Taking McDowell’s Anti-Humeanism and Particularism Seriously:
What This Means and Why It Matters

By James C. Taggart
B.A., Northwestern University, 1986
J.D., University of Chicago, 1990
A.M., Brown University, 1997

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy
in the Department of Philosophy at Brown University

Providence, Rhode Island
May 2009
COPYRIGHT NOTICE

Copyright © James C. Taggart, 2009
All rights reserved
This dissertation by James C. Taggart is accepted in its present form by the Department of Philosophy as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date__________________ __________________________________________
  James Dreier, Advisor

Recommended to the Graduate Council

Date__________________ __________________________________________
  Dan Brock, Reader

Date__________________ __________________________________________
  Jaegwon Kim, Reader

Date__________________ __________________________________________
  Bernard Reginster, Reader

Approved by the Graduate Council

Date__________________ __________________________________________
  Sheila Bonde, Dean of the Graduate School
ABOUT THE AUTHOR

After defending his dissertation in August 2000, James Taggart spent two years as a Greenwall Fellow in Bioethics and Health Policy at Johns Hopkins University in Baltimore, Maryland. He has also taught at Brown University, Wellesley College, Georgetown University, and the University of Maryland in College Park. His current research focuses on the intersection of ethics, metaethics, and the philosophy of literature. He lives in Silver Spring, Maryland with his wife and two children.
ACKNOWLEDGEMENTS

Thanks to Jamie for playing devil’s advocate all these years: I don’t know anyone who performs this role quite so well. Of course, I learned a lot from our discussions, though perhaps not always what you intended. Thanks to my family – Mom and Dad, Larry and Bobbi, Billy and Kira – for all the patience you have shown with me. I know that I have tried it far too often. And thanks most of all to Lor who has stood by me and helped in so many ways throughout the whole process.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td><strong>Chapter One</strong></td>
<td></td>
</tr>
<tr>
<td>Cognitive Pro-Attitudes:</td>
<td>18</td>
</tr>
<tr>
<td>An Explication and Defense of Anti-Humeanism</td>
<td></td>
</tr>
<tr>
<td><strong>Chapter Two</strong></td>
<td></td>
</tr>
<tr>
<td>Having Your (Layered) Cake and Eating It Too:</td>
<td>44</td>
</tr>
<tr>
<td>Saving Nonreductivism from Trouble with Mental Causation</td>
<td></td>
</tr>
<tr>
<td><strong>Chapter Three</strong></td>
<td></td>
</tr>
<tr>
<td>Waking Moral Particularism from its Dogmatic Slumbers:</td>
<td>87</td>
</tr>
<tr>
<td>From Murdochian Description to Wittgensteinian Criteria &amp; Back Again</td>
<td></td>
</tr>
<tr>
<td><strong>Chapter Four</strong></td>
<td></td>
</tr>
<tr>
<td>Defending Criterial Particularism as Constitutive Generalism:</td>
<td>146</td>
</tr>
<tr>
<td>(One Small Step) Towards Rehabilitating Analytic Naturalism</td>
<td></td>
</tr>
<tr>
<td><strong>Chapter Five</strong></td>
<td></td>
</tr>
<tr>
<td>On the Very Idea of Nondefinitive Analytic Naturalism:</td>
<td>162</td>
</tr>
<tr>
<td>Rule-Following and an Alternative Account of Moral Reasoning</td>
<td></td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>216</td>
</tr>
</tbody>
</table>
INTRODUCTION

Both anti-Humeanism and moral particularism have received a great deal of attention in the wake of John McDowell’s suggestive writings on these topics. However, most accounts of anti-Humeanism and particularism proceed differently from what a close reading of McDowell would imply. This dissertation rectifies this imbalance by developing an anti-Humean conception of motivation and a particularistic account of ethical thought that takes seriously McDowell’s central insights in these areas.

Chapter One elaborates McDowell’s distinctive anti-Humeanism. Since I provide a detailed map of my argument at the beginning of Chapter One, I do not repeat that here. However, about the broader place and importance of this argument, I should mention three things. First of all, I do not attempt in Chapter One – nor do I suspect it even possible – to provide a knock-down argument for McDowell’s (neo-Aristotelian) anti-Humeanism. Rather, I have as my primary aim something more modest: to show not only that the anti-Humean position suggested by McDowell’s writings makes sense but what making sense of it involves.¹

This first general point leads to a second. If anti-Humeanism is possible – if there could be a kind of thought capable (in itself) of moving us to act – then this kind of

---

¹ Hereafter, I talk simply of anti-Humeanism without reference to McDowell. However, unless otherwise noted, I should still be understood as referring to the anti-Humeanism suggested by McDowell’s writings and explored in Chapter One. As I explain there, contemporary anti-Humeans can be broadly categorized as either neo-Aristotelian or neo-Kantian, and McDowell falls into the former category.
thought could not be mere belief. For there are many ways of arriving at belief and, correspondingly, many ways that belief can fail to engage us personally.  

Therefore, I spend a significant part of Chapter One exploring the kind of thought that anti-Humeanism requires, ultimately arriving at an important insight: the kind of thinking capable (in itself) of moving us to act can only be an immediate, first-person experiential/quasi-perceptual kind of thinking, not mere belief. Moreover, while this kind of thinking may seem esoteric and obscure, it is ubiquitous not only if anti-Humeanism is correct but also since, as I discuss in Chapter Five, essentially the same kind of thinking is involved in the successful, genuine appreciation of art and literature.

Now emphasizing, as I have here and do below in Chapter One, that what anti-Humeanism requires cannot be mere belief but can only be an immediate, first-person experiential/quasi-perceptual kind of thinking – and, thereby, avoiding the relatively straightforward refutation that follows if the cognition in anti-Humeanism (or internalist moral cognitivism) is taken to be mere belief – can prove underwhelming to skeptics of anti-Humeanism. And the skeptics have a point. For it matters not what we call this supposed form of cognition that can (in itself) move us to act but what this so-called thinking really involves. Skeptics about anti-Humeanism suspect that this so-called thinking – this “new” form of cognition – is really the same kind of noncognitive

---

2 That mere belief can fail to engage us personally makes anti-Humeanism (or ‘internalist moral cognitivism’) a relatively easy target if belief were the only kind of thinking upon which anti-Humeanism could depend. See Alfred Mele, "Internalist Moral Cognitivism and Listlessness," Ethics, 106 (1996): 727-753.

3 As reflected in the title of Chapter One, I call the distinctive kind of thinking involved in anti-Humeanism ‘cognitive pro-attitudes’ (or ‘CPA’s’ for short).

4 I put scare quotes around ‘new’ because whatever its merits as way to make sense of anti-Humeanism along the lines I propose, the idea that there is a distinctive kind of immediate, first-personal thinking is not a new one.
motivation that they have always recognized, only now masquerading under a new name. In brief, skeptics object that I am pushing a distinction without a difference. What would make a difference? It would make a difference if the so-called thinking that I urge as anti-Humeanism’s savior really possessed (some of) the criteria of genuine thinking, to wit, that it was responsive to and could play a role in reasoning. With this as background, then, my third general comment about Chapter One is this: without more, my argument there is incomplete. As we just saw, if the so-called thinking upon which anti-Humeanism depends is to deserve the name and, thereby, make anti-Humeanism possible, then this kind of thinking must be shown to be responsive to reasoning and play a part in it. Not coincidently, filling in this lacuna in the argument of Chapter One constitutes one of the (other) motivations for the alternative account of moral reasoning developed in Chapter Five.5

More generally, apart from filling in the lacuna just discussed, Chapter One suggests two further lines of investigation. The first I pursue in Chapter Two, and it concerns another obstacle that anti-Humeanism must surmount: making room for mental causation. Broadly speaking, mental causation constitutes a problem for anti-Humeanism because it is a form of nonreductive physicalism and, as I discuss in Chapter Two, accommodating mental causation has posed difficulties for nonreductive physicalism. More specifically, mental causation poses a problem for anti-Humeanism given two seemingly innocuous points. First, it allows that various sciences of the brain (e.g. neurology) study the underpinnings of our mental lives, trying to find the causal basis of this behavioral output or that behavioral capacity by looking to brain states picked out in

5 See the final footnote and the corresponding discussion in this Introduction for one of the other motivations behind Chapter Five’s alternative account of moral reasoning.
terms of each science’s preferred vocabulary, respectively. But anti-Humeanism also grants what the first point implies: that we have other ways of conceptualizing “our” causal interactions with the world besides the scheme that anti-Humeanism recommends involving cognitive pro-attitudes (or CPA’s). Taken together, then, these two points threaten to preempt from below the rational-causal story that anti-Humeanism would tell about what moves us to act.

Still, by themselves, these two points may not appear to pose any more trouble for anti-Humeanism than they do for nonreductive physicalism generally. However, the trouble presses harder against anti-Humeanism because, unlike other forms of nonreductive physicalism, it cannot adopt what is perhaps the main strategy for accommodating mental causation. This strategy purports to accommodate mental causation by at least allowing different levels of causal-explanatory relations even if the fundamental physical level is the only level with actual, full-blooded causal relations.⁶ In other words, the thought behind this strategy goes: even if nonreductive physicalism cannot accommodate causal relations between mental items, it can at least accommodate causal explanations involving such items. Anti-Humeanism, however, cannot adopt this strategy because doing so would undermine the integrity and, thereby, the very possibility of CPA’s. The integrity of CPA’s can be maintained only if the same transaction between agent and world that anti-Humeanism would account for in terms of one mental state (i.e. a CPA) cannot also be explained in terms of, say, two brain states. But if the explanatory pluralism that this strategy envisions were true and such a parallel, lower-level explanation were possible for the same phenomena that anti-Humeanism would

---

⁶ I say ‘purports to’ because, as discussed in Chapter Two, this common strategy for accommodating mental causation fails for reasons set forth by Jaegwon Kim.
explain with CPA’s, then the integrity that CPA’s require – not merely nominal but real integrity – would be unavailable. Consequently, anti-Humeanism has even greater reason to find a better way to accommodate mental causation than other forms of nonreductive physicalism have yet done. And it is precisely this task that I take up in Chapter Two.

The second main line of investigation suggested by Chapter One’s defense of anti-Humeanism concerns moral particularism. Not only McDowell but others too have suggested a connection between anti-Humeanism and particularism. Initially, I may have been inclined to explore particularism primarily as a means to further elucidate – and perhaps even support – anti-Humeanism. However, particularism deserves its own treatment, independently of any connection it might have to anti-Humeanism. Indeed, from where I stand now, acceding to my initial inclination would have had the anti-Humean tail wagging the particularistic dog. Thus, working out the view that I call ‘criterial particularism’, including defending this view as superior to other kinds of particularism (e.g. the currently popular defeasibility view), constitutes the primary focus of Chapters Three to Five.

This focus on particularism, however, does not mean that these chapters have nothing to do with anti-Humeanism. Quite the contrary, the insight about anti-Humeanism mentioned above – that the kind of understanding capable (in itself) of moving us to act can only be an immediate, first-person experiential/quasi-perceptual kind of understanding, not mere belief – resonates with much of the argument in Chapters Three and Five. For example, in Chapter Three, I discuss Iris Murdoch’s distinctive conception of practical reasoning to help illuminate particularism in part because her
view stresses the required immediacy. In Chapter Five, while developing an alternative account of moral reasoning for criterial particularism, I take up the somewhat daunting topic of Wittgenstein’s rule-following considerations in part because these considerations arguably show how someone can engage in moral reasoning without relying (even implicitly) upon an independently formulable principle to guide one’s deliberations, again preserving the immediacy that anti-Humeanism requires. Finally, also in Chapter Five, Wittgenstein’s ideas of aspect perception and perspicuous representation prove attractive in developing an alternative account of moral reasoning in part because they easily accommodate the crucial anti-Humean insight about immediacy.

More broadly, apart from these specific examples of how anti-Humeanism shapes the account of criterial particularism in Chapters Three to Five, particularism and anti-Humeanism can also be seen as intertwined in light of the analogy sometimes drawn between particularism and art appreciation. Indeed, I have taken the art appreciation analogy to heart, and it has influenced much of my approach in Chapters Three to Five. While this analogy has usually been drawn to illuminate particularism’s commitment to holism about reasons, it also throws light on anti-Humeanism while suggesting how anti-Humeanism and particularism might be connected. Roughly, the idea is this: you do not understand – at least not fully – a particular work if you fail to respond to it in a certain

---


8 My recognition of the art appreciation analogy’s importance owes something to a conversation that occurred at my oral defense. Specifically, I recall a discussion involving James Dreier, Bernard Reginster, and myself. This discussion encouraged me to take the analogy more seriously and pursue it further.
way. So, for example, if *Romeo and Juliet* prompts you to howls of laughter, then you are missing something: you have not understood Shakespeare’s play correctly.\(^9\)

Moreover, cases of successful, genuine art appreciation reveal an immediate, first-person experiential/quasi-perceptual understanding at work, confirming merit in the crucial anti-Humean insight.

However, despite the importance of the art appreciation analogy for my account of particularism, the analogy has its limits. Attending to these limits helps bring out what makes the variant of particularism that I defend – criterial particularism – distinctive. As discussed in Chapter Three, particularism typically invokes the art appreciation analogy to explain (and obliquely support) some form of holism about reasons: unreconstructed particularism urges holism without default reasons; defeasibility particularism urges holism with default reasons. Nevertheless, according to both these kinds of particularism and their respective forms of holism, we cannot tell whether certain features of an action unqualifiedly count for (or against) doing the action without considering *all* the features of the action given the circumstances. Thus, orthodox particularism characteristically holds that: just as whether adding a patch of red makes a painting more aesthetically appealing must wait upon our examination of the red patch in the context of the work as a whole, so whether burning a child by extinguishing one’s spent cigarettes on her skin counts against doing such a thing must wait upon our consideration of the full context of the action.

---
\(^9\) This example derives from a scene in Aldous Huxley’s *Brave New World*. Examples of this link between one’s response to art and one’s appreciation of it according to which certain failures of response imply a failure of understanding appear ubiquitous.
From the perspective of criterial particularism, however, orthodox particularism takes the art appreciation analogy too far. Criterial particularism grants that some situations may need to be morally assessed in this holistic way. Indeed, it even allows that some situations taken holistically might serve as touchstones (or criteria) for understanding moral concepts. Nevertheless, criterial particularism would wake orthodox particularism from its dogmatic slumbers by rejecting its leading thought (or dogma), namely, that the holism suggested by the art appreciation analogy should be our exclusive model for understanding the behavior of moral reasons. As I discuss in Chapter Three, rejecting this dogma manifests itself initially in the recognition that moral criteria include not only entire situations taken holistically but also situations more-or-less richly described (as in the description above of someone’s burning a child with cigarette butts). But further: recognizing that rich descriptions, not just whole situations, can serve as criteria for thick moral terms means acknowledging not only moral principles (contra unreconstructed particularism) but unhedged moral principles (contra defeasibility particularism).

Now the unhedged moral principles that criterial particularism acknowledges – e.g. ‘Burning a child with cigarette butts is cruel’ – prove even more distinctive and interesting. As I discuss in Chapter Four (and also, to some extent, in Chapter Five), these principles collectively represent at least a partial rehabilitation of analytic

---

10 I intend ‘the holism suggested by the art appreciation analogy’ broadly to include the form of holism that allows for default reasons, that is, the form endorsed by defeasibility particularism (in addition to the form of holism without default reasons endorsed by unreconstructed particularism).
naturalism.\textsuperscript{11} For, on the view being explored, these principles express criteria for various moral terms (e.g. ‘cruel’) and criteria in the relevant sense help give the meaning of their respective moral terms. Thus, the unhedged moral principles that criterial particularism admits should be understood as analytic. Needless to say, endorsing analytic moral principles raises a number of issues, and I address the most pressing of these in Chapter Four. However, related to criterial particularism’s rehabilitation of analytic naturalism, two other issues also deserve mention.

The first of these, which I only want to flag here, concerns how the unhedged analytic moral principles that criterial particularism endorses fit into a broader moral semantics. As I emphasize in Chapter Four, criterial particularism only endorses unhedged analytic moral principles involving thick moral terms. Regarding thin moral terms, criterial particularism envisions some less restrictive but nevertheless conceptual constraints. This raises the obvious question: what are these (“less restrictive but nevertheless conceptual”) constraints and how do they relate to the conceptual constraints on thick moral terms represented by various unhedged analytic moral principles? I only flag this issue because the full implications of criterial particularism for moral semantics, while no doubt interesting and important, lie beyond the scope of this book.\textsuperscript{12} Still, to preview the general approach to this issue that I would pursue: criterial particularism

\textsuperscript{11} I say ‘at least partial’ because I believe that it is unclear whether such putative analytic naturalists as Foot and Anscombe ever intended anything more comprehensive than the sort of sufficiency conditions that criterial particularism endorses. If they did not intend any more than this, then – at least relative to their views – perhaps the ‘partial’ qualifier should be removed. I acknowledge, however, that my belief about how best to understand Foot and/or Anscombe may not be widely shared. So, I keep the ‘partial’.

\textsuperscript{12} Given its importance to criterial particularism’s rehabilitation of analytic naturalism, I would like to take up this issue elsewhere. Indeed, the issue of criterial particularism’s full implications for moral semantics seems like an excellent topic for future research.
should be understood as a species of what Susan Hurley calls “[n]on-centralism about reasons for action” according to which “the general concepts of right and ought are [not] logically prior to and independent of specific reason-giving concepts such as just and unkind.”

The other issue, which I raise in Chapter Four but do not directly answer there, is first suggested by the revelation, both in this Introduction and in Chapter Three, that criterial particularism rejects the holism about reasons espoused as “dogma” by orthodox particularism. But it is also highlighted in the title of Chapter Four: “… Criterial Particularism as Constitutive Generalism …” Thus, the other issue is: should my so-called ‘criterial particularism’ really count as a kind of particularism? On one hand, most proponents have assumed that particularism requires some form of principled holism, either with or without default reasons. On the other hand, some of its most prominent critics have assumed that particularism sits diametrically opposed to constitutive generalism, the view that knowledge of moral principles is (at least partially) constitutive of competence with moral concepts. My so-called ‘criterial particularism’, then, appears stuck between a rock and a hard place. It flouts the understanding of particularism offered by both major proponents and major critics.

Despite this – that is, despite how dubious it might appear to call the view developed in Chapters Three through Five a kind of particularism – the rough scheme just

---

14 I talk of ‘principled’ (or ‘across the board’) holism to contrast how criterial particularism allows that some moral reasons may need to be understood holistically but denies that all moral reasons should be understood in the same holistic way. Whether we call criterial particularism’s ad hoc position a form of holism matters less than whether we bear in mind how its position differs from that of orthodox particularism.
described for delimiting particularism gets it wrong. Using the holism (hurray) versus constitutive generalism (boo) dichotomy to say what counts as *bona fide* particularism is too quick, emphasizing relatively superficial features and missing what is more important.¹⁵ Let me explain. First, the principled (or across the board) holism that most proponents of particularism feel compelled to endorse should be seen as relatively superficial. Likewise, the opposition to moral principles – including opposition to moral principles that are unhedged, analytic, or both – should also be understood as relatively superficial. More fundamental – and what both holism about moral reasons and opposition to moral principles can be seen as manifestations of – is what might be called ‘the uncodifiability commitment’, a commitment to the effect that morality (or, less broadly, the meaning of moral concepts) does not admit of codification, not even in principle.¹⁶

Somewhat ironically, since they arguably misunderstand its real significance, Sean McKeever and Michael Ridge correctly pinpoint the *locus classicus* of particularism’s uncodifiability commitment. That is, they acknowledge the primacy of John McDowell’s claim that, “[T]o an unprejudiced eye it should seem quite implausible that any reasonably adult moral outlook admits of … codification” (where ‘codification’ is understood as capturing “how, in general, one should behave in principles apt for

---

¹⁵ I do not mean to suggest that the proponents and critics whose views I amalgamate to arrive at this rough classificatory scheme would necessarily find it acceptable. Still, this rough scheme does capture something important about the views of many who have thought and written about particularism.

¹⁶ For most purposes here, ‘codification’ can be interchanged with ‘reductive definition’ because I am primarily concerned with the prospects for reductively defining specific moral terms. Talk of ‘codification’ seems more specifically appropriate when talking about the prospects for reductively defining a network of moral concepts.
serving as major premises in syllogisms”). But McKeever and Ridge as well as others, including particularists like Jonathan Dancy and Margaret Little, miss the point that endorsing uncodifiability need not mean rejecting unhedged moral principles. Or, put another way, one can arrive at uncodifiability other than by adopting holism about moral reasons or by opposing unhedged moral principles. Indeed, I maintain in Chapter Four that even unhedged analytic moral principles are compatible with uncodifiability.

While Chapter Four takes this compatibility for granted, Chapter Five defends it. Focusing on thick moral terms – the moral terms that, according to criterial particularism, appear in unhedged analytic moral principles – the basic idea is this. Even unhedged analytic moral principles do not imply the codifiability of moral terms if (i) each such principle merely states some non-moral condition sufficient for a moral term’s application and (ii) all such principles involving the same term, taken together, fail to state a necessary condition of the term’s application. In short, codifiability (or better, definability) requires both sufficient and necessary conditions, but the unhedged analytic moral principles that criterial particularism endorses only yield the former, not the latter. Accordingly, I call this analytic naturalism ‘nondefinitive’. The unhedged moral principles that criterial particularism endorses link non-moral conditions with moral terms and these conditions conceptually constrain our use of the associated moral terms, but taken together these principles do not define the moral terms involved.

Of course, this is just the basic insight behind the view. The main question concerns the very idea of nondefinitive analytic naturalism. Does it make sense? That is,

---

can we recognize conceptual constraints on our use of moral terms yet claim (consistently) that such terms cannot be defined? The coherence of my so-called ‘criterial particularism’ depends on an affirmative answer. Moreover, if a commitment to uncodifiability fundamentally characterizes particularism, then an affirmative answer will also secure criterial particularism’s right to its name. Therefore, it is precisely what most critics and proponents of particularism miss that needs to be explained: the possibility of another route to uncodifiability, one that allows for non-moral conceptual constraints on moral terms – such as the constraints that various unhedged analytic moral principles represent – but rejects the possibility of non-moral definitions of moral terms. Exploring the possibility of such a route to uncodifiability – and, correspondingly, the very idea of nondefinitive analytic naturalism – frames the agenda of Chapter Five.

How, then, does this agenda lead into Wittgenstein’s rule-following considerations and an alternative account of moral reasoning? To see the connections here, we should first consider that naturalism in ethics, whether analytic or not, centrally involves something like the following claim: objective, non-moral features of the world rationally guide our use of moral terms. But what does the rational guidance envisioned by naturalism require? Specifically, how much structure must the relevant non-moral features of the world admit if they are to rationally guide our use of some moral term?

---

18 I take ethical naturalism to be a species of cognitivism (or realism) in contrast to, say, ethical non-naturalism. By this usage, however, I do not mean to impugn another, broader sense of ‘naturalism in ethics’ according to which most, if not all, noncognitivist (or expressivist) views count as naturalistic.

19 I am reminded here of the so-called Canberra argument against (unreconstructed) particularism put forward in Frank Jackson, Philip Pettit and Michael Smith, "Ethical Particularