in housing (and consequent vulnerability to homelessness) is not caused in any identifiable way by aberrant vices of (e.g.) malice or sexism on the part of landlords, nor by the purely random churn of events we call 'bad luck', nor by identifiable unreasonable actions (or vices) of her own, nor by any particular identifiably poor governmental policy decision(s) (43-52). Nevertheless, it seems unjust that Sandy is at severe risk of homelessness through no fault of her own, given that the resources which would remove this vulnerability exist in the broader social context, and given that the social structures which shape the distribution of these resources are malleable in principle (45). She gives evidence that her narrative is realistic, and so it supplies a good motivation for her claim that structural injustice, as she conceives it, is a genuine and widespread problem. Later, Young provides an excellent theoretical discussion of the ways that dynamic social practices can constrain and otherwise shape the possibilities for action of differently situated individuals, and, over time and across space, create enduringly divergent social positions (2011, 52-64). As well as problems arising from economic and other paradigmatically institutional structures, Young also argues that large parts of the social injustices of racism, sexism and heterosexism can be identified as structural, in this sense<sup>32</sup>. For example, just as innumerable small-scale interactions within the housing market, in conjunction with the large-scale institutional features which structure it, can give rise to individual economic vulnerability, so can many (apparently isolated) interactions give rise to (e.g.) racial or sexual disadvantage (Young 2009, esp. 363-9). For instance, an individual's apparently benign disposition to react more positively to more familiar features of others (of dress or dialect, for

---

32 Presumably, some portion of these injustices is more straightforwardly caused by those varieties of unequally-distributed individual malice or explicit prejudice which commonly attract the names 'racism', 'misogyny', 'homophobia', etc. But, arguably at least, these wrongings (or dispositions to wrong) are only reducibly structural, if they are structural at all.

example), or to features of conventionally higher status, can reinforce racial (and sexual/gendered, and disability-based) disparities when ramified over larger periods of time and greater masses of people.

On Young's account, examples such as these abound. What unites them seems to be the fact that it is impossible (at least in practice) to reduce the ethically-relevant aspects of these structures to the individual acts (e.g. of wronging) which are elsewhere the paradigmatic subject-matter of ethics. At the very least, the task set by the epistemology of a traditional ethical account of human interactions, requiring as it does an untangling of different lines of social influence to establish individual causal contributions, is intractable when faced with structures of these kinds. As a consequence, she argues that accommodating structural injustice requires numerous revisions to the way we construct normative theories. Firstly, we need to replace the individual-centric 'liability model' of responsibility, which aims to assign blame and punishment according (mostly) to causal contribution to an injustice, with a 'social connection' model, which aims instead to motivate changes in the lives of all those socially implicated by an injustice, with the degree of change warranted depending both on the extent of any individual benefit from the injustice, or from its prospective removal, and on the degree of ability to enact such change (Young 2011, 95-187, esp. 142-51). Rather than a backward-looking perspective, on which responsibility is always connected to blame, she proposes a forward-looking conception, for which responsibility is connected to the means of rectification. Secondly, on Young's understanding, many instances of injustice-relevant harm (discrimination, exclusion, deprivation, etc.) should be represented as the normal state of affairs within a social practice, requiring a disruptive response, rather than a rare aberration that requires a normalising response (2011, esp. 106-9). Finally, because this understanding of

injustice does not require an identifiably tight connection to (abnormally) harmful current acts, it is able to incorporate analyses of injustices which typically exhibit few such connections. The most important categories of these are global injustices such as trade inequality or cross-border patterns of deprivation (2011, 123-51, cf. Young 2004; 2006), and the enduring relevance of historical injustices such as slavery and colonialism (2011, 171-87).

Before I make explicit the role that Young's account of structural injustice plays in my argument, it will be helpful to sketch some criticisms of her attempt to argue for a change in our conception of ethical responsibility and/or its practical consequences. This is because some of these potential problems are apt to be inherited by my view<sup>33</sup>. Firstly, several thinkers (e.g. Barry and Ferracioli 2013, Jubb 2013) have argued that Young's attempt to replace the liability model is unsuccessful, as she presents it, because it relies on an overly simplistic understanding of liability, and because it is vague about the notion of social connection, which must impute *some* kind of causal influence to agents in order for her to claim that they should respond by changing their behaviour. In particular, Robert Jubb argues that many cases that Young classifies as beyond the reach of liability-based conceptions are really not so, because individuals are liable for 'cases where individual agents have their wills implicated in the patterning institutions achieve by their acceptance of them' (702)<sup>34</sup>. For example, he notes that even someone who actively hopes that their employer's (injustice-

---

33 Several others are not, including many strands in Barry and Ferracioli 2013, and Jubb 2013, which I need not discuss here. For the most part, criticisms in this category are irrelevant because I need not endorse the claim that the liability model is inappropriate *in every case* where irreducible structures are implicated, because given the distinctions I shall make, it is possible (*prima facie* at least) that both the liability model and a more forward-looking model could coherently be applied to the same case. Other examples, such as the effective critiques of elements of Young's revised 'social connection' model of responsibility at (e.g.) Jubb 2013, 705-6 and Barry and Ferracioli 2013, 255-6 are irrelevant because I need not endorse Young's specific criteria for greater responsibility for rectification; particular judgements about responsibility are best left for non-ideal theory, rather than any species of theorising conducted at the holistic, evaluative level that is my focus.

34 Cf. the central arguments (concerning the 'negative duties' of developed-world citizens) of Pogge 2008, which mostly trade on this indirect, participatory sort of liability.

contributory) projects will fail may not be automatically removed from complicity; they may not intend the employer's bad ends, but they intend to work, where this work supports those ends (703-4). However, it remains highly questionable how far this mode of response could take us, even if we grant that some individuals can reasonably be expected to refuse injustice-involving employment<sup>35</sup>. It might be suggested here, against Jubb, that many injustices which are largely caused by bystandership or other forms of traceable complicity<sup>36</sup> are still not helpfully dealt with by a liability model, because the alternative courses of action which would reject the institutional forms in question are not things that individuals can reasonably be expected to choose, or sometimes even to know about, before significant social change has occurred. For example, there are probably alternatives to the various gender norms that currently govern sexual relationships (or, equally, divisions of household labour, or many sites of academic discussion), and which would not ramify to create structural gender injustices. However, it is not clear that many contemporary individuals are in a position to readily inculcate these different norms in themselves; they are unlikely to vividly know<sup>37</sup> what forms such behaviours could take, and the process of inculcation may involve heavy personal costs, because many other individuals are likely to refuse to co-operate in, or even respond punitively to, such attempts. Sometimes, this epistemic lack may follow *a fortiori*, since it may be that *no-one* has a live, detailed sense of these possible norms. However, it is open to Jubb to respond that this counterargument motivates a different kind of conclusion to Young's own: rather than implying a change in the conception of responsibility, it can be taken to show that such individuals are *excused* for the complicity for which they would

---

35 Personally, I doubt that most individuals have such employment options available to them, because almost all employment can be expected to require participation in injustice at some point. This is just a reflection of the larger point that most action of any kind involves participation in injustice in the contemporary world. But I cannot argue for this hypothesis here.

36 My thought is that these are paradigmatic members of Jubb's set, i.e. 'cases where individual agents have their wills implicated in the patterning institutions achieve by their acceptance of them'.

37 Cf. the discussion of Little's point against the Humean theory of motivation in Ch.4, §2.1.

otherwise be liable; this conclusion arises from within the liability model itself<sup>38</sup>.

Secondly, Justin Weinberg (2011) has noted that some of the considerations which lead Young to postulate irreducibly structural injustices as a distinctive category also threaten her own conclusions. On the one hand, it is questionable whether the fact that individuals are systematically ignorant about the ways in which their actions impact global others necessarily rules out culpability related to structure-involving injustices. Perhaps such individuals are culpable for their continued ignorance, even if not for the injustice-contributory actions that must result? On the other, if ignorance of the precise dynamics of social structures is really as pervasive and stubborn as Young's account implies, how are we to arrive at conclusions about possible paths to rectification (227)?<sup>39</sup>

Many aspects of Young's account can be deployed in support of the position I am developing here. Others are inessential, and can be put aside, as would independently be desirable in order to avoid criticisms such as those I have adduced. Most importantly for my project: given the irreducibility of the structures themselves, it does seem plausible that backward-looking, blame-implying culpability cannot be assigned to the individuals who participate in structural injustices (*per se*), even though – from the wider perspective I've developed – the personal justice of these individuals is precluded by them. The category of irreducible structural injustices incorporates several distinct types, which are united by their resistance to the individual treatments of contribution and/or prescribed rectification that typify the

---

38 Barry and Ferracioli also make this sort of point. Cf. 2013, 254.

39 He also wonders whether it is obvious that structural injustices *necessarily* require collective solutions – it might well be the case that systematically different individual patterns of response would reduce or eliminate many injustices even absent institutional or other collective change (Weinberg 2011, 228). Since my account need not build in the idea that collective change is necessarily the only means to rectification, I am not vulnerable to this point.

liability model of responsibility. Some cases are structural because alternative courses of action are epistemically unavailable (at least to individuals), others because alternatives are too costly to be reasonable at the individual level, and yet more because the social mechanisms involved in the injustice itself, rather than (or as well as) the just alternatives, are unknowable at the requisite level of detail. It is also plausible, as Young suggests, that many contemporary structures should be represented as *normally* unjust, and thus that the kinds of relational wrongdoing that they mediate should be regarded as equally normal. This fits with the generally anti-conservative, disruptive orientation that I'll advocate when discussing some of the practical implications of frameworks like mine (in §3.2 below). It also serves to re-emphasise the futility of attempts to pick out individual acts of participation as instances of wrongdoing: although some may be worse than others, all involve wrongdoing, and so none can be presented as different in kind to the norm. Any individual agents that are singled out for blame will necessarily be scapegoats, at least in some respect. Instead, it makes sense to see at least some part of the ethical weight of global structural injustice (considered as a whole) as resting on every human being alive<sup>40</sup>.

However, other elements of Young's account are inessential for my purposes. Firstly, the possibility that the notion of exculpation can be used to accommodate irreducible structural injustice within the liability model, while preserving the general connection to individualised attributions and to blame that I want to deny, needs to be taken seriously. I do this, after

---

40 Some (e.g. Lawford-Smith 2012) have argued that argumentation which proceeds via consideration of the causal contribution of individuals to global injustices may be more effective at motivating rectificatory action than humanity-based arguments. This may well be correct, *ceteris paribus*. However, given the limited practical aspirations of my humanity-based claims, this is largely irrelevant to my position; it has not been shown that humanity-based arguments are ineffective at motivating regret or an impulse to imagination and continued engagement. Secondly, it also seems doubtful that things will always *be* equal – it seems evident that humanity-based arguments are inherently more flexible than causality-based ones, so they are likely to apply in more cases. Cf. also Tronto 2012.

presenting my alternative in more detail, in §2.3. Secondly, the concession that some kind of causal influence over injustices must be present needs to be made. Fortunately, this does nothing to damage the basic case against traditional ways of thinking about responsibility, because that case rests on the intractability of attempts to attribute strong causal influence in a particularly fine-grained way, not on the inability to attribute influence in weaker or more general senses. I have already noted that it would make little sense to connect the notion of justice to states of affairs which do not depend upon human inter-relational form, and the acknowledged fact that each individual's form could be different does enough to attribute some general kind of influence to them, even though this falls short of the stronger causal power that liability models require. Thirdly, I can accommodate Weinberg's criticism by allowing that individuals can be liable for failures to act in such a way as to lessen their ignorance, at least to the extent that such actions would not be so costly as to provide exculpation<sup>41</sup>. Finally, there is a genuine threat that acknowledging ignorance in structure-involving domains leaves normative theories without prescriptive power, and it needs to be clarified how my approach avoids this threat. Both of these last two problems will be discussed further in §3.2 below.

Returning to the more general level, acknowledging the severe limits of our knowledge in the social domains that are most relevant to normative questions concerning global justice may seem to leave views like the capability approach with little or nothing to say about individuals in the contemporary world. As a result, it may seem that the capability approach must fail as an ethical perspective, if not necessarily as a higher-level political one<sup>42</sup>. In the rest of this

---

41 As elsewhere, this invocation of exculpation is compatible with the liability model (at least *prima facie*), so the liability model remains in effect, if only on one level of evaluative analysis.

42 Most of the accounts that may seem relevant to my target questions – about which actions (courses of action, policies, etc.) are possible, and what kinds of difference possibility makes to theories about justice – are actually unhelpful because they do not even attempt to include the discussions of individual agency that

section and throughout the next, I'll provide an account which shows why this is not so.

## §2.2 Structural Injustice as a Barrier to Virtue

There may be several different ways of integrating a Young-style analysis of structural injustice with a relational conception of global justice like that which I've described. However, in the remainder of the Thesis, I'll defend one which seems to incorporate all of the desiderata I've motivated hitherto. I have already gestured at the core idea that I'll develop. On this account, structural injustices (whether they are reducible or not) are barriers to the personal virtue of justice, and the constitutive and external goods it enables. Where the world is structurally unjust, at least some of the (at least (essentially) possible) relations which every individual has will be wrongly constituted, at least in some respects. As a corollary, no individual can be completely just – can instantiate the virtue of justice<sup>43</sup> – while global structural injustice persists. And, given that everyone has reason to live a flourishing life, it will be the case that every individual has reason – of at least some kind – to desire the elimination of such injustices, and to undertake whatever actions are reasonable means to that

---

are necessary to link justice in the *polis* to ethics *tout court*. As an illustrative example, Gilabert and Lawford-Smith 2012 consistently discusses what 'we' can do, and what is feasible for 'us'. If 'we' denotes, as it generally seems to, an abstract representation of the entire *polis* as a single agent, this is of some use for my purposes, since what the *polis* can do seems to be identical to what depends on the shape of human relational form, as I've used that notion. The (probably very spare) limits on this are what Gilabert and Lawford-Smith call 'hard constraints' (811-13). Cf. fn.54 below. But, of course, what is feasible for 'us' will only extremely rarely be feasible for 'me'.

- 43 A less stringent concept of virtue might be able to avoid this result: if virtue is merely the best instrumental means, rather than something which embodies constitutive value of its own, it might make sense to exploit the non-sufficiency of virtue for the good here. Given such a view, it might appear that structural injustices could be considered barriers to goods that are external to agential constitution alone. This strategy, however, precludes the thought that structural injustices prevent *constitutive* goods of justice from coming into being, rather than just blocking the normal external-good-promoting effects of the virtue (/disposition). This position should only be palatable to those who are prepared to accept that there is nothing constitutively worse about being unable to relate rightly to others, but doing the best that can be expected regardless, than doing the best that can be expected, and relating rightly to others as a (perhaps partial) result. If such a position is plausible, much of the most distinctive argument in this Chapter, and some of my earlier arguments too, can probably be rejected in favour of weaker alternatives. Cf. Tessman 2009b, which offers some argument for maintaining this strictness.

end. There are two ways in which structural injustices constitute barriers<sup>44</sup> to just flourishing of our interrelations, which correspond to my earlier distinction between reducibly and irreducibly structural injustices. Most clearly, where injustice is irreducible to the relational form of any individual human, no agent will be able to act with a reasonable expectation that just relations will be the result. However, even those structures whose justice-relevant inadequacies can be attributed to the agency of some set of individuals other than the current agent prevent the current agent from rectifying things, ensuring that elements of wrongdoing will persist in the way she relates (/would relate) to the affected others. In both kinds of case, structures make it impossible (however contingently) for rectification to occur by the current agent's action alone<sup>45</sup>. And, given the prevalence of such structures, it is overwhelmingly plausible that all contemporary agents are incapable of the virtue of justice *simpliciter*<sup>46</sup>.

One possible problem with this line of argument is that it may seem to be the symptom of an inwardness that is entirely inappropriate to the subject-matter of justice. It may seem self-obsessive to even *consider* personal virtue or personal flourishing when severe deprivation and appalling inequality are in the frame. It is worth elucidating why this thought is misplaced before I continue. In line with the flexible McDowellian concept of virtue argued for in earlier chapters, the theoretical function of virtue is not to provide an independent or

---

44 This kind of barrier to virtue is different to at least many of those postulated by Lisa Tessman in her important account of the 'burdens' that oppression imposes on agents, burdens which militate against virtue (see especially her 2005 monograph), and consequently against flourishing (see especially her 2009b). These are often different in two ways – firstly, in that many of these barriers are specific to the most oppressed, unlike those I postulate, which are universal to all participants in an irreducibly unjust social structure; secondly, in that (again, most of) these barriers involve the absence of specific circumstances necessary to inculcate virtue, rather than preclusions of the conditions necessary to instantiate it more broadly.

45 It is worth noting that this follows even for actions aimed at co-operation with others. Even such actions depend for success on external conditions, so this negative result obtains even when the current agent is as well disposed as can reasonably be expected towards collective action.

46 As I have noted before, this is compatible with its being the case that they are rightly related to certain individuals.

foundational ethical idea, but to unify the otherwise disparate elements of our ethical perspectives. Bearing this in mind reveals that the connotations of self-regarding moralism that might accompany much of the language that I have been using are entirely misleading<sup>47</sup>. The claim that it is important to recognise structural injustice as a barrier to personal virtue and (consequently) to individual flourishing should not distract attention from our (*de re*) wrongings of others, or the plight that this may leave them in; rather, it places those external facts in the context of our whole lives, and helps to make clear their full significance. On such a way of thinking, there is no conflict between the thoughts 'racism is bad because it prevents me from living well' and 'when I perpetuate racist structures I wrong/harm others terribly'. The benefit of a well-ordered conception of virtue is that it reveals these to be two parts of an integrated whole.

The revisions that I am arguing are required in order to incorporate the insights of a view like Young's are philosophically difficult, relying on the generation of distinctions that, in some ways, go against the grain of (at least large sections of) traditional practical thought. However, there are links to be made which help to alleviate such difficulties. In particular, I have already argued<sup>48</sup> that there is a category of *reasons* ('V-reasons') of which it is true that, when they are distinctively present<sup>49</sup>, they cannot function so as to motivate action within their immediate contexts of deliberation. Likewise, I have accepted that it is not plausible to connect such reasons to practices of blame. Unsurprisingly, my suggestion is then that, although the existence of structural injustices drives a wedge between claims about *D-reasons* (and best-possible action, and potential blameworthiness), and claims about virtue

---

47 This claim would bear comparison to the more general line of criticism of virtue ethics that it is fundamentally narcissistic. See (e.g.) Hursthouse 1999, 190; Cox 2006. Unfortunately, I cannot examine this tangent in the Thesis.

48 See Chapter 4, §2.2.

49 That is, when they are non-identical to the D-reasons which are present.

(and the internal goods of justice, and the sort of non-tragic action that we should *accept* rather than merely *tolerate*<sup>50</sup>) there need be no such disconnect with *V-reasons*<sup>51</sup>. Thus, it is still possible, despite the difficulties which social complexity generates, to make claims about the reasons of all individuals on the basis of the conjunction of two facts alone: their humanity, and the fact that the bundle of functionings I am calling 'dignity' is not universally attained in the contemporary world. If this is so, the capability approach remains an extremely promising ethical/political framework. The price is that we have to restrict the scope of the theoretical claims that the framework includes to a comparatively holistic and idealised kind of evaluation, rather than necessarily aspiring to provide specific action-guidance for the actual, non-ideal world. My conclusion at this point is this: to the extent that irreducible structural injustices are a persistent feature of our world, we should view theories of global justice such as the capability approach as, most basically<sup>52</sup>, theories about *V-reasons* and about *acceptably*, rather than *tolerably*, just action.

An additional benefit of the connection can be seen by reflecting on the fact that although *V-reasons* are reasons for action<sup>53</sup>, the point of noting them is not necessarily to motivate action,

---

50 See §3.1 below for more on this distinction.

51 The correspondence between D-reasons and non-structural features, and V-reasons and structural ones, is not a perfect one. As I noted in my original discussion in Ch.4, there are some examples of gaps between D- and V-reasons which result from circumstances that, although they are non-ideal in some respects, are not structurally unjust (or, at least, not obviously so). Examples would include the lack of needed skills which take time to acquire, or the presence of ingrained vices or biases on the part of an agent. The claim is that *all* structural injustices generate divergences between D- and V-reasons, not that *only* structural injustices can do so.

52 There are many who deploy the capability approach as what would, in my terms, be very much a non-ideal theory. Much of Sen's work has fit this general pattern (see especially Sen 2009), even when it is not used to, for example, contribute to UN poverty measurement strategies (ul Haq 1990, esp. 26-9). See also Robeyns 2006 for discussion of several foci of non-ideal deployment. There is no reason why the broad notion of 'capability' cannot find a place in both sorts of theory.

53 It is important to stress this, since it might seem natural to interpret them merely as reasons to have emotional responses such as regret (for example). Although I am claiming that people do have such reasons in the sort of non-ideal contexts in question, it is important to preserve the idea that people have reasons to act well even when they cannot, precisely because the failure that doing so registers is what *generates* the reasons to regret. If people were not failing to act well (e.g., justice-wise) – failing to do something they have reason to do – it would not be clear why they have reason to regret the actions that they must perform

given the contexts in which they are reasons. Instead, the function of postulating V-reasons (in those non-ideal contexts where D-reasons and V-reasons come apart) is more holistic and evaluative than atomic or deontic. Rather than primarily aiming to describe what individuals are obliged to do, or what is best for them to do, situated as they are, accounts of V-reasons ('ideal theories', in the sense I'll explicate) mostly regulate our agency less directly. In §3, I shall do more to elucidate this distinctive (but also strongly limited) function.

### §2.3 Two Alternatives: a Conditional Interpretation, and Exculpation-Maximalism

Before I move to consider the consequences of this conception of the role of a capability-theoretic framework, it will be helpful to quickly distinguish it from two possible rival attempts to accommodate some of the desiderata I've noted. The different ways in which they fail to do so are illustrative of the distinctive benefits of the 'barrier'-based interpretation, despite its perhaps more radical departure from philosophical orthodoxy.

The first of these alternatives is a development from the tradition of 'fact-insensitive' idealising political theory epitomised by G.A. Cohen (esp. 2008, 2009), but excellently summarised in recent work by Pablo Gilabert (esp. Gilabert 2011). Cohen's distinctive contribution to debates about ideal and non-ideal political theorising involves the insistence that there are fundamental principles which do not depend (either epistemically or metaphysically) on whether or not a wide range of facts about the actual world obtain. For example, Cohen thinks that the truth of many (the 'fundamental') normative principles does not presuppose any facts, including most importantly facts about 'the human situation' – e.g.

---

to do the best they can.

about existing economic structures, or concerning the values espoused by the majority (2008, 229-273). One implication of this is that normative principles are true whether or not they can feasibly be put into practice. This may seem paradoxical, since the feasibility of an action involves whether or not an agent can perform it, and, in the classic Kantian slogan, 'ought implies can'. Gilabert urges us to resolve this tension by reading Cohen's fundamental principles as making 'evaluative' rather than 'prescriptive' claims. The distinction is this: evaluative claims are of the form 'A ought to  $\Phi$  in C *if they can*', while prescriptive claims have the form 'A ought to  $\Phi$  in C *given that they can*'. As a consequence, while prescriptive claims are defeasible by (because inconsistent with) certain facts about what can be done, evaluative claims are not so defeasible. (Gilabert 2011, 55-9). In Cohen's words: 'we find fundamental<sup>54</sup> justice within claims of the form: if it is possible to do A, then you ought to do A' (2008, 252). For Cohen and those within this tradition, there is a distinctive, 'ideal' kind of theorising in political philosophy, whose focus is on evaluation, rather than prescription, construed as above. One alternative interpretation of the content of the capability framework I've offered, then, is as an ideal theory of this sort.

The problem with a Gilabert/Cohen-style proposal, if it were reintroduced in an attempt to accommodate the need for ideal and non-ideal species of normative thought, is that it does not make good sense of the impossibility of good (evaluatively adequate) action. It is probably the case that there are many true conditionals of this 'evaluative' form. Recognising these truths might also be linked to an appropriate wishful sentiment, to the effect that it

---

54 Although the position I am developing in this Chapter agrees with most of Cohen's claim that 'what's true in "ought implies can"' does not show that fundamental normative truth is constrained by what it is possible for people to do' (2008, 250), I do not claim that the truths of ideal theory are more 'fundamental' than those of non-ideal theory. In particular, unlike Cohen I cannot claim that non-ideal theory is in some sense a derivation from ideal theory, in conjunction with numerous empirical truths; I think that no such derivation is possible. Also, my position may be less radical than Cohen's, in that I accept the restriction of justice to what depends on human-relational form.

would be better (justice-wise) if some (or all) individuals were able to do certain things that they cannot. However, the difficult conclusion that I have argued is motivated by the recognition of structural injustice is that all the possible options are *inadequate* ('bad') justice-wise, not merely that they could be better. A conditional claim is not sufficient to recognise this, precisely because the relevant conditionals do not, actually, warrant negative evaluation of action in non-ideal circumstances; the claims 'S was virtuous in not- $\Phi$ ing', and 'S's not- $\Phi$ ing was just' are perfectly compatible with the conditional 'S ought to  $\Phi$  if she can', given that she cannot<sup>55</sup>. This is not what I have argued we want: it cannot make full sense of what is regrettable about the tragically unjust action that inevitably accompanies structural injustice.

The second alternative exploits the possibility that wrongful action in contexts of structural injustice might be systematically exculpated, which I noted in §2.1 above. As I implied there, if this possibility can be utilised across the board, this would remove the motivation for the claim that just action is genuinely impossible in such contexts, by providing an alternative within the confines of the liability model of responsibility. The most promising version of such a broad-based exculpatory strategy would work by noting that the difficulty of structural injustice arises because of the serious epistemic limits it involves. Against such a background, knowledge of any possibly rectificatory actions would be extremely difficult to come by, and this explains the inappropriateness of blame: rather than being because the best available actions are still bad, it is because they are sufficiently hard to know about that everyone who fails to perform them has an excuse. My response to this has two parts. Firstly, at least while

---

55 I should make clear why the account I'm developing does not, in fact violate 'Ought implies Can', any more than Cohen's must. The different kinds of 'ought' judgement that would accompany the two kinds of 'reason' judgement simply refer to correspondingly different reference-classes for 'can'. There are conceptual restrictions on what can be asserted using 'V-ought', just as there are using 'D-ought'. My references to what depends on the shape of human interrelations provides an (intuitive, vague) reference-class for 'V-ought'. I am indebted here to Amy Berg, who has (in Chapter 2 of her forthcoming Thesis (2015)) given an account of one way these different levels of 'ought's and 'can's could work. Cf. fn.42 above.

we remain at the individual level, it just does not seem likely that there will be *any* available actions that it would make sense to think acceptable, whether or not the situated individuals could know about them if they did exist. To think this involves supposing that there is a way in which an individual could act such that, as a result of their action(s) alone, they will realise a disposition to be rightly related to all human beings. It seems wildly unlikely that such options exist with any regularity in the contemporary world<sup>56</sup>. Secondly, as a result, it does not seem true that those agents who behave *most* justly in the contemporary world are, in fact, acting massively far from optimally, such that they would be liable to blame if only their epistemic position were not so poor. (Think of Martin Luther King, for example...) It does not seem plausible, then, that the general explanation for the inappropriateness of blame in circumstances of global structural injustice is the fact that otherwise liable actions are systematically exculpated, rather than the fact that all available options are unacceptable, and tragic choices must nevertheless be made.

Given that these alternatives are inadequate in the ways that I've indicated, I conclude that my way of incorporating a recognition of structural injustice into an ethical framework for global justice is well-motivated. In the next section, I'll move on to consider the consequences that acceptance of the idea of barriers to virtuous (acceptable, etc.) action has for the practical functions of theorising about global justice.

### §3 What are the Practical Consequences?

Most of what I've said thus far concerns the internal structure of a capability approach to global justice, rather than what pragmatic roles it can play in issuing normative verdicts on

---

<sup>56</sup> Note that, in formulating this challenge, I do not need to refer to the individual's epistemic situation, only their actual set of options.

actions, policies, personal dispositions, and other contributors to the shape of human interrelations. In this section, I remedy this deficit by further locating the approach I've argued for relative to recent debates about ideal and non-ideal theory, and so bringing together some elements of the approach that might otherwise remain inchoate.

As I've already hinted, I think the best interpretation of a Nussbaum-style capability approach is as an ideal theory. On the distinction I'm deploying here, ideal theories aim at relatively holistic evaluation of situations and exhibit relatively low levels of epistemic dependence on detailed description of existing conditions, while non-ideal theories draw heavily upon such description in order to generate (potentially fairly fine-grained) prescriptions for immediate action. These prescriptions, unlike the evaluations that ideal theories are apt to issue, may be appropriately connected to detailed, situation-specific judgements of liability, blameworthiness, and obligation, in addition to their central judgement about which of the immediately available actions are most choice-worthy (most supported by D-reasons)<sup>57</sup>. However, the degree of abstraction from the non-ideal features of the contemporary world that is necessary to produce an account of acceptable agency – agency consistent with the disposition towards right-relationship – makes it unlikely that ideal theorising could produce reliable judgements of these kinds about similarly specific cases. Instead, as I have already emphasised, the central role of ideal theories of global justice should be to deliver negative evaluations of contemporary actions, and to provide as coherent and unified an account of the grounds for such evaluations as possible. Doing this will involve precisely the kind of reflection on the human goods that justice implicates – on 'dignity' – that Nussbaum-style

---

<sup>57</sup> As this should make clear, I am not claiming that it is plausible that contemporary individuals have few obligations of global justice, or that they are not liable for their failures to satisfy these obligations, or that these do not provide them with powerful D-reasons. All I am claiming is that it is not plausible that a framework such as the ideal-theoretic capability approach I've developed can be expected to generate a good account of such things to the extent that they *do* exist.

capability approaches have at their methodological centre. This may appear to be a frustratingly thin remit for ideal theory, however, so in the subsections below (§§3.1-2) I note two important subsidiary functions that such theory can play too<sup>58</sup>.

Before I move on to that, it is important to make explicit why methodological aspirations like those of my proposed 'ideal' theory-type need not be scuppered by the ignorance of structures and their connection to individual agency that it avows. Primarily, this is because the ignorance here is of the detailed ways in which individual agency gives rise to the non-ideal structures that actually exist. It is not, in general, clear why ignorance about the precise mechanisms that current injustice involves would vitiate apparent knowledge about what justice requires. However, some knowledge of the ideal *might* depend on knowledge of social mechanisms<sup>59</sup>. Specifically, it is likely that such ignorance does rule out very *detailed* knowledge of capability-theoretic dignity; exactly which labour regulations are compatible with the capability to enter into relationships of meaningful recognition with other workers<sup>60</sup>, for example? Fortunately, though, capability approaches modelled on Nussbaum's own have little reason to aspire to this fineness of grain<sup>61</sup>. Such detail would only be of significant practical relevance at the point where choices between multiple *prima facie* good options<sup>62</sup> became realistic; because such feasibility would require dramatic convergence towards the

---

58 Because my proposed framework deploys a sufficientarian distributive principle at the ideal-theoretic level, but remains neutral about which kinds of distribution are warranted in non-ideal domains, it can more clearly accommodate prioritarianism about the kinds of case for which it is most plausible. That is, it can accommodate the thought that non-ideal actions aimed at rectification of injustices should prioritise the worst affected by those injustices. See Arneson 2000 (55-9), for a critique of Nussbaum's view on this issue.

59 Given that I am conceding that justice-claims can only be made against the background of the judgement that the relevant circumstances are affected by the shape of human relations, such claims also have empirical presuppositions, which are defeasible in principle. I think it is defensible to neglect this, however, because I take the claim that human-relational form affects most aspects of human life to be extremely well-supported, and because, to the extent that it might be the case that no amount of change in human-relational form would eliminate some (putative) injustices, we do not have good evidence that this is so. Compare Gheaus 2013; Gilibert 2011, 60-63.

60 See Appendix A.

61 Compare the discussion of epistemic limits in Nussbaum's theory in Chapter 2, §2.3.2.

62 Which is to say, apparently non-tragic choices; choices apparently between genuinely acceptable options.

ideal, we could expect our epistemic position with respect to such fine-grained ideal-theoretical questions to improve concurrently<sup>63</sup>. As a result, a view with the limited aspirations I have proposed faces relatively little threat from contamination by our ignorance of structure and its agential correlates.

A conception of global justice that builds within the framework I've presented can function as a regulative ideal for global-justice discourse and practice more generally, in the sense that it provides a standard against which the world can be assessed. However, as I've stressed, this does not mean that it can generate many particular prescriptions about how people – or institutions – should act so as to make the world more just<sup>64</sup>. This contrasts with Nussbaum's strategy in recent years, which has been to accept (as I do) that her approach cannot generate fine-grained prescriptions even regarding the policies of fairly large institutions – much less individual agents – but to insist that it does generate prescriptions at the extremely abstract level of constitutional principle. One reason to prefer my strategy is that it does not build in the notion that any particular institutional form (the nation-state, for example) will be sufficient for justice to be done<sup>65</sup>; all that matters is that everyone be assured a life of equal dignity, it does not matter whether, for example, this is the outcome of national-constitutional guarantees.

### §3.1 Regulating Regret

---

63 After all, at the limit of such progress, ideal and non-ideal theory would collapse into one another, and the distinction would cease to be of any use.

64 Certainly, a filled-in picture of the just world in the vein of my suggestions should not be taken as an end-point in a consequentialistic sense, with the only questions about justice then concerning the best way to get from here to there. This kind of 'ideal theory' would *indeed* be vitiated by the extent of our ignorance about social mechanisms; because the absence of such ignorance would signal the end of structural injustice, this would also lead to the collapse of the two theory types into one another.

65 Compare Ch.2, §1.3.

The first subsidiary function that ideal theories can play involves the articulation of reasons for regret. As I have already suggested<sup>66</sup>, cases where one has acted, or faces the prospect of acting, unacceptably are cases where it is appropriate to regret this fact<sup>67</sup>. In this subsection, I shall draw on work by Lisa Tessman<sup>68</sup> to elaborate on and support this contention.

Much of the background to Tessman's argument is reflected also in what I have said above. Tessman's work is located in a recent tradition of proposals for non-ideal theorising which is associated most strongly with Charles Mills' 2005 paper 'Ideal Theory as Ideology'. There, Mills provides an effective argument for the conclusion that ideal theories – construed, in Rawls' sense<sup>69</sup>, as theories which justify principles using a model that assumes full compliance of participants to its governing norms (169)<sup>70</sup> – are systematically unreliable and will tend to perpetuate structural injustices. For example, he argues that because Rawlsian principles are designed for a world in which there would be no racist biases (arguably, a world containing no races at all), they will not generate useful principles for the actual world, with its history of racial oppression and its continuing effects (169-182). Such a constraint militates against (e.g.) justifications for potentially rectificatory policies like affirmative action programmes, because such programmes must involve discrimination on (ideal-theoretically) arbitrary grounds. As Mills puts it, 'by assuming the ideal...one is only guaranteeing the perpetuation of the non-ideal' (182).

---

66 Throughout this Chapter, and also in Ch.4, fn.88.

67 Arguably, the connection to regret is also implicit in Young's account (e.g. at 2011, 106).

68 Cf. also Gheaus 2013, 454-5.

69 See, e.g. Rawls 1999, 8.

70 Mills actually lists a variety of features of theories that makes the ideal in the sense he is concerned to attack (2005, 168-9). However, all of these can be seen as aspects or consequences of the full compliance assumption. For example, it is only necessary to assume that everyone has idealised capacities of rationality because such capacities will (on the relevant views) be required for everyone to reliably comply with the norms of justice (whatever those are then argued to be).

Tessman (like myself) accepts the general thrust of this argument (2010, 805-8; cf. esp. 2009a, xvi-xviii), and so accepts that the methodological stance of the non-ideal theorist is largely correct. In particular, she accepts, as I do, that 'action-guidance' – the issuance of prescriptions – needs to be the exclusive province of non-ideal theorising (2010, 808)<sup>71</sup>. However, she thinks that because non-ideal theorising is so much better at performing this role, its proponents tend to fail to recognise that this is not the only thing that we should want from a theory:

What is displaced by this sort of non-ideal theory is an acknowledgment that there are irrectifiable wrongs, irreparable damage, and uncompensatable losses, as well as ways in which oppositional acts aimed at challenging (or surviving) some aspect of oppression may conflict (as in a dilemma) with other such acts to produce a situation in which new or continued moral wrongdoing<sup>72</sup> is unavoidable. Just as Mills points out that aiming at the ideal of a society with no (acknowledged) history of injustice obscures the wrongs that need to be righted, I am suggesting that aiming (only) to rectify the wrongs of oppression can eclipse the inevitability of failing at this task (809).

Rather, she proposes:

an addition to action-guiding normative theorizing: some kind of theorizing that can witness and evaluate the particular piece of moral experience that lies in the gap between a nonideal world and the unattainable, worthy ideals of someone who has not suffered the adaptation of normative expectations (813-14).

One of the functions of ethical thought is to motivate, and make sense of the motivations for, the distinctive experience of someone who must act unacceptably: the experience of tragic regret. The capability approach to global justice, construed as an ideal theory in my sense, is well-equipped to perform this role.

---

71 It is less clear what she thinks about deontic statuses, attributions of liability, and so on.

72 One issue that remains undecided here concerns the choice between linking 'right' action to the balance of D-reasons – which is the interpretation that seems closest to ordinary usage and philosophical orthodoxy – and linking it to V-reasons. On the former interpretation, what I have been saying in this Chapter is that right action often involves wronging someone, and that right action comes apart from virtue in such non-ideal instances. On the latter, right action as well as virtue (and action from virtue) becomes impossible in situations of structural injustice. Tessman generally seems to take this latter line, arguing, for example, that '[i]t is especially repugnant, however, to use a cost-benefit analysis that produces a prescription for moral wrongdoing but to think of the prescribed action as morally right' (2010, 810-11). Cf. Ch.5 fn.70.

### §3.2 Motivating Imaginative Engagement

The second function is even more important. The recognition that we must currently tolerate bad actions, by ourselves and others, can play a role in pressing us to look beyond the present moment for opportunities for more radical change that might disrupt the normality of injustice. This is of especial importance when, as I have urged is the case, the immediate (unjust) situations in which individuals find themselves are fringed by significant uncertainty about the extent of change that may be possible, and thus about which, perhaps radically better, ways of living might become available over time.

A recent paper by Anca Gheaus (2013) is particularly suggestive of an effective argument here<sup>73</sup>. Her (most relevant) aim is to argue that 'a willingness to apply the label 'unjust' to regrettable situations that we cannot fix is going to *enhance* the action-guiding potential<sup>74</sup> of a conception of justice, by providing an aspirational ideal' (448.) The critical point that Gheaus makes is that we do not securely know what the limits of possibility for human relations are; indeed she suggests that we may well be better at knowing what we ought to do than what we can do, at least at such a broad level (460). Since this is so, we will be better at acting, given the

---

73 One potentially important difference between my position and that of Gheaus is that I do not need to accept that 'there is a gap "between the idea of injustice, on the one hand, and the idea that some people are treating others unjustly, on the other"' (448, the embedded quote is from Stemplowka and Swift, 2012). On my account, there is no such gap, because the existence of structural injustice necessarily renders individual relationships (and the interactions they include) unjust (dispositionally/virtue-theoretically speaking). On the other hand, elsewhere Gheaus puts the question in terms of whether or not 'justice necessarily involves rights violations' (448). Especially if rights are given a Hohfeldian interpretation – as something with inherent deontic relevance – or are viewed as inherently connected to a reliable means of enforcement, it would make sense to view me as accepting this version of the gap. A second important difference is that Gheaus is concerned to argue that even with claims of justice 'addressed to humanity', it would be useful to propose 'ideal' (infeasible) conceptions of justice (449). This invites contradiction, because I have been accepting that justice is the first virtue of (and depends upon the possible shape of) human relations. To avoid confusion, I run a more moderate version of her argument, rather than one conducted at the level of humanity as a collective.

74 This conclusion may seem to be at odds with my position, because Gheaus regards the (relevantly) 'ideal' theory as generating action-guidance, which I have explicitly denied. This apparent tension is illusory: it is not that Gheaus thinks that the ideal theory can directly tell us what to do (she is implicitly conceding that the ideal-theoretic prescriptions are probably impossible for us, at least as things stand), rather, the existence of a currently-infeasible ideal generates pressure to make things such that it is no longer infeasible.

presence of larger numbers of agents, and over larger periods of time, if we devote some resources to making the infeasible feasible, than simply trying, at each given level of 'zoom', to secure the most possible justice from that narrow vantage point. Because there are different levels of feasibility, depending on how far out we 'zoom' (the individual, a community, a territory, humanity itself), and which time-scale we restrict ourselves too, there is no contradiction between conceding that injustice is a tragic necessity *for individuals, here and now*, and holding that nevertheless justice is a possible condition *for humanity, in the unforeseeable future*. As a result, any uncertainty that there may be about the limits of the possible at a wider level than the level of action (for my purposes in this Chapter, the level of the individual agent) generates some degree of (D-)reason, at the level of action, to pursue activities that have the potential to expand those limits. Given that, as I have argued, the limits of possibility for individuals in situations of structural injustice are often themselves due to epistemic limitations, these feasibility-expanding actions will frequently involve attempts to improve our epistemic position, for example by generating better *non-ideal* theories of justice<sup>75</sup> and the social structures it involves, or attempting to invent new ways of relating to others that might be less corrupting.

One illustration of this might refer to the gender norms that contribute to structural gender injustice, as referred to in §2.1 above. As I said, it is likely that few contemporary individuals have a live sense of exactly which patterns of interrelation would, if widely inculcated, eliminate gender injustice. However, the socially-constructed and historically fluid nature of such norms suggests that even very rapid change is possible (at least, if we 'zoom out' a little).

This generates a motive for imaginative experimentation and open-minded, potentially

75 In this way, somewhat paradoxically, ideal theories may function to focus more attention on the activity of non-ideal theorising *itself*, while removing some attention from the other prescriptions those non-ideal theories might generate.

transformative<sup>76</sup> engagement with diverse others, which would not be present on a view that restricted itself to current epistemic horizons. Such pressure increases the likelihood of improvements in our epistemic positions, which may in turn make much more just relationships a possibility.

Because most such actions are not very costly, and do not (at least initially) require massive changes to the patterns of activity in most individuals' lives, it is unlikely that they will be subject to exculpation<sup>77</sup>. In one respect, then, an ideal, sufficientarian capability approach would provide some input into non-ideal theory: it would describe the conditions in which there would no longer be strong, justice-based reason<sup>78</sup> to seek the relevant improvements in one's epistemic position. In so doing, it makes clear that such (D-)reasons are still powerful. If the rest of the components of a non-ideal theory are 'co-operative', so to speak – if the actions that imaginative engagement involves are not so costly as to generate exculpation, and would not interfere with other D-rational actions such as to become unjustified (e.g. sub-optimal, or impermissible) – then the content of an ideal theory can (indirectly) affect D-reasons even for non-ideal circumstances.

These two subsidiary roles effectively complement the general, holistically evaluative function of ideal theories that has been my emphasis throughout §3. In both cases, it is not that the ideal theory directly justifies particular actions, it is more that the shaping of general

---

76 That is, of the kind that I identified in Ch.4, §2.2.

77 Cf. Weinberg's first criticism of Young, as discussed in §2.1 above. Here I am conceding that – because action-guidance is the domain of non-ideal theory, even if ideal theory can affect it indirectly – it may be that people are liable (and, e.g., blameworthy) for any failure to do at least low-cost things to improve their epistemic position and those of related others over time.

78 Of course, this is not to say that there would no longer be *any* such reason: inquiry may be valuable for its own sake, and it may well be that the virtuous would persist in such reflection and engagement even in the absence of structural injustices.

motivational forces (dissatisfaction, curiosity, regret) that the content of the theory regulates has knock-on effects within most immediate contexts of deliberation (which remain nevertheless the basic province of non-ideal theory). In concluding the section, I'd like to return to the bigger picture this Chapter has laid out, and sound a relatively harmonious note. As this section has clarified: even in the face of all the difficulties I have noted, an agent can, intelligibly, aim at a just life within a just world – as with the other components of *eudaimonia* – at least in the sense that she will not rest content unless we get there. In this sense, it can be the *telos* of our action even when our action does not (because it cannot) 'aim' at it by being a predictably effective means to that end.

#### §4 Capability-Theoretic Justice Revisited

The theoretical framework I've presented in this Thesis is not, in most respects, a radical departure from existing proposals. In particular, as well as the work in feminist and neo-Aristotelian metaethics (especially in epistemology), the recent work in political philosophy that has been the focus of this Chapter, and the starting point that Nussbaum's work provided, there is a further connection to a dispute within cosmopolitanism concerning the basis for cosmopolitan reasoning. Roughly, one faction in this dispute argues that cosmopolitan claims are based upon the existence of (universal) contingent social connections, while the other argues that such connections are not necessary for cosmopolitan conclusions to be warranted. Because my framework takes a side in this dispute, I shall begin the final section of the Thesis by locating it within this terrain. In doing so, I shall further describe the universal human relationship which I've postulated, urging that, far from being an outlandish proposal, it is an exegesis of a commitment which many – if not most – cosmopolitans already possess. After that, I shall complete the process of drawing the Thesis to a close, by stating some final

conclusions.

#### §4.1 The Wider Context: Cosmopolitan Universalism

I have claimed that the right way to interpret claims about cosmopolitan justice – justice at the global level, without essential reference to sub-global institutions – is as claims about the right-/wrong-relationship of individuals to one another, with the universal relevance of these considerations explained in virtue of an essentially human disposition towards such relationships, whether actual causal social connections exist at a given time or not. In turn, the capability approach, interpreted as an ideal theory, provides the prospect of a reasonably systematic articulation of the sufficient conditions for compatibility with right-relationship, in terms of a diverse bundle of functionings and capabilities. If worldly conditions are not such that every individual possesses this bundle – possesses 'dignity' – no individual can, at a personal level, realise the virtue of justice. This, abstractly, is the harm which structural global injustice does to every human alive; the source of the reasons that we all have to eliminate it.

Because this interpretation does not base the demands of cosmopolitan justice on the contingent existence of overarching social practices, but on the fact of humanity itself, it is 'humanity-based' rather than 'interdependence-based'<sup>79</sup>, in the sense given by Simon Caney (2009, 390-1)<sup>80</sup>. As such, it fits within a significant tradition, many of whose supporting arguments can be relied upon to defend the claim of essential disposition-to-relationship which is at the core of the framework I propose. Unfortunately, I cannot provide an expansive

---

79 Of course, this terminology is misleading, since the account of 'humanity' I've been sketching claims, precisely, that all humans are necessarily 'interdependent'.

80 The cosmopolitan view I've defended is a version of 'juridical' and 'ethical', but not necessarily 'political' cosmopolitanism, in the senses stipulated by Caney (389).

comparison of approaches here, but must rest content with the bare recognition that substantial similarities exist across the tradition<sup>81</sup>. I will not describe the arguments given by Caney against 'interdependence-based' approaches<sup>82</sup> (391-4), but rather respond to an illuminating challenge that has been offered to 'humanity-based' interpretations by a proponent of an 'interdependence-based' model. Darrel Moellendorff has argued<sup>83</sup> that a person is no more bound by duties of justice 'to intervene into the affairs of [humans] with whom we have no intercourse but only an awareness of their existence...say, [humans] on the second planet orbiting some distant star', than they have duties of justice to intervene in the affairs of 'rocks and plants' (2002: 31). The *prima facie* implausibility of claiming that such duties exist may seem to damage the prospects for a humanity-based cosmopolitanism. However, on my view all that would follow from the fact that such isolated humans did not have equal dignity with us is that this should be regretted, that we would not genuinely have the constitutive goods of justice unless it were remedied, and that we might have reason to perform actions with the potential to bring us closer together in terms of mutual influence over time. Even these things would only follow given certain assumptions: it might be that no possible alteration in the form of human interrelations could bring such causal influence about, and if so justice (for us) would genuinely fail to extend to these particular people. This is not an absurd implication at all, in fact it has considerable plausibility in its own right<sup>84</sup>. I conclude that a humanity-based version of cosmopolitanism, such as the one I've provided, has good prospects.

---

81 For example, as well as much of Nussbaum's work which has already been cited, and that of Caney himself, see especially the argument of Barry 1989, 238-41.

82 But cf. fn.8 above, which gives a schematic argument which has some similarities to those which Caney provides.

83 This is charitable. Really, he simply describes a possible implication of an alternative view in such a way as to invite the interpretation that it is absurd.

84 I have slightly modified Moellendorff's argument: his wording refers to 'intelligent beings'; I have changed this to 'humans' so that I would more clearly be targeted by it. Cf. Caney, 397-9.

## §4.2 Drawing the Threads Together

I began this Thesis with the aspiration to provide a clarified, better defended framework for a capability approach to global justice. In each Chapter, I have progressed towards this aim by integrating further elements into this structure.

After introducing the historical debates around justice in Chapter 1, I moved on to discuss several problematic aspects of Nussbaum's development. I argued that a plausible version of the capability approach must deploy what I called 'strong capability', as part of a bundle including both capabilities and other valuable functionings. By the end of the first half of Chapter 2, I had concluded that although Nussbaum's sufficientarian deployment of the notion of 'dignity' could be used to do important work, the domain of justice needed to be given a more expansive interpretation than Nussbaum has offered. In the second half of that Chapter, I introduced a distinction between 'constructivist' and 'realist' understandings of methodology, and argued, both by identifying unclarities in Nussbaum's account, and then by taking up a critique by Alison Jaggar, that more needed to be done to preserve a realist understanding of methodology.

In Chapter 3, I deepened my critical focus on the relative merits of constructivism and realism, and supplied a number of motivations to seek what I called a 'culture-laden' realist understanding of ethical epistemology. I argued that constructivism had little prospect of satisfyingly accounting for the phenomenon of adaptive preferences (§1), and that it had no significant advantages, and actually several disadvantages, when attempting to account for the role of first-order values within ethical epistemology (§2). I also (§2.5) focussed more

tightly on John Rawls' 'political constructivism', arguing that this paradigm of the type does not hold much promise for the development of a solid methodology. I concluded that we have reason to seek a well-founded culture-laden realist orientation, since that stood the best chance of cohering with the insights from feminist epistemology that I had introduced throughout the Chapter.

In Chapter 4, I presented just such an orientation, drawing heavily upon my understanding of the metaethics of John McDowell to provide a philosophical background capable of supporting culture-laden realism. I argued that his metaethical stance – perhaps uniquely – fits perfectly with the requirements of ethical epistemology (§1). Then, in §2, I noted further connections between his ideas and feminist-epistemic claims, in particular by providing an account of two different kinds of practical reasons. The anti-reductionist flexibility of this metaethical perspective has formed the background for everything that has followed.

In Chapter 5, I offered a reinterpretation of the recently revived tradition of neo-Aristotelian naturalism, which offers a promising means of making sense of the basis for the ethical universalism that is presupposed by the capability approach. Firstly, I gave a sketch of the way in which capability-theoretic justice can be placed within an Aristotelian account. Then, I argued such accounts should not, on pain of serious implausibility, be interpreted as attempts at an independent grounding for ethical claims, but as an application of ethical thought to the domain of species and their natures. I closed with a clarifying discussion of the relation between virtue – as I'd already placed it within a McDowellian frame – and the human good.

Finally, in this Chapter I have returned to the question of the roles that a capability approach

is suited to play within our thought about, and practice of, global justice. After giving a more detailed description of the core Aristotelian understanding of justice I favour, I noted a major threat to it, in the form of distinctively *structural* injustice. In the ensuing discussion, I have distinguished between two kinds of political-philosophical theory, 'ideal' and 'non-ideal', and argued that sufficientarian capability approaches are best placed to play the former, rather than the latter, role.

The capability approach has, since its inception, been the site of wide-ranging philosophical controversy. In this Thesis, I have moved through several areas of this controversy, developing a consistently Aristotelian account of methodology and epistemology, of the distinctive first-order contribution that capability approaches make, and of the kinds of theoretical aspiration that are appropriate to an approach with this nature. I conclude that, if it is interpreted in the way that I have urged, the capability approach is more philosophically promising than ever.

# Appendix A:

## Nussbaum's List of Essential Capabilities

- *Life*. Being able to live to the end of a human life of normal length; not dying prematurely, or before one's life is so reduced as to be not worth living.
- *Bodily Health*. Being able to have good health, including reproductive health; to be adequately nourished; to have adequate shelter.
- *Bodily Integrity*. Being able to move freely from place to place; to be secure against violent assault, including sexual assault and domestic violence; having opportunities for sexual satisfaction and for choice in matters of reproduction.
- *Senses, Imagination, and Thought*. Being able to use the senses, to imagine, think, and reason—and to do these things in a "truly human" way, a way informed and cultivated by an adequate education, including, but by no means limited to, literacy and basic mathematical and scientific training. Being able to use imagination and thought in connection with experiencing and producing works and events of one's own choice, religious, literary, musical, and so forth. Being able to use one's mind in ways protected by guarantees of freedom of expression with respect to both political and artistic speech, and freedom of religious exercise. Being able to have pleasurable experiences and to avoid non-beneficial pain.
- *Emotions*. Being able to have attachments to things and people outside ourselves; to love those who love and care for us, to grieve at their absence; in general, to love, to grieve, to experience longing, gratitude, and justified anger. Not having one's emotional development blighted by fear and anxiety. (Supporting this capability means supporting forms of human association that can be shown to be crucial in their development.)
- *Practical Reason*. Being able to form a conception of the good and to engage in critical reflection about the planning of one's life. (This entails protection for the liberty of conscience and religious observance.)
- *Affiliation*.
  - Being able to live with and toward others, to recognize and show concern for other

humans, to engage in various forms of social interaction; to be able to imagine the situation of another. (Protecting this capability means protecting institutions that constitute and nourish such forms of affiliation, and also protecting the freedom of assembly and political speech.)

- Having the social bases of self-respect and non-humiliation; being able to be treated as a dignified being whose worth is equal to that of others. This entails provisions of non-discrimination on the basis of race, sex, sexual orientation, ethnicity, caste, religion, national origin and species.
- *Other Species*. Being able to live with concern for and in relation to animals, plants, and the world of nature.
- *Play*. Being able to laugh, to play, to enjoy recreational activities.
- *Control over one's Environment*.
  - *Political*. Being able to participate effectively in political choices that govern one's life; having the right of political participation, protections of free speech and association.
  - *Material*. Being able to hold property (both land and movable goods), and having property rights on an equal basis with others; having the right to seek employment on an equal basis with others; having the freedom from unwarranted search and seizure. In work, being able to work as a human being, exercising practical reason and entering into meaningful relationships of mutual recognition with other workers.

(Nussbaum 2006, 76-8)

## Appendix B:

### Neo-Aristotelian *Naturalism*?

In Chapter 5, my central conclusion was that the only plausible interpretation of neo-Aristotelian naturalist theses views them as moves within ethical discourse, broadly construed, rather than as claims about an extra-ethical domain of reality. The straightforward consequence of this is that the meta-level status of such naturalistic normative philosophy cannot differ from that of ethics more generally. In this Appendix, I build on this claim, with two main aims. Firstly, I shall complete the task of incorporating the basic neo-Aristotelian account of the human life-form and its good within the McDowellian metaethical framework set out in Chapter 4. Secondly, I shall provide a defence of my account against what may be the most immediately pressing meta-level objection to it: the objection that it is not really a form of naturalism at all, and thus falls foul of standard naturalist objections.

The Appendix comprises five sections. In §1, I make explicit the meta-level status of neo-Aristotelian naturalism as I understand it, thus largely fulfilling the first of the above aims in short order. §2, 3, and 4 then provide the bulk of the discussion about naturalism. §2 sets out a distinction between two sorts of naturalist challenge – methodological and metaphysical – and argues that the first kind of critique does not pose a major threat to the view I propose. §3 and §4 then consider challenges based upon the metaphysical variety of naturalism. In §3, I bring out a dilemma for neo-Aristotelian accounts, and conclude that my version sides firmly with one horn; this renders worries about metaphysical naturalism more pressing for me than they might otherwise be. In §4, I examine McDowell's 'naturalism of second nature', evaluating its worth as a means of defence here, and concluding that it stands or falls with McDowellian answers to broader metaphilosophical questions. §5 then briefly considers those answers, and argues for a minimal conclusion: that, although the questions at hand are far beyond the possible scope of this thesis (or any other), it is far from clear that the perspective that motivates reductive naturalism should win out. I conclude, then, that the meta-level framework for a capability approach that I propose is secure enough to be effective.

## §1 What are the Meta-level claims?

In Chapter 5, I discussed a number of difficult issues of interpretation that are implicated by some recent varieties of neo-Aristotelian naturalism. I concluded that the best hope of a plausible neo-Aristotelian framework lies with a robust non-reductive position. However, this might appear to some to threaten the realist aspirations of such positions. This is a particularly pressing problem because both Foot and Hursthouse have indicated – to some extent at least<sup>1</sup> – that they think that one important role of neo-Aristotelian theses is to support realism in this area. My response to this is predictably brief, given the arguments of Chapter 4: there is no clear connection between realist status and non-ethical domains of thought, such as those of the natural sciences, or the typical (perhaps stereotypical) features of those domains.<sup>2</sup> There is no conflict between realist status and – for example – being emotion-involving, or being linked to agency by functionally broadcasting intentions, or being dependent on specific patterns of enculturation. So the apparent tension between preserving the realist appearances of ethical thought and acknowledging the role of these features within it was never genuine in the first place. If this was the only motivation for deploying neo-Aristotelian naturalist concepts, they would be of no value. If this confused motivation is dispensed with, then, what are the correct things to say about the meta-level status of life-form judgements and the connected judgements about the human good?

Applications of life-form concepts are cognitive, aiming at worldly facts, and having roles in the ordinary logical and pragmatic relations of implication, probification, contradiction etc. that discourse and practice involves. However, they are also norm-involving: to say (with adequate understanding) that a particular organism is 'human'<sup>3</sup> is to commit to responding in

---

1 See esp. §§2.3.1 and 2.3.2 in Chapter 5.

2 In any case, there are other ways that a version of neo-Aristotelian naturalism might be used to do pro-realist meta-level work. In particular, even insistently non-reductive versions of neo-Aristotelian Naturalism can help to support 'companions in guilt' argumentation. To wit: if Thompson is right in claiming that natural-goodness concepts are deeply implicit in ordinary thought, discourse, and practice, and especially if he is right that we cannot make sense of everyday concepts of life, species, etc. without them, then a powerful companions in guilt argument will be available here. (However, see §2.1 below on this point.) Briefly, if ethical claims are rendered suspect (probably, by anti-realist argumentation) their loss might be a cost that some would be prepared to bear. Once the deep connection between ethics and natural-goodness concepts is revealed, however, the cost increases greatly. Interestingly, the availability of this kind of argumentation actually *relies* on the meta-level equivalency of natural-goodness concepts and ethical ones that the reductive sort of argument must contradict.

3 In the relevant sense, that is: neo-Aristotelians need not claim to be providing an exclusive analysis of

paradigmatically norm-related ways to certain facts about the functioning of that organism. If certain functionings are present, that will be judged good, while their absence will be judged bad; if others are present, that will be judged bad, while their absence will be judged good. In this way, the psychology and semantics of life-form judgements – properly construed – resembles that of ethical judgements *per se*. The metaphysics of life-form and 'natural goodness' facts should also be thought of in much the same way that McDowellians think of ethical facts generally. These are *sui generis* facts, supervening on but irreducible to the facts presented by biology, physics, etc. These facts are independent of individual minds, but their metaphysical nature includes a connection to one particular set of human responses – the responses of the virtuous.

What is more, these normative judgements about proper functioning should be expected to be fully bound up with the rest of the array of ethical/political concepts which play a role in our lives. Thus, this assimilation of life-form concepts to the ethical domain extends also to their epistemology. I have suggested in particular (especially in Ch.5 §1.2) that not obviously connected concepts like 'distributive justice' or 'equality of opportunity' must be viewed as parts of the same Neurathian epistemic web. Arguments based upon new evidence from (e.g.) evolutionary biology will be, in principle, no more and no less valid in altering conceptions of the human life-form than arguments based on (e.g.) considerations of institutional equality. Finally, we should also expect the general dependence of ethical cognition on enculturation to be replicated here. As a result, different social/cultural locations may be more conducive to understanding different aspects of the human life-form, and some such locations might overall be better suited to appreciating the human life-form than others. Regardless of this, facts about humanity are facts concerning all humans, regardless of the cultural locations they contingently have.

Hopefully, enough has already been said to suggest that these claims are coherent: the naturalist and McDowellian realist elements do not contain any obvious tensions. Furthermore, I hope that they can be seen to exhibit a harmony with one another that goes beyond mere compatibility; I aim to build upon this sense of harmony in the remainder of the

---

ordinary word-use. Depending on their response to the dilemma posed in §3 below, indeed, they might not want to claim to be saying very much about ordinary language at all. Cf. Ch.4 fn.36 here.

thesis. In Chapter 6, I do this by returning to the issue of justice with which I began, and integrating the neo-Aristotelian basis with a capability approach more explicitly. The rest of this Appendix, though, will focus on an important challenge to this set of claims, which threatens its plausibility, if not its coherence.

## §2 Living up to the Label

I have, throughout Chapters 5 and 6, used the term 'naturalism' in the name I give to the view I have been developing. To some extent, this word-choice would make sense on historical grounds alone, since the Aristotelian tradition that my view belongs to has consistently been referred to as a form of naturalism in recent literature. However, one might hope that it is also defensible as a *description* of my view, highlighting one of its important philosophical characteristics, and thus aligning it with some tendencies in philosophical thought, and distancing it from others. The question of what makes a view count as a form of naturalism has been fraught, however, especially over the last few decades since naturalism has become a predominant philosophical stance, and so it is to be expected that one major variety of criticism that my proposal will face will be putatively internecine: 'naturalist' against 'naturalist'.

This is compounded by the fact that the McDowellian inspiration for much of what I've said comes with relevant baggage; some philosophers who strongly identify as naturalists have claimed that McDowell's thought is, if anything, determinedly anti-naturalistic<sup>4</sup>, and McDowell has endeavoured to defend himself against these charges. There are two broad ways of thinking about 'naturalism'; one concerns methodology, and the other metaphysics<sup>5</sup>. I'll now provide preliminary characterisations of each in turn.

Methodological naturalism can be summarised as the following thesis. Philosophy, as a

---

4 A good example is Fodor 1995.

5 This taxonomy is used by Miller 2003 (he calls metaphysical naturalism 'substantive naturalism'), amongst others. In a number of places (e.g. 178-84), he cites Railton, who disavows any deep metaphysical commitment of this kind, claiming that his explicitly reductive metaethical naturalism relies on methodological considerations alone. (One general worry is that methodological naturalism might undermine metaphysical naturalism, insofar as the former might be supposed to involve a disavowal of any irrevisable empirical commitments, which the latter seems to involve.) This manoeuvre might appear to make views like his immune to the considerations I draw upon later. I can't give a response to this here, for reasons of space.

distinctive set of questions, is methodologically continuous with the sciences, such that it possesses no significant unique methods, and such that evidence from those sciences will in principle be relevant in answering philosophical questions<sup>6</sup>. There are three potential threats arising from adherents of this sort of naturalism<sup>7</sup>. Firstly, the view I've presented requires that the domain of ethical inquiry be *sui generis* relative to all other domains of inquiry, and that the involvement of full-fledged ethical concepts is necessary to even begin such inquiry. This demand might appear to violate methodological naturalism, because there is a distinctive constraint in operation here that is not in operation for the natural sciences<sup>8</sup>. This is an illusion, however. The constraint is only on which *specific* sets of concepts can enter into inquiry on particular questions; it does not amount to a fundamentally distinct kind of method. In principle, ethical enquiry could still be as fundamentally empirical as the sciences are<sup>9</sup>. I contend that, if people are suspicious of this response, it can only be because their naturalism is tacitly metaphysical, and not just methodological: it is characteristic of metaphysical naturalism to insist that certain 'natural' sets of concepts are privileged and must be at the centre of every theoretical story.

Secondly, there is an additional threat insofar as ethical concepts arguably<sup>10</sup> bring in additional mental faculties, beyond the minimal cognitive set that is involved in thought as such, and/or play different roles in individual psychology or social practice. It is less clear

---

6 This can take two forms. Either the proponent can regard this as compatible with traditional philosophical inquiry, but just hold that some past philosophers have misdescribed their own enterprise, or they can regard this as a radically revisionary discovery, replacing the methods of philosophy with alien imports from the sciences. Either way, philosophy done, and conceived of, correctly will be methodologically continuous with science.

7 There are other threats falling within this category that are merely apparent, but which might have arisen from too close an association between my view and that of other neo-Aristotelian naturalists. Thompson, for example, has claimed that the epistemology of life-form concepts must be distinctively non-empirical, since they involve deploying a distinct logical category, rather than engaging in ordinary acts of predication. See Thompson 2004a. As my acceptance of epistemological naturalism indicates, this is not something I need to claim. However fraught the status of logical categories within thoroughgoing empiricism will be, it should be clear that these must lie near the 'core' of any constructive-empiricist web, and are not readily revisable in practice.

8 Compare, at this point, distinctions between different sciences. If (as is currently popular – see Cat 2013) we take the view that not all sciences are reducible to physics, this kind of epistemological constraint will also operate within natural science itself. A powerful companions in guilt argument would be made available thereby.

9 Note here that methodological naturalism is just a continuity-claim. It is compatible with an array of views about the proper methodology of natural science, just so long as that methodology is much the same as that for philosophy; in practice, however, it tends to be associated with thoroughgoing Quinean empiricism.

10 I say 'arguably' because I do not want to commit to the claim that non-ethical concepts *cannot* constitutively involve affective/conative elements of psychology; I am not treating it as arguable that ethical concepts do. See Ch.4 §2.1.

that this does not add up to a methodological difference. However, this problem can be avoided by noting that, strictly speaking, it is only cognitive content that plays a direct role in ethical epistemology, even on a McDowellian picture. The point here is not that the non-cognitive/social-pragmatic elements can be stripped off from an independent cognitive remainder; clearly it would violate shapelessness to suppose this could be done (cf. Ch.4 §2.1). But we do not need to be able to dissect a concept from the outside to see it as playing some of the roles it plays in virtue of some of the kinds of faculty it involves and not others. Indeed, this is presupposed when we note, for example, that the emotional content of ethical concepts has a special connection to agential motivation. In this case, the cognitive aspect of ethical concepts is what does the work of putting them into Neurathian contact with other concepts, and thus enables epistemology to occur<sup>11</sup>. So, as above, there is not any clear discontinuity between ethical methodology and natural-scientific methodology, apart from differences in the content of ethical and scientific concepts, to which only a metaphysical naturalist can object.

The previous two threats targeted the claimed epistemic autonomy of ethics. However, a third potential problem concerns the methods used within metaethics itself, with which that autonomy claim is supported. Some methodological naturalists will insist that even with reference to meta-level questions, the methods used in the natural sciences are appropriate. This threat can be dealt with by drawing attention to the quietist nature of McDowellian metaethics. On this view, the 'questions' with which much of metaethics is concerned are not genuine questions at all; as such, they do not require substantive answers. Instead, the relevant parts of metaethics should be regarded as mired in confusion, requiring therapeutic rather than substantive treatment. I cannot go into further detail here<sup>12</sup>, but any version of quietism worthy of the name will hold that the areas of discourse it targets do not involve genuine questions. Since methodological naturalism is a claim about the right way to *answer philosophical questions*, it appears to be compatible with quietism. The combination of a quietist anti-substantialism for many of the traditional 'questions' of philosophy, and

---

11 This is not to say, contra the claims I made in Ch.4, §2.1, that facts about an agent's affective states, or their position in a particular social practice, cannot play an evidential role in ethical practice; they can, since given the inseparable connection between these aspects and the cognitive content, an agent without e.g. proper affective responses will thereby not possess the cognitive content either (they will not be relevantly virtuous). This could make a difference, i.e. to the evidential status of their testimony. The point is just that in every case the presence of cognitive content is required for a concept to enter into epistemic relations.

12 See Ch.4 §1.1, and §5 below on this topic.

(roughly) a substantive empiricism for everything else, seems to be entirely coherent.<sup>13</sup>

The discussion above suggests that methodological naturalism is not a significant threat to a McDowellian view like mine. The other form of naturalism, on the other hand, is far more challenging. As I have said, this second type can be called 'metaphysical naturalism'. Its defining characteristic is an insistence that everything in reality has to be related to some 'natural' kind of entities, modelled on – if not identical to – those which populate natural-scientific theories. Corrolarily, if it is thought that ethical facts/concepts are not 'natural' in their own right, then they must be satisfactorily related to non-ethical ones, or some subset thereof<sup>14</sup>. The kinds of view in metaethics that would be most clearly compatible with metaphysical naturalism, then, would be either eliminativisms about ethics, or views that reduced ethical facts/concepts to non-ethical ones, such that they could then be seen as identical. As I have explained in Ch.4 §1.3 however, McDowell holds that although there will be some determinate set of non-ethical features upon which each ethical fact supervenes, it does not make sense to regard the non-ethical features and the ethical facts as identical, since there is no prospect of reducing ethics to something that can be independently understood, because we cannot even recognise the ethical facts' existence in the absence of an acquired ethical sensibility. If the non-ethical base facts – but not the supervening ethical ones – are considered to be 'natural', the absence of any constructive theoretical account of how these levels are related will make this view a kind of metaphysical non-naturalism. Unlike the methodological kind, I do not think that the McDowellian view I've proposed can readily be shown to comply with the constraint of metaphysical naturalism, and thereby avoid naturalist attacks of this sort. Because his view is so closely linked with anti-reductionist realism, there is little prospect of satisfying metaphysical naturalists on their own terms. Instead, the strategy McDowell uses, and which I'll broadly follow, aims to remove the motivation for thinking of ethical facts/properties as 'non-natural' from the start, and so support the conclusion that, if metaphysical naturalism comes to anything at all, a realist ethics will fall

---

13 It might even be possible to reconceive of quietism within the Quinean frame, by treating it as an attempt to resolve confusions whereby 'core' theses are thought of as more peripheral than they really are. (Cf. Ch.3 §2.4 above.) If this is possible, even McDowellian metaethics would be methodologically continuous with natural science. I won't pursue this line of thought further here, however.

14 It is unclear whether only clearly natural-scientific concepts will suffice, or whether it is enough simply for non-ethical ones to be involved. The latter option seems to be more commonly accepted, although it is unclear why this is so, given naturalism's motivational origins. This is a moot point when it comes to McDowell, however, since his strong anti-reductionism would have significant trouble satisfying either requirement.

within it by default.

For the rest of this Appendix, then, I shall be concerned with metaphysical naturalism, the threat it poses, and McDowell's attempt to defuse that threat. In the next section, I consider a tension that arises for versions of neo-Aristotelian naturalism concerning the extent to which life-form judgements and judgements of natural good can be closely tied in to broader conceptual domains – both pre-theoretical and natural-scientific – with attention to the possibility that placing them in this wider context might help to assuage metaphysical-naturalist concerns.

### §3 Anti-Reductionism vs. Pre-Theoretic Ubiquity

Neo-Aristotelian naturalists commonly emphasise that the conceptual domains that they describe are ubiquitous in ordinary thought, and, *prima facie*, also play a significant role in biological theorising. Examples of this abound in the second chapter of Foot's *Natural Goodness* (2001, 25-37), as well as throughout Thompson's work. If the distinctive characteristic of metaphysical naturalism is the attempt to demonstrate continuity between (e.g.) ethics and (e.g.) biology not merely in terms of methods, but also in terms of utilised concepts, postulated properties, and so on, this ubiquity thesis will likely reduce the degree to which neo-Aristotelians are vulnerable to metaphysical naturalist critique. This section will firstly flesh out this thought, and then argue that this way of avoiding confrontation with metaphysical naturalism is not available to an anti-reductionist Aristotelian view like mine, because the relevant kind of ubiquity thesis vitiates any plausible ethical role for life-form concepts.

To see the potential for ubiquity claims to support aspirations to metaphysical naturalism, consider the sort of judgements that Thompson focuses on, such as 'the female bobcat gives birth in the spring' (2008, 63). *Prima facie*, this kind of judgement is required to begin biological theorising about how different animals live, and (thereby) about the processes that determine *how* animals live, including over evolutionary time-scales. Famously, Darwin's close observations of different features of animal life, both morphological and ethological, were essential to the development of his theories of evolution. On the other hand, there is considerable room for doubt about how deeply embedded these sorts of judgements are in

contemporary biology<sup>15</sup>, since cladistic models of species have largely displaced morphological ones, amongst other things.<sup>16</sup> This might well reduce the extent to which neo-Aristotelians can claim to be making proposals that are continuous with the current biological mainstream, but ubiquity claims are somewhat likely to improve their standing vis-a-vis metaphysical naturalism, nevertheless.<sup>17</sup>

This does not show that a view like mine is on similarly firmer ground, however, because there are good reasons to think that an ethically-laden conception of life-form like mine cannot help itself to this ubiquity thesis. Some points developed by Elijah Millgram illustrate why. Part of his argument focusses on what he calls 'the method of the nature documentary'. This names the central epistemological practice that pro-ubiquity neo-Aristotelian naturalists would have in mind for thinking about life-forms. It involves, at its base, straightforward observation of animals (including observation of other humans), and deployments of the logical category of life-form to make sense of what are seen to be typical traits.<sup>18</sup> (Millgram 2009). Notably, it is only in virtue of the thought that life-form judgement involves such straightforwardly observational methods that the metaphysical continuity claim has any plausibility: if the account of the logical structure of life-form judgement were uncoupled from activities that natural science probably involves, there would no longer be reason to think that natural scientists (as such) use life-form judgements at all. Both elements are, then, required for the ubiquity thesis to have any prospect of improving neo-Aristotelianism's standing as a metaphysical naturalism.

However, Millgram presents a convincing case that the method of the nature documentary is spectacularly poorly suited to provide for the kind of description of humanity-as-such that

---

15 At least one strand in some anti-realist metaphysical naturalisms seems to be defeated by ubiquity claims, regardless: if life-form concepts have played a role in biological science, even if this were confined to the 19<sup>th</sup> century, it would follow that they are not a metaphysically queer sort of entity, given that 'queerness' in this sense picks out certain kinds of difference from the entities that natural science postulates. Compare Mackie 1977.

16 See FitzPatrick 2000, and Lewens 2012 for argument along these lines.

17 As I noted in the previous section, some kinds of metaphysical naturalism do not require continuity with natural science specifically, but just continuity with non-ethical concepts, properties, etc. If this is so, the potential for ubiquity to support Aristotelianism's metaphysical-naturalist credentials will clearly be much greater. Regardless of which view is in play, the considerations against ubiquity below will undermine this sort of argument.

18 As Millgram emphasises, this does not have to involve a crude empiricism of the kind that Thompson explicitly rejects (2009, 559-60).

might do normative work, in the way that neo-Aristotelians require. His point can be put quite simply: it is extremely implausible that what would result from an observational method such as this would be a conception of the human life-form that would have tolerable ethical implications. He gives compelling examples of this from the work of anthropologists and other scientists, who have argued on such observational grounds that infanticide, rape, and a desire to hierarchically dominate others are all species-normal, at least in certain circumstances (2009, 361-2). If the method of the nature documentary has the ultimate implication that such functionings are good in the humans who instantiate them, it should be rejected completely.

Finally, it is clear that a view like mine will not have such abhorrent implications, but this is just because a view like mine does not involve the method of the nature documentary. Instead, observations of human functioning are undertaken from within a perspective that is always-already ethical to its core. This has the potential cost, however, of precluding the ubiquity thesis from being available for anti-reductionist views<sup>19</sup>. As a straightforward consequence, any help that ubiquity theses would provide in avoiding the threat from metaphysical naturalism will be equally unavailable.

In this section I've shown that claims about the ubiquity of life-form judgement, whatever their plausibility for some forms of neo-Aristotelianism, are of no help to varieties of it which take its ethical role seriously. This reinforces the thought that such anti-reductionist forms of neo-Aristotelian naturalism are seriously divergent from metaphysical naturalism, of the kind I've described, and will face considerable challenge from that quarter, unless there is some other way of ameliorating the threat. The next section, then, will consider McDowell's own, very different, attempt to head off naturalist criticisms, under the banner of a 'naturalism of second nature'.

## §4 On Second Naturalism

---

<sup>19</sup> It is possible that emphasising the ethical dimension here will reduce the scope to claim that lots of everyday thought and talk can be captured by this Aristotelian account, but if metaphysical naturalism aim were disavowed, this would not be very problematic. It is not clear that there is any reason to think that life-form judgements of the relevant type have to be ubiquitous, or that we cannot tolerate a split between one type of life-form judgement and another. At any rate, ubiquity is of little relevance to the central role of life-form judgement in this thesis: providing a sound basis for a capability approach to global justice.

In §2 above, I claimed that the threat of metaphysical naturalism for a McDowellian view like mine arises because no constructive theoretical account is provided of the relationship between ethical facts – and ethical psychological states/processes – and the non-ethical counterparts which appear in (e.g.) biology. Instead, all that is postulated is some species of supervenience, which is very unlikely to satisfy critics who hold that features like normativity, or a connection to emotion and motivation, are 'queer', and need to be demystified through a comprehensive reduction to the entities that already populate natural-scientific theory<sup>20</sup>. This reductionist impetus has two sorts of target in an approach to a region of human cognition such as metaethics. Firstly, it can target the kind of facts that the approach postulates, demanding a reduction of these facts to more 'natural' ones. Alternatively, it can target the psychological processes that the approach identifies as being in play. Of these two strategies, the second has received far more attention, and as a result I'll concentrate on this aspect in what follows.<sup>21</sup> In the rest of this section, then, I'll briefly describe McDowell's notion of 'second nature', which is, precisely, intended to defuse reductionist threats of this second kind. I'll then consider the impact of an important critique of McDowell's tactics here, and conclude that the notion of second nature is of little independent use, apart from the way that it fits with other areas of his thought.

Roughly, the idea of second nature is as follows. Human beings, as animals governed by biological (physical, etc.) laws, have a 'first nature' consisting of the space of functionings that these laws include as possibilities.<sup>22</sup> There will then be explanations for certain classes of events in human life, certain kinds of differences between humans, etc., based upon these

---

20 As I mentioned above, this is not quite right: some metaphysical naturalisms don't obviously involve reduction to natural-scientific properties, and some metaphysical naturalists probably don't think this is required. Nevertheless, they require reduction to non-ethical concepts/properties in a way that shapelessness prevents, and so all such views will be incompatible with McDowellian realism.

21 It remains quite unclear why this is; if people are willing to countenance *sui generis* worldly facts, why are they repulsed by the idea of *sui generis* mental processes? Both supervene in the same way on otherwise-describable parts of the world, and neither are reducible in the same way to that supervenience base; ultimately, cogniser-side shapelessness and cognised-side shapelessness then appear to be on a par. This thought notwithstanding, however, in what follows, I'll concentrate on the cogniser-side processes, following McDowell and his main critics.

22 Relative to the bounds of humanity itself, that is: at least, we are always implicitly committed to some set of states of affairs or patterns of events that would *not* be compatible with counting as human, or continuing to do so (e.g. spontaneously decomposing into a soup of one's component atoms). However compatible such events would be with the laws of nature, they will not be compatible with human first nature *ex hypothesi*. It is possible that a non-Aristotelian conception of life-form is at work here (perhaps a statistical or cladistic one), at least superficially. (It is not plausible however that ultimately all such conceptions derive their power from an implicit Aristotelian conception. Cf. Thompson 2008 (33-48) here.)

laws. For example, certain differences in human genital morphology can be explained by referring, for the most part at least, to facts about chromosomes and the role these play in natural laws. *Prima facie*, however, this 'first' human nature radically explanatorily underdetermines the psychological states and dispositions – including intrinsically other-involving states and dispositions such as those that social practices involve – that actual human beings come to have as a result of their normal<sup>23</sup> process of development and enculturation. Correspondingly, it also underdetermines the ways that actual human beings interact with the rest of the world, including in cognising it, at least where these interactions involve such complex psychological states. Instead, at least much of fully-developed human psychology can be conceived as a kind of 'second nature', the result of processes of formation that carve out particular paths through the space of possibilities set up by first nature, but not further explicable by it alone.

A better grasp on the notion of second nature can be gained by considering the following. One immediate contrary response to the appearance of underdetermination might be to claim that some future science, epistemologically continuous with contemporary non-ethical thought (but, accordingly, epistemologically discontinuous with current ethical thought), would fill in the requisite gaps. This misses the point, however: the issue is not simply that we do not yet have an adequate first-natural understanding of psychology and (especially) the processes of enculturation and cultural change that it is bound up with, however transparent the inadequacy of current theory in this area may be. The issue is also that it does not appear that increasing knowledge and explanatory power of *these* kinds does anything at all to preserve the distinctive features of (e.g.<sup>24</sup>) ethical thought and practice. To a large extent, to assert this is simply to reiterate that the characteristics in question (most importantly normativity) only appear from *within* a developed sensibility of the relevant kind. Further study of the supervened-upon (first-natural) level is possible, and will very probably be fruitful and important in its own terms. But it will never generate a faithful account *of* the

---

23 This could be read either as a statistical claim, or as a neo-Aristotelian-style norm-involving claim. On either interpretation it is extremely plausible.

24 This is actually a radically general point: the features that belong essentially to second nature and the points of view it creates are implicated far more widely than in ethics alone. Normativity, for example, is certainly involved in epistemology, and the massive array of social practices which involve it; it may also be involved in semantic meaning itself. I do not strictly have to take stands on these issues here, but the extent to which a comparably anti-reductionist stance is mandatory in those other areas too will be the extent to which there will be a powerful companions in guilt argument for McDowellian metaethical realism in particular.

second-natural dispositions (their phenomenology, pragmatic roles, etc.), without using information generated by those dispositions themselves. Thus, elements of human psychology whose *explananda* involve the relevant sort of feature (normativity, etc.) will not be fully explicable using first-natural *explanantia* alone. Instead, these elements can only be explained as parts of an acquired 'second nature'; the events and processes that they cause and are caused by are all themselves *sui generis*, and so no more comprehensive program of explanation will be available.

This approach looks initially promising. 'Second nature' seems to be no more or less a variety of human nature than 'first nature' – in particular, it (or rather its specific components) seems to play the same role in explanations of different human functionings that 'first nature' does – and resistance to this claim appears to rest (question-beggingly) on metaphysical naturalism itself, since the *sui generis* claim appears to be a straightforward corollary of anti-reductionism. However, there are a number of important criticisms of the notion of second nature that cast doubt on its utility here. For my purposes here, however, I'll only consider one: that of David Forman (2008). Forman begins by noting that McDowell's concept of 'second nature' involves an analogy with the case of language-learning:

The basic idea is that the initiation into the practices of a linguistic tradition is something that is clearly part of the normal and therefore natural maturation of human beings even though the norms embodied in that tradition cannot somehow be constructed out of elements that are merely natural in the relevant sense: the norms embodied in that tradition are not merely natural since they are not intelligible with reference to nature considered apart from that tradition. (567)

This is essentially the same as the explication I've just given. However, Forman makes two important critical points.

Firstly, there is a sense in which an Aristotelian concept of 'second nature' cannot serve a helpful dialectical purpose in discourse with someone – a metaphysical naturalist – who doubts whether the components or features of non-reductive ethics are natural. This is because the relevant notion of 'natural' seems to mean something like 'not subject to interruption by acts of mind': on the Aristotelian account, the understanding of (e.g.) the concepts that are involved in virtue is second nature because the comprehending agent no longer needs to deliberate afresh each time to come up with an answer in accord with the

norm (Forman 2008, 573).<sup>25</sup> Thus, second nature simply refers to acquired instincts, rather than innate ones. However, someone who is skeptical (e.g.) of the natural position of normativity as such will be unimpressed by this: norms when adequately internalised become *like* natural entities, in this respect at least. But if normativity<sup>26</sup> as such is queer, this will not provide good reason to think that they become (or are revealed to always have been) natural *tout court*, since the (queer) normativity remains untouched by this. This consideration suggests that dialectically speaking, a 'naturalism of second nature' can do no independent work; the only people who might find it reassuring will be those who are already comfortable with the idea that normativity *just is* a feature of the natural world.

Secondly, Forman notes that there is a difficulty in using the notion of habituation in talking about *sui generis* conceptual contents, and expecting this to form an intelligible bridge between one uncontroversial region of nature, and a second more controversial one. This is because in order for habituation to lead someone to possess a fuller, and more thoroughly integrated, disposition to instantiate some functioning it already has to be possible to instantiate the functioning in a primitive form. But:

[i]f we want to say that the most basic conceptual abilities we draw upon in perception are acquired by means of developing a kind of habitual second nature, then we will have to say that such perceptual habits differ from Aristotle's habits in one key respect: the perceptual habit and the corresponding conceptual ability would have to be considered coeval, e.g., the habit of responding to red things in a certain conceptually informed way would have to be considered coeval with the ability to see something as red in a particular case. For if we possess the conceptual ability before we have the habit, then the habit cannot itself be what distinguishes a conceptually informed perception from a merely natural response to the environment. And, in that case, the account of a habitual second nature has no role to play in making us comfortable with the idea that we can acquire conceptual abilities that are responsive to an autonomous space of reasons. (575.)

That is, when faced with someone who doubts whether the stuff of concepts (i.e. placement in a normative nexus) is part of nature as such, it can't help to appeal to the way that habit integrates concepts into human life in much the same way that (e.g.) motor response to a pin pricking a finger is integrated, because, on the face of it, habits can only be formed from ingredients that are already present. If normativity is not natural – not already

---

25 Compare Pippin 2002 and Wright 1996, on this point.

26 Forman consistently talks of autonomy at these junctures. Although this accords with some of McDowell's Kantian talk, I suspect he reads too much (or the wrong) Kant into the concepts, such that spontaneity and norm-responsiveness begin to appear like separate features.

unproblematically present, at least as a potentiality waiting only on the right kind of responder to reveal itself – habituation cannot make it so; if normativity *is* natural in this sort of way, habituation is not needed<sup>27</sup> to tell us how it comes to play a role in human life.

These two considerations illustrate the extent to which McDowell's 'naturalism of second nature' is quite limited in its efficacy at defusing the threat from metaphysical naturalism. Indeed, it begins to seem that this should never have been surprising: nothing can defuse this threat because metaphysical naturalism is simply brutally different from other philosophical stances. As such, perhaps the reductionist impulse cannot be accommodated, but only refused. In the next section, I'll turn more broadly to some metaphilosophical questions that bear on this issue, and argue that McDowell's alternative to metaphysical naturalism is compelling, but only as part of his overall perspective. Neo-Aristotelian naturalism, then, cannot be defended by taking metaphysical naturalist motivations for granted.

## §5 Quietism and the Point of Metaethics Redux

This section, finally, concludes this Appendix by re-emphasising the metaphilosophical origins of naturalist worries. Because the role of metaethical (including the relevant metaphilosophical) claims in this Thesis is strongly limited<sup>28</sup>, I have little scope for an expansive discussion of these (sometimes quite obscure) topics. However, I can do a little more to link the impasse reached in the previous section to McDowell's broader philosophy.

In this Appendix, I have suggested that there is a species of naturalism, 'metaphysical' naturalism, which cannot be productively engaged with by views like McDowell's, because it begins from entirely incompatible metaphilosophical assumptions. These generate an all-or-nothing methodology which insists upon reduction to a privileged core of metaphysical posits. Because McDowell's view is distinguished by its distinctively anti-reductionist

---

27 This is not to say that habituation isn't a useful, or even essential, concept for other purposes. It simply has no dialectical efficacy here.

28 Throughout the Thesis, there are many points at which I have implicated questions to which I have either not given answers, given very vague answers, or given answers without completely adequate support. However, this reflects the upshot of my discussion in Chapter 4 (§1.4): that many of the issues involved in the background to theories of global justice such as Nussbaum's capability approach, although they may appear to be traditionally meta-level problems, cannot be resolved in abstraction from the sort of complex, substantive, and controversial first-order ethical/political thought and debate that is necessarily beyond the scope of my project here.

orientation, no real argument seems to be possible in either direction: the metaphysical naturalist insists that without reduction to the privileged class nothing is explained; the anti-reductionist insists that the domain is simply irreducible, and that questions which cannot be answered are not well-formed questions at all.

This conclusion invites a brief return to the idea of quietism. I have already, in Chapter 4 (§1.4), argued that quietism about meta-level philosophical topics does not imply that philosophy cannot perform any useful tasks, or reveal any substantial truths, in the connected first-order domains. What is more, the articulation of a coherent and helpful non-reductionist meta-ethical perspective might play the 'therapeutic' role that Wittgenstein urged for philosophy, removing the impulse to demand reductive theories at its root. What is left will be a species of metaethical framework which proceeds in harmony with first-order ethics, rather than deploying a contentious conception of philosophical method to force ethical conclusions from above.

Ultimately, I suggest that if there is a way of resolving these issues more conclusively, it will not be separable from more broadly metaphilosophical debates, which I can't address in detail. I conclude that although there are many extremely difficult and all-encompassing issues in play in metaethical debates, the challenge from metaphysical naturalism is not powerful enough to discredit a neo-Aristotelian approach, and so its role in an account of objectivity in judgement about universalist justice remains tenable.

# Bibliography

- Ackerly, Brooke. 2000. *Political Theory and Feminist Social Criticism*. Cambridge: Cambridge University Press.
- Anderson, Elizabeth. 1995a. 'Knowledge, Human Interests, and Objectivity in Feminist Epistemology' in *Philosophical Topics* Vol.23 No.2: 27-58.
- Anderson, Elizabeth. 1995b. 'Feminist Epistemology: An Interpretation and Defence' in *Hypatia* Vol.10 No.3: 50-84.
- Anderson, Elizabeth. 2003. 'Sen, Ethics, and Democracy' in *Feminist Economics* Vol.9 No.2-3: 239-61.
- Anderson, Elizabeth. 2010. 'Justifying the Capability Approach to Justice' in *Measuring Justice: Primary Goods and Capabilities*, Brighouse, Harry, and Robeyns, Ingrid (eds.). Cambridge: Cambridge University Press. 81-100.
- Anderson, Elizabeth. 2011. 'Feminist Epistemology and Philosophy of Science' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/fall2012/entries/feminism-epistemology>.
- Anderson, Elizabeth. 2014. 'Dewey's Moral Philosophy' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/spr2014/entries/dewey-moral/>.
- Andreou, Chrisoula. 2006. 'Getting on in a Varied World' in *Social Theory and Practice* Vol.32: 61-73.
- Anscombe, G.E.M. 1958. 'Modern Moral Philosophy' in *Philosophy* Vol.33 No.124: 1-19.
- Antony, Louise. 2000. 'Natures and Norms' in *Ethics* Vol.111 No.1: 8-36.
- Antony, Louise. 2008. 'Quine as Feminist: The Radical Import of Naturalised Epistemology' in *Epistemology: An Anthology*, Ernest Sosa, Jaegwon Kim, Jeremy Fantl, and Matthew McGrath (eds.). Oxford: Blackwell. 552-84.
- Antony, Louise, and Witt, Charlotte (eds.). 2002. *A Mind of One's Own*. Boulder, CO: Westview.
- Aristotle. 1998. *Politics*, C.D.C. Reeve (trans.). Indianapolis: Hackett.
- Arneson, Richard. 2000. 'Perfectionism and Politics' in *Ethics* Vol.111 No.1: 37-63.
- Arneson, Richard. 2010. 'Two Cheers for Capabilities' in *Measuring Justice: Primary Goods and Capabilities*, Brighouse, Harry, and Robeyns, Ingrid (eds.). Cambridge: Cambridge University Press. 101-28.
- Audi, Robert. 2008. 'Ethical Generality and Moral Judgement' in *Challenging Moral Particularism*, Mark Lance, Matjaz Portc, Vojko Strahovnik (eds.). London: Routledge. 31-52.
- Baber, H.E. 2010. 'Worlds, Capabilities, and Well-Being' in *Ethical Theory and Moral Practice* Vol.13: 377-92.
- Barry, Brian. 1989. *Theories of Justice: A Treatise on Social Justice*, Vol. 1. London: Harvester Wheatsheaf.
- Barry, Christian, and Ferracioli, Laura. 2013. 'Young on Responsibility and Structural Injustice' in *Criminal Justice Ethics* Vol.32 No.3: 247-57.

- Bell, Daniel. 2012. 'Communitarianism' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/fall2013/entries/communitarianism>.
- Benhabib, Seyla. 1993. 'Feminist Theory and Hannah Arendt's Concept of Public Space' in *History of the Human Sciences* Vol.6 No.2: 97-114.
- Benhabib, Seyla. 2006. 'The "Claims" of Culture Properly Interpreted: Response to Nicholas Kompridis' in *Political Theory* Vol.34 No.3: 383-8.
- Berg, Amy. 2015. *Studies in Ideal and Non-Ideal Theory*. (Forthcoming).
- Besch, Thomas M. 2013. 'On Political Legitimacy, Reasonableness, and Perfectionism' in *Public Reason* Vol.5 No.1: 58-74.
- Blackburn, Simon. 1981. 'Rule-Following and Moral Realism' in *Wittgenstein: To Follow a Rule*, Steven Holtzman and Christopher Leich (eds.). London: Routledge and Kegan Paul: 163-87.
- Blackburn, Simon. 1984. *Spreading the Word*. Oxford: Oxford University Press.
- Blackburn, Simon. 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- Bohman, James, and Richardson, Henry S. 2009. 'Liberalism, Deliberative Democracy, and "Reasons that all can accept"' in *The Journal of Political Philosophy* Vol.17 No.3: 253-74.
- Boyd, Richard N. 1988. 'How to be a Moral Realist' in *Essays on Moral Realism*, Geoffrey Sayre-McCord (ed.). Ithaca, NY: Cornell University Press. 181-228.
- Brewer, Talbot. 2009. 'The Foundations of Neo-Aristotelianism: Critical Notice of Michael Thompson, Life and Action' in *Philosophical Books* Vol.50 No.4: 197-212.
- Brighouse, Harry, and Robeyns, Ingrid (eds.). 2010. *Measuring Justice: Primary Goods and Capabilities*. Cambridge: Cambridge University Press.
- Brink, David O. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Brooks, Thom. 2014. 'A New Problem with the Capabilities Approach' in *Harvard Review of Philosophy* Vol.20: 100-106.
- Brooks, Thom. 2015. 'The Capabilities Approach and Political Liberalism' in *Rawls' Political Liberalism*, Thom Brooks and Martha C. Nussbaum (eds.). New York: Columbia University Press. 139-174.
- Calder, Gideon. 2008. 'Ethics and Social Ontology' in *Analyse & Kritik* Vol.30: 427-43.
- Calder, Gideon. 2011. 'Climate Change and Normativity: Constructivism versus Realism' in *Critical Review of International Social and Political Philosophy* Vol.14. No.2: 153-69.
- Caney, Simon. 2009. 'Cosmopolitanism and Justice' in *Contemporary Debates in Political Philosophy*, Thomas Christiano and John Christman (eds.). Oxford: Blackwell. 387-407.
- Cat, Jordi. 2013. 'The Unity of Science' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/sum2013/entries/scientific-unity/>.
- Christensen, Anne-Marie S. 2009. 'Getting It Right in Moral Experience: John McDowell and Virtue Ethics' in *The Journal of Value Inquiry* Vol.43: 493-506.
- Claassen, Rutger. 2011. 'Making Capability Lists: Philosophy versus Democracy' in *Political Studies* Vol.59 No.3: 491-508.

- Cohen, Gerald A. 1993. 'Equality of What? on Welfare, Goods, and Capabilities' in *The Quality of Life*, Martha C. Nussbaum and Amartya Sen (eds.). Oxford: Clarendon. 9-29.
- Cohen, Gerald A. 2008. *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press.
- Cohen, Gerald A. 2009. *Why Not Socialism?*. Princeton, NJ: Princeton University Press.
- Crocker, David A. 1992. 'Functioning and Capability: the Foundations of Sen and Nussbaum's Development Ethic' in *Political Theory* Vol.20 No.4.: 584-612.
- Code, Lorraine. 1991. *What Can She Know? Feminist Theory and the Construction of Knowledge*. Ithaca, NY: Cornell University Press.
- Code, Lorraine. 1996. 'What is Natural about Epistemology Naturalised?' in *American Philosophical Quarterly* Vol.33 No.1: 1-22.
- Copp, David, and Sobel, David. 2004. 'Morality and Virtue: An Assessment of some Recent Work in Virtue Ethics' in *Ethics* Vol.114: 514-54.
- Cox, Damian. 2006. 'Agent-Based Theories of Right Action' in *Ethical Theory and Moral Practice* Vol.9: 505-15.
- Cudd, Ann. 2006. *Analysing Oppression*. New York: Oxford University Press.
- Dancy, Jonathan. 2004. *Ethics Without Principles*. Oxford: Oxford University Press.
- D'Arms, Justin, and Jacobson, Daniel. 2000. 'Sentiment and Value' in *Ethics* Vol.110 No.4: 722-48.
- Davidson, Donald. 2001. *Essays on Actions and Events*. Oxford: Clarendon.
- Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.
- Douglas, Heather. 2004. 'The Irreducible Complexity of Objectivity' in *Synthese* Vol.138: 453-73.
- Dreier, James. 2004. 'Meta-Ethics and the Problem of Creeping Minimalism' in *Philosophical Perspectives* Vol.18 No.1: 23-44.
- Driver, Julia. 2004. *Uneasy Virtue*. Cambridge: Cambridge University Press.
- Dummett, Michael. 1991. *The Logical Basis of Metaphysics*. London: Duckworth.
- Dworkin, Ronald. 2000. *Sovereign Virtue. The Theory and Practice of Equality*. Cambridge: Harvard University Press.
- Elster, Jon. 1983. *Sour Grapes*. Cambridge: Cambridge University Press.
- Enoch, David. 2005. 'Why Idealize?' in *Ethics* Vol.115 No.4: 759-787.
- Enoch, David. 2006. 'Agency, Schmagency: Why Normativity Won't Come from What is Constitutive of Agency' in *Philosophical Review* Vol.115: 169-98.
- Enoch, David. 2009. 'Can There Be a Global, Interesting, Coherent Constructivism about Practical Reason?' in *Philosophical Explorations* Vol.13 No.3: 319-39.
- Enoch, David. 2011a. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.
- Enoch, David. 2011b. 'Schmagency Revisited' in *New Waves in Metaethics*, Michael Brady (ed.). Basingstoke: Palgrave Macmillan. 208-33.
- Enoch, David. 2013. 'The Disorder of Public Reason' in *Ethics* Vol.124 No.1: 141-76.
- FitzPatrick, William. 2000. *Teleology and the Norms of Nature*. New York: Garland.
- Fodor, Jerry. 1995. 'Encounters with Trees' in *London Review of Books* Vol.17 No.8: 10-11.

- Foot, Philippa. 2001. *Natural Goodness*. Oxford: Clarendon.
- Foot, Philippa. 2002. 'Moral Arguments' in *Virtues and Vices*. Oxford: Oxford University Press.
- Forman, David. 2008. 'Autonomy as Second Nature: On McDowell's Aristotelian Naturalism' in *Inquiry* Vol.51 No.6: 563-80.
- Fricker, Miranda. 2003. 'Epistemic Injustice and a Role for Virtue in the Politics of Knowing' in *Metaphilosophy* Vol.34 No.1-2: 154-73.
- Fricker, Miranda. 2006. 'Powerlessness and Social Interpretation' in *Episteme* Vol.3 No.1-2: 96-108.
- Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Politics of Knowing*. Oxford: Oxford University Press.
- Fricker, Miranda. 2014. 'What's the Point of Blame? A Paradigm-Based Explanation' in *Noûs* (forthcoming): 1-19.
- Friedman, Marilyn. 2000. 'John Rawls and the Political Coercion of Unreasonable People' in *The Idea of a Political Liberalism: Essays on Rawls*, Victoria Davion and Clark Wolf (eds.). Lanham, MD: Rowman and Littlefield. 16-33.
- Gaus, Gerald F. 1999. 'Reasonable Pluralism and the Domain of the Political: How the Weaknesses of John Rawls' Political Liberalism Can be Overcome by a Justificatory Liberalism' in *Inquiry* Vol.42 No.2: 259-82.
- Gaus, Gerald F. 2012. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.
- Gheaus, Anca. 2013. 'The Feasibility Constraint on the Idea of Justice' in *The Philosophical Quarterly* Vol.63 No.252: 445-64.
- Gilibert, Pablo. 2011. 'Debate: Feasibility and Socialism' in *The Journal of Political Philosophy* Vol.19 No.1: 52-63.
- Gilibert, Pablo, and Lawford-Smith, Holly. 2012. 'Political Feasibility: A Conceptual Exploration' in *Political Studies* Vol.60: 809-25.
- Grasswick, Heidi E. 2004. 'Individuals-in-Communities: The Search for a Feminist Model of Epistemic Subjects' in *Hypatia* Vol.19 No.3: 85-120.
- Grasswick, Heidi E. 2013. 'Feminist Social Epistemology' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/spr2013/entries/feminist-social-epistemology>.
- Guyer, Paul. 1992. 'Introduction: the starry heavens and the moral law' in *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press.
- Haddock, Adrian. 2008. 'McDowell and Idealism' in *Inquiry* Vol.51 No.1: 79-96.
- Haslanger, Sally. 2000. 'Gender and Race: (What) Are They? (What) Do We Want Them to Be?' in *Noûs* Vol.34 No.1: 31-55.
- Held, Virginia. 2002. 'Moral Subjects: The Natural and The Normative' in *Proceedings and Addresses of the American Philosophical Association* Vol.76 No.2: 7-24.
- Held, Virginia. 2006. *The Ethics of Care: Personal, Political, and Global*. Oxford: Oxford University Press.
- Hicks, Daniel. 2011. 'Is Longino's Conception of Objectivity Feminist?' in *Hypatia* Vol.26 No.2: 333-51.
- Hope, Simon. 2013. 'Neo-Aristotelian Social Justice: An Unanswered Question' in *Res Publica* Vol.19: 157-72.

- Horgan, Terrance, and Timmons, Mark. 1996. 'From Moral Realism to Moral Relativism in One Easy Step' in *Critica* Vol.38: 3-39.
- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.
- Intemann, Kristen. 2005. 'Feminism, Underdetermination, and Values in Science' in *Philosophy of Science* Vol.72 No.5: 1001-1012.
- Intemann, Kristen. 2010. '25 Years of Feminist Empiricism and Standpoint Theory: Where Are We Now?' in *Hypatia* Vol.25 No.4: 778-96.
- Jacobson, Daniel. 2005. 'Seeing by Feeling: Virtues, Skills, and Moral Perception' in *Ethical Theory and Moral Practice* Vol.8 No.4: 387-409.
- Jaggar, Alison M. 1989. 'Love and Knowledge: Emotion in Feminist Epistemology' in *Inquiry* Vol.32 No.2: 151-76.
- Jaggar, Alison M. 1991. 'Introduction to Part VI: The Personal as Political' in *Living with Contradictions: Controversies in Feminist Social Ethics*, Alison M. Jaggar (ed.). Boulder, CO: Westview.
- Jaggar, Alison M. 1998. 'Globalizing Feminist Ethics' in *Hypatia* Vol.13 No.2: 7-31.
- Jaggar, Alison M. 1999. 'Multicultural Democracy' in *The Journal of Political Philosophy* Vol.7 No.3: 308-29.
- Jaggar, Alison M. 2000a. 'Feminism in Ethics: Moral Justification' in *The Cambridge Companion to Feminism in Philosophy*, Miranda Fricker and Jennifer Hornsby (eds.). Cambridge: Cambridge University Press. 225-44.
- Jaggar, Alison M. 2000b. 'Ethics Naturalised: Feminism's Contribution to Moral Epistemology' in *Metaphilosophy* Vol.31 No.5: 452-68.
- Jaggar, Alison M. 2004. 'Feminist Politics and Epistemology: the Standpoint of Women' in *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, Sandra Harding (ed.). New York: Routledge. 55-66.
- Jaggar, Alison M. 2006. 'Reasoning about Well-Being: Nussbaum's Methods of Justifying the Capabilities' in *The Journal of Political Philosophy* Vol.14 No.3: 301-22.
- James, Aaron. 2007. 'Constructivism about Practical Reasons' in *Philosophy and Phenomenological Research* Vol.74 No.2: 302-25.
- Johnson, Robert N. 1997. 'Reasons and Advice for the Practically Rational' in *Philosophy and Phenomenological Research* Vol.57 No.3: 619-25.
- Johnson, Robert N. 1999. 'Internal Reasons and the Conditional Fallacy' in *The Philosophical Quarterly* Vol.49 No.109: 53-71.
- Joyce, Richard. 2006. *The Evolution of Morality*. Cambridge, MA: MIT Press.
- Jubb, Robert. 2013. 'Social connection and practice dependence: some recent developments in the social justice literature' in *Critical Review of International Social and Political Philosophy* Vol.16 No.5: 698-713.
- Kalderon, Mark E. 2005. *Moral Fictionalism*. Oxford: Oxford University Press.
- Kawall, Jason. 2004. 'Moral Response-Dependence, Ideal Observers, and the Motive of Duty: Responding to Nick Zangwill' in *Erkenntnis* Vol.60: 357-69.
- Kawall, Jason. 2014. 'Qualified-agent and agent-based virtue ethics and the problems of right action' in *The Handbook of Virtue Ethics*, Stan van Hooft et. al. (eds.). Oxford: Routledge. 130-40.
- Khader, Serene J. 2011. 'Beyond Inadvertent Ventriloquism: Caring Virtues for Anti-Paternalist Development Practice' in *Hypatia* Vol.26 No.4: 742-61.

- Klosko, George. 1997. 'Political Constructivism in Rawls' *Political Liberalism*' in *The American Political Science Review* Vol.91 No.3: 635-46.
- Koppelman, Andrew. 2009. 'The Limits of Constructivism: Can Rawls Condemn Female Genital Mutilation?' in *The Review of Politics* Vol.71 No.3: 459-82.
- Korsgaard, Christine M. 1996. 'From Duty and for the Sake of the Noble: Aristotle and Kant on Morally Good Action' in *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, Stephen Engstrom and Jennifer Whiting (eds.). Cambridge: Cambridge University Press.
- Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.
- Kukla, Rebecca. 2002. 'Attention and Blindness: Objectivity and Contingency in Moral Perception' in *Canadian Journal of Philosophy* Vol.32 Sup.1: 319-46.
- Kukla, Rebecca, and Ruetsche, Laura. 2002. 'Contingent Natures and Virtuous Knowers: Could Epistemology be "Gendered"?' in *Canadian Journal of Philosophy* Vol.32 No.3: 389-418.
- Kymlicka, Will. 1988. 'Liberalism and Communitarianism' in *Canadian Journal of Philosophy* Vol.18. No.2: 181-203.
- Lance, Mark N., and Little, Margaret O. 2006. 'Defending Moral Particularism' in *Contemporary Debates in Moral Theory*, James Dreier (ed.). Oxford: Blackwell. 305-22.
- Lance, Mark N., and Little, Margaret O. 2007a. 'Where the Laws Are' in *Oxford Studies in Metaethics*, Vol.2, Russ Schafer-Landau (ed.). Oxford: Oxford University Press. 149-171.
- Lance, Mark N., and Little, Margaret O. 2007b. 'From Particularism to Defeasibility' in *Challenging Moral Particularism*, Mark Lance, Matjaz Portc, Vojko Strahovnik (eds.). London: Routledge. 53-74.
- Lang, Gerald. 2001. 'The Rule-Following Considerations and Meta-Ethics: Some False Moves' in *European Journal of Philosophy* Vol.9 No.2: 190-209.
- Langton, Rae. 1993. 'Beyond a Pragmatic Critique of Reason' in *Australasian Journal of Philosophy* Vol.71: 364-84.
- Langton, Rae. 2007. 'Objective and Unconditioned Value' in *The Philosophical Review* Vol.116 No.2: 157-85.
- Lawford-Smith, Holly. 2012. 'The Motivation Question: Arguments from Justice and from Humanity' in *British Journal of Political Science* Vol.42 No.3: 661-78.
- LeBar, Mark. 2004. 'Good For You' in *Pacific Philosophical Quarterly* Vol.85: 195-217.
- LeBar, Mark. 2005. 'Three Dogmas of Response-Dependence' in *Philosophical Studies* Vol.123: 175-211.
- LeBar, Mark. 2008. 'Aristotelian Constructivism' in *Social Policy and Philosophy* Vol.25 No.1: 182-214.
- Leiter, Brian. 2004. 'The Hermeneutics of Suspicion: Recovering Marx, Nietzsche, and Freud' in Brian Leiter (ed.), *The Future for Philosophy*. Oxford: Clarendon Press. 74-105.
- Lenman, James. 2012. 'Expressivism and Constructivism' in *Constructivism in Practical Philosophy*, James Lenman and Yonatan Shemmer (eds.). Oxford: Oxford University Press. 213-25.

- Lewens, Tim. 2010. 'Foot Note' in *Analysis* Vol.70 No.3: 468-73.
- Lewens, Tim. 2012. 'Human Nature: The Very Idea' in *Philosophy and Technology* Vol.25 No.4: 459-74.
- Lewis, David, Johnston, Mark, and Smith, Michael. 1989. 'Dispositional Theories of Value' in *Proceedings of the Aristotelian Society, Supplementary Volumes* Vol.63: 89-111, 113-137, 139-174.
- Lillehammer, Hallvard. 1997. 'Smith on Moral Fetishism' in *Analysis* Vol.57 No.3: 187-95.
- Little, Margaret O. 1995. 'Seeing and Caring: The Role of Affect in Feminist Epistemology' in *Hypatia* Vol.10 No.3: 117-37.
- Little, Margaret O. 1997. 'Virtue as Knowledge: Objections from the Philosophy of Mind' in *Noûs*, Vol.31 No.1: 59-79.
- Lott, Micah. 2012a. 'Moral Virtue as Knowledge of Human Form' in *Social Theory and Practice* Vol.38 No.3: 407-31.
- Lott, Micah. 2012b. 'Have Elephant Seals Refuted Aristotle? Nature, Function, and Moral Goodness' in *Journal of Moral Philosophy* Vol.9: 1-23.
- Lovibond, Sabina. 2008. 'Practical Reason and its Animal Precursors' in *John McDowell: Experience, Norm, and Nature*, Jakob Lindgaard (ed.). Oxford: Blackwell. 112-23.
- Lugones, María C., and Spelman, Elizabeth V. 1983. 'Have We Got a Theory for You! Feminist Theory, Cultural Imperialism, and the Demand for "The Woman's Voice"' in *Women's Studies International Forum* Vol.6 No.6: 573-81.
- MacIntyre, Alasdair. 1999. *Dependent Rational Animals: Why Human Beings Need the Virtues*. Chicago: Open Court.
- MacIntyre, Alasdair. 2007. *After Virtue: A Study in Moral Theory, Third Edition*. Notre Dame, IN: Notre Dame University Press.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.
- Markovits, Julia. 2011. 'Why be an Internalist about Reasons?' in *Oxford Studies in Metaethics* Vol.6, Russ Schafer-Landau (ed.). Oxford: Oxford University Press. 255-79.
- McDowell, John. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- McDowell, John. 1998a. 'Wittgenstein on Following a Rule' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 221-62.
- McDowell, John. 1998b. 'Criteria, Defeasibility, and Knowledge' in *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press. 369-94.
- McDowell, John. 1998c. 'Knowledge and the Internal' in *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press. 395-413.
- McDowell, John. 1998d. 'Non-Cognitivism and Rule-Following' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 198-218.
- McDowell, John. 1998e. 'Values and Secondary Qualities' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 131-150.
- McDowell, John. 1998f. 'Anti-Realism and the Epistemology of Understanding' in *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press. 314-43.
- McDowell, John. 1998g. 'The Constitutive Ideal of Rationality: Davidson and Sellars' in *Critica* Vol.30 No.88: 29-48.

- McDowell, John. 1998h. 'Virtue and Reason' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 50-76.
- McDowell, John. 1998i. 'Are Moral Requirements Hypothetical Imperatives' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 77-94.
- McDowell, John. 1998j. 'Some Issues in Aristotle's Moral Psychology' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 23-49.
- McDowell, John. 1998k. 'Might There Be External Reasons?' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 95-111.
- McDowell, John. 1998l. 'Knowledge by Hearsay' in *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press. 414-44.
- McDowell, John. 1998m. 'Two Sorts of Naturalism' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 167-97.
- McDowell, John. 1998n. 'Aesthetic Value, Objectivity, and the Fabric of the World' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 112-30.
- McDowell, John. 1998o. 'Projection and Truth in Ethics' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 151-66.
- McDowell, John. 1998p. 'The Role of Eudaimonia in Aristotle's Ethics' in *Mind, Value, and Reality*. Cambridge, MA: Harvard University Press. 3-22.
- McDowell, John. 2006. 'Response to Pettit and Smith' in *McDowell and his Critics*, Cynthia Macdonald and Graham Macdonald (eds.). Oxford: Blackwell. 170-79.
- McDowell, John. 2008a. 'Avoiding the Myth of the Given' in *John McDowell: Experience, Norm, and Nature*, Jakob Lindgaard (ed.). Oxford: Blackwell. 1-14.
- McDowell, John. 2008b. 'Responses' in *John McDowell: Experience, Norm, and Nature*, Jakob Lindgaard (ed.). Oxford: Blackwell. 200-267.
- McDowell, John. 2009a. 'Gadamer and Davidson on Understanding and Relativism' in *The Engaged Intellect: Philosophical Essays*. Cambridge, MA: Harvard University Press. 134-151.
- McDowell, John. 2009b. 'Deliberation and Moral Development in Aristotle's Ethics' in *The Engaged Intellect: Philosophical Essays*. Cambridge, MA: Harvard University Press. 41-58.
- McDowell, John. 2009c. 'Eudaimonism and Realism in Aristotle's Ethics' in *The Engaged Intellect: Philosophical Essays*. Cambridge, MA: Harvard University Press. 23-40.
- McDowell, John. 2009d. 'Naturalism in the Philosophy of Mind' in *The Engaged Intellect: Philosophical Essays*. Cambridge, MA: Harvard University Press. 457-78.
- McDowell, John. 2009e. 'Wittgensteinian "Quietism"' in *Common Knowledge* Vol.15 No.3: 365-72.
- McDowell, John. 2011. *Perception as a Capacity for Knowledge*. Milwaukee, WI: Marquette University Press.
- McKeever, Sean, and Ridge, Michael. 2005. 'The Many Moral Particularisms' in *Canadian Journal of Philosophy* Vol.35 No.1: 83-106.
- McNaughton, David. 1988. *Moral Vision: An Introduction to Ethics*. Oxford: Blackwell.
- Mill, J.S. 1904. 'On Nature' in *Nature, The Utility of Religion, and Theism*. London: Watts. 7-33.
- Miller, Alexander. 2003. *An Introduction to Contemporary Metaethics*. Cambridge:

Polity.

- Miller, Alexander. 2010. 'Realism' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/spr2012/entries/realism/>.
- Millgram, Elijah. 2009. 'Life and Action' in *Analysis Reviews* Vol.69 No.3: 557-64.
- Millikan, Ruth G. 1993. 'White Queen Psychology' in *White Queen Psychology, and other Essays for Alice*. Cambridge, MA: MIT Press. 279-364.
- Mills, Charles W. 2005. "'Ideal Theory'" as Ideology' in *Hypatia* Vol.20 No.2: 185-84.
- Mills, Charles W. 2012. 'Occupy Liberalism! Or, Ten Reasons Why Liberalism Cannot Be Retrieved For Radicalism (And Why They're All Wrong)' in *Radical Philosophy Review* Vol.15 No.2: 305-23.
- Mišćević, Nenad. 1998. 'The Aposteriority of Response-Dependence' in Peter Menzies (ed.), 'Secondary Qualities Generalized' in *The Monist* Vol.81 No.1: 69-84.
- Moellendorf, Darrel. 2002. *Cosmopolitan Justice*. Boulder, CO: Westview.
- Mohanty, Chandra T. 1991. 'Under Western Eyes: Feminist Scholarship and Colonial Discourses' in *Third World Women and the Politics of Feminism*, Chandra T. Mohanty, Ann Russo, Lourdes Torres (eds.). Bloomington: Indiana University Press. 52-80.
- Nelson, Eric. 2008. 'From Primary Goods to Capabilities: Distributive Justice and the Problem of Neutrality' in *Political Theory* Vol.36 No.1: 93-122.
- Neurath, Otto. 1944. 'Foundations of the Social Sciences' in *International Encyclopedia of Unified Science* Vol.2 No.1, O. Neurath, R. Carnap, Ch. Morris (eds.). Chicago: University of Chicago Press.
- Newman, Micah. 2014. 'A Realist Sexual Ethics' in *Ratio* (forthcoming): 1-18.
- Nussbaum, Martha C. 1986. 'Non-Scientific Deliberation' in *The Fragility of Goodness: Luck and Ethics in Greek Philosophy*. Cambridge: Cambridge University Press. 290-317.
- Nussbaum, Martha C. 1987. 'Nature, Function, and Capability: Aristotle on Political Distribution' in *WIDER Working Papers*, No.31.
- Nussbaum, Martha C. 1990a. 'Aristotelian Social Democracy' in *Liberalism and the Good*, R. Bruce Douglass, Gerald M. Mara, and Henry S. Richardson (eds.). London: Routledge. 203-52.
- Nussbaum, Martha C. 1990b. 'The Discernment of Perception: An Aristotelian Conception of Public and Private Rationality' in *Love's Knowledge*. Oxford: Oxford University Press. 54-105.
- Nussbaum, Martha C. 1992. 'Human Functioning and Social Justice: In Defence of Aristotelian Essentialism' in *Political Theory* Vol.
- Nussbaum, Martha C. 1993. 'Non-Relative Virtues' in *The Quality of Life*, Martha C. Nussbaum and Amartya Sen (eds.). Oxford: Clarendon. 242-69.
- Nussbaum, Martha C. 1995. 'Aristotle on Human Nature and the Foundations of Ethics' in *World, Mind, and Ethics: Essays on the Philosophy of Bernard Williams*, J.E.J. Altham and Ross Harrison (eds.). Cambridge: Cambridge University Press. 86-131.
- Nussbaum, Martha C. 1997. 'Capabilities and Human Rights' in *Fordham Law Review* Vol.66: 273-300.
- Nussbaum, Martha C. 1998. 'The Good as Discipline, The Good as Freedom' in *Ethics of Consumption: The Good Life, Justice, and Global Stewardship*, David A. Crocker

- and Toby Linden (eds.). Lanham, MD: Rowman and Littlefield.
- Nussbaum, Martha C. 1999a. 'The Feminist Critique of Liberalism' in *Sex and Social Justice*. Oxford: Oxford University Press. 55-80.
  - Nussbaum, Martha C. 1999b. 'Objectification' in *Sex and Social Justice*. Oxford: Oxford University Press. 213-39.
  - Nussbaum, Martha C. 2000a. *Women and Human Development: The Capabilities Approach*. New York: Cambridge University Press.
  - Nussbaum, Martha C. 2000b. 'Aristotle, Politics, and Human Capabilities: A Response to Antony, Arneson, Charlesworth, and Mulgan' in *Ethics* Vol.111 No.1: 102-40.
  - Nussbaum, Martha C. 2001a. 'Symposium on Amartya Sen's philosophy 5: Adaptive Preferences and Women's Options' in *Economics and Philosophy* Vol.17 No.1: 67-88.
  - Nussbaum, Martha C. 2001b. *The Fragility of Goodness, Revised Edition*. Cambridge: Cambridge University Press.
  - Nussbaum, Martha C. 2003. 'Capabilities as Fundamental Entitlements: Sen and Social Justice' in *Feminist Economics* Vol.9 No.2-3: 33-59.
  - Nussbaum, Martha C. 2006a. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: Harvard University Press.
  - Nussbaum, Martha C. 2006b. 'Reply: In Defence of Global Political Liberalism' in *Development and Change* Vol.37 No.6: 1313-28.
  - Nussbaum, Martha C., and Sen, Amartya. 1993. 'Introduction' in *The Quality of Life*, Martha C. Nussbaum and Amartya Sen (eds.). Oxford: Clarendon. 1-8.
  - Okin, Susan M. 1989. *Justice, Gender, and the Family*. New York: Basic Books.
  - Okin, Susan M. 2003. 'Poverty, Well-Being, and Gender: What Counts, Who's Heard?' in *Philosophy and Public Affairs* Vol.31 No.3: 280-316.
  - Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
  - Parfit, Derek. 2011a. *On What Matters* (Vol.1.). Oxford: Oxford University Press.
  - Parfit, Derek. 2011b. *On What Matters* (Vol.2.). Oxford: Oxford University Press.
  - Phillips, Anne. 2007. *Multiculturalism Without Culture*. Princeton: Princeton University Press.
  - Pippin, Robert. 2002. 'Leaving Nature Behind: Or, Two Cheers for Subjectivism' in *Reading McDowell: Essays on Mind and World*, Nick Smith (ed.). New York: Routledge. 58-76.
  - Plato. 1997. 'Euthyphro' in *Plato: Complete Works*, John M. Cooper and D.S. Hutchinson (eds.). Indianapolis, IN: Hackett. 1-16.
  - Pogge, Thomas. 2008. *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms*. Cambridge: Polity Press.
  - Poole, Ross. 2005. *Morality and Modernity*. Taylor and Francis e-Library.
  - Putnam, Hilary. 1993. 'Objectivity and the Science-Ethics Distinction' in *The Quality of Life*, Martha C. Nussbaum and Amartya Sen (eds.). Oxford: Clarendon. 141-57.
  - Railton, Peter. 1986. 'Moral Realism' in *The Philosophical Review* Vol.95: 163-207.
  - Randell-Koons, Jeremy. 2003. 'Why Respondence-Dependence Theories of Morality Are False' in *Ethical Theory and Moral Practice* Vol.6: 275-94.
  - Rawls, John. 1997. 'The Idea of Public Reason Revisited' in *The University of Chicago Law Review* Vol.64 No.3: 765-807.
  - Rawls, John. 1999a. *A Theory of Justice: Revised Edition*. Cambridge, MA: Harvard University Press.

- Rawls, John. 1999b. *The Law of Peoples*. Cambridge, MA: Harvard University Press.
- Rawls, John. 2001. *Justice as Fairness: A Restatement*, Erin Kelly (ed.). London: Belnap.
- Rawls, John. 2005. *Political Liberalism: Expanded Edition*. New York: Columbia University Press.
- Richardson, Henry S. 2000. 'Some Limitations of Nussbaum's Capabilities' in *Quinnipiac Law Review* Vol.19: 309-32.
- Ricoeur, Paul. 1970. *Freud and Philosophy: An Essay on Interpretation*, trans. Denis Savage. New Haven, CT: Yale University Press.
- Ridge, Michael. 2014. 'Moral Non-Naturalism' in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Accessed on 30/09/14 at <http://plato.stanford.edu/archives/fall2014/entries/moral-non-naturalism/>.
- Roberts, Debbie. 2011. 'Shapelessness and the Thick' in *Ethics* Vol.121 No.3: 489-520.
- Robeyns, Ingrid. 2003. 'Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities' in *Feminist Economics* Vol.9 No.2-3: 61-92.
- Robeyns, Ingrid. 2005. 'Selecting Capabilities for Quality of Life Measurement' in *Social Indicators Research* Vol.74: 191-215.
- Robeyns, Ingrid. 2006. 'The Capability Approach in Practice' in *The Journal of Political Philosophy* Vol.14 No.3: 351-376.
- Ruetsche, Laura. 2004. 'Virtue and Contingent History: Possibilities for Feminist Epistemology' in *Hypatia* Vol.19 No.1: 73-101.
- Sandel, Michael. 1984. 'The Procedural Republic and the Unencumbered Self' in *Political Theory* Vol.12 No.1: 81-96.
- Saul, Jennifer M. 2003. *Feminism: Issues and Arguments*. Oxford: Oxford University Press.
- Sen, Amartya. 1982. 'Rights and Agency' in *Philosophy and Public Affairs* Vol.11, No.1: 3-39.
- Sen, Amartya. 1983. 'Evaluator Relativity and Consequential Evaluation' in *Philosophy and Public Affairs* Vol.12 No.2: 113-132.
- Sen, Amartya. 1993. 'Capability and Well-Being' in *The Quality of Life*, Martha C. Nussbaum and Amartya Sen (eds.). Oxford: Clarendon. 30-53.
- Sen, Amartya. 2009. *The Idea of Justice*. London: Penguin.
- Setiya, Kieran. 2010. 'Does Moral Theory Corrupt Youth?' in *Philosophical Topics* Vol.38 No.1: 205-22.
- Setiya, Kieran. 2012. 'Introduction: Internal Reasons' in *Internal Reasons: Contemporary Readings*, Kieran Setiya and Hille Paakkunainen (eds.). Cambridge, MA: MIT Press. 1-34.
- Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Schaffer, Jonathan. 2009. 'On What Grounds What' in *Metametaphysics: New Essays on the Foundations of Ontology*, David Manley, David J. Chalmers, and Ryan Wasserman (eds.). Oxford: Oxford University Press. 347-83.
- Scheman, Naomi. 2001. 'Non-negotiable Demands: Metaphysics, Politics, and the Discourse of Needs' in *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy*, Juliet Floyd and Sanford Shieh (eds.). Oxford: Oxford University Press.

314-338.

- Schroeder, Mark. 2005. 'The Humean Theory of Reasons' in *Oxford Studies in Metaethics*, Vol.2, Russ Schafer-Landau (ed.). Oxford: Oxford University Press. 195-219.
- Skorupski, John. 2012. 'Aristotelianism and Modernity: Terence Irwin on the Development of Ethics' in *European Journal of Philosophy* Vol.20 No.2: 312-37.
- Slote, Michael. 2010. *Moral Sentimentalism*. Oxford: Oxford University Press.
- Smith, Michael. 1993. 'Objectivity and Moral Realism: On the Significance of the Phenomenology of Moral Experience' in *Reality, Representations, and Projection*, John Haldane and Crispin Wright (eds.). Oxford: Oxford University Press. 235-55.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.
- Smith, Michael. 1995. 'Internal Reasons' in *Philosophy and Phenomenological Research* Vol.55: 109-31.
- Smith, Michael. 1996. 'The Argument for Internalism: Reply to Miller' in *Analysis* Vol.56: 175-84.
- Stark, Susan. 2001. 'Virtue and Emotion' in *Noûs* Vol.35 No.3: 440-455.
- Stemplowka, Zofia, and Swift, Adam. 2012. 'Ideal and Non-Ideal Theory' in *The Oxford Handbook of Political Philosophy*, David Estlund (ed.). Oxford: Oxford University Press. 373-92.
- Strawson, P.F. 1962. 'Freedom and Resentment' in *Proceedings of the British Academy* Vol.48: 1-25.
- Street, Sharon. 2010. 'What is Constructivism in Ethics and Metaethics?' in *Philosophy Compass* Vol.5 No.5: 363-84.
- Street, Sharon. 2012. 'Coming to Terms with Contingency: Humean Constructivism about Practical Reason' in *Constructivism in Practical Philosophy*, James Lenman and Yonatan Shemmer (eds.). Oxford: Oxford University Press. 40-59.
- Stichter, Matt. 2007. 'Ethical Expertise: The Skill Model of Practice' in *Ethical Theory and Moral Practice* Vol.10 No.2: 183-94.
- Sturgeon, Nicholas. 2006. 'Ethical Naturalism' in *The Oxford Handbook of Ethical Theory*, David Copp (ed.). Oxford: Oxford University Press. 91-121.
- Svensson, Frans. 2010. 'Virtue Ethics and the Search for an Account of Right Action' in *Ethical Theory and Moral Practice* Vol.13: 255-71.
- Swanton, Christine. 2003. *Virtue Ethics: A Pluralistic View*. Oxford: Oxford University Press.
- Tamir, Yael. 1996. 'Hands Off Clitoridectomy' in *Boston Review* (Summer 1996).
- Taylor, C. 1991. 'What's Wrong with Negative Liberty' in *Liberty*, David Miller (ed.). Oxford: Oxford University Press. 141-62.
- Tessman, Lisa. 2005. *Burdened Virtues: Virtue Ethics for Liberatory Struggles*. Oxford: Oxford University Press.
- Tessman, Lisa. 2009a. 'Introduction' in *Feminist Ethics and Social and Political Philosophy: Theorising the Non-Ideal*, Lisa Tessman (ed.). New York: Springer. xiii-xxvi.
- Tessman, Lisa. 2009b. 'Feminist Eudaimonism: Eudaimonism as Non-Ideal Theory' in *Feminist Ethics and Social and Political Philosophy: Theorising the Non-Ideal*, Lisa Tessman (ed.). New York: Springer. 47-58.
- Tessman, Lisa. 2010. 'Idealizing Morality' in *Hypatia* Vol.25 No.4: 797-824.

- Thomas, Laurence M. 1992-3. 'Moral Deference' in *Philosophical Forum* Vol.24: 233-50
- Thompson, Michael. 1995. 'The Representation of Life' in *Virtues and Reasons*, Rosalind Hursthouse, Gavin Laurence and Warren Quinn (eds.). Oxford: Oxford University Press. 247-97.
- Thompson, Michael. 2003. 'Three Degrees of Natural Goodness' in *Iride* Vol.16 No.38: 179-200.
- Thompson, Michael. 2004a. 'Apprehending Human Form' in *Modern Moral Philosophy*, Antony O'Hear (ed.). Cambridge: Cambridge University Press. 47-74.
- Thompson, Michael. 2004b. 'What is it to Wrong Someone? A Puzzle About Justice' in *Reason and Value*, R. Jay Wallace, Phillip Pettit, Samuel Scheffler, and Michael Smith (eds.). Oxford: Oxford University Press. 333-84.
- Thompson, Michael. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge, MA: Harvard University Press.
- Thornton, Tim. 2004. *John McDowell*. Chesham: Acumen.
- Tronto, Joan. 2012. 'Partiality Based on Relational Responsibilities: Another Approach to Global Ethics' in *Ethics and Social Welfare* Vol.6 No.3: 303-16.
- ul Haq, Mahbub et al. 1990. *Human Development Report 1990*. Oxford: Oxford University Press.
- Vallentyne, Peter. 2005. 'Debate: Capabilities versus Opportunities for Well-Being' in *The Journal of Political Philosophy* Vol.13 No.3: 359-71.
- Vallier, Kevin. 2011. 'Consensus and Convergence in Public Reason' in *Public Affairs Quarterly* Vol.25 No.4: 261-79.
- Velleman, J. David. 2009 *How We Get Along*. New York: Cambridge University Press.
- Velleman, J. David. 2013. *Foundations for Moral Relativism*. Open Book Publishers. Accessed 09/09/14 at <http://www.openbookpublishers.com/product/181>.
- Väyrynen, Pekka. 2007. 'Usable Moral Principles' in *Challenging Moral Particularism*, Mark Lance, Matjaz Portc, Vojko Strahovnik (eds.). London: Routledge. 75-106.
- Väyrynen, Pekka. 2014. 'Shapelessness in Context' in *Noûs* Vol.48: 573-93.
- Vogler, Candace. 2006. 'Modern Moral Philosophy Again: Isolating the Promulgation Problem' in *Proceedings of the Aristotelian Society* Vol.106: 347-64.
- Walsh, Vivian. 2003. 'Sen after Putnam' in *Review of Political Economy* Vol.15 No.3: 315-94.
- Wedgwood, Ralph. 1997. 'The Essence of Response-Dependence' in *European Review of Philosophy* Vol.3: 31-54.
- Weinberg, Justin. 2011. 'Review: Iris Marion Young. Responsibility for Justice' in *Ethics* Vol.122 No.2: 224-8.
- Wenar, Leif. 1995. 'Political Liberalism: An Internal Critique' in *Ethics* Vol.106 No.1: 32-62.
- Wiggins, David. 1987. 'A Sensible Subjectivism?' in *Needs, Values, Truth: Essays in the Philosophy of Value*. New York: Oxford University Press. 185-214.
- Williams, Bernard A.O. 1973. 'A Critique of Utilitarianism' in *Utilitarianism: For and Against*, Bernard A.O. Williams and J.J.C. Smart. Cambridge: Cambridge University Press. 77-150.
- Williams, Bernard A.O. 1995a. 'Internal Reasons and the Obscurity of Blame' in

- Making Sense of Humanity and other philosophical papers*. Cambridge: Cambridge University Press. 35-45.
- Williams, Bernard A.O. 1995b. 'Replies' in *World, Mind, and Ethics: Essays on the Philosophy of Bernard Williams*, J.E.J. Altham and Ross Harrison (eds.). Cambridge: Cambridge University Press. 185-224.
  - Williams, Bernard A.O. 2011. *Ethics and the Limits of Philosophy*. Oxford: Routledge.
  - Williams, Bernard A.O. 2012. 'Internal and External Reasons' in *Internal Reasons: Contemporary Readings*, Kieran Setiya and Hille Paakkunainen (eds.). Cambridge, MA: MIT Press. 37-50.
  - Wittgenstein, Ludwig. 1969. *On Certainty*, Elizabeth Anscombe and G.H. von Wright (eds.). Oxford: Blackwell.
  - Wittgenstein, Ludwig. 2001. *Philosophical Investigations*, Elizabeth Anscombe (trans., ed.). Oxford: Blackwell.
  - Wright, Crispin. 1988. 'Moral Values, Projection, and Secondary Qualities' in *Proceedings of the Aristotelian Society, Supplementary Volumes* Vol.62: 1-26.
  - Wright, Crispin. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.
  - Wright, Crispin. 1996. 'Human Nature?' in *European Journal of Philosophy* Vol.4 No.2: 235-53.
  - Wright, Crispin. 2007. 'Rule-Following without Reasons: Wittgenstein's Quietism and the Constitutive Question' in *Ratio* Vol.20 No.4: 481-502.
  - Wylie, Alison. 2003. 'Why Standpoint Matters' in *Science and other Cultures: Issues in philosophies of Science and Technology*, Robert Figueroa and Sandra Harding (eds.). New York: Routledge. 26-48.
  - Young, Iris Marion. 2004. 'Responsibility and Global Labour Justice' in *The Journal of Political Philosophy* Vol.12 No.4: 365-88.
  - Young, Iris Marion. 2006. 'Responsibility and Global Justice: A Social Connection Model' in *Social Philosophy and Policy* Vol.23 No.1: 102-30.
  - Young, Iris Marion. 2009. 'Structural Injustice and the Politics of Difference' in *Contemporary Debates in Political Philosophy*, Thomas Christiano and John Christman (eds.). Oxford: Blackwell. 362-84.
  - Young, Iris Marion. 2011. *Responsibility for Justice*. Oxford: Oxford University Press.
  - Zangwill, Nick. 2003. 'Against Moral Response-Dependence' in *Erkenntnis* Vol.59: 285-90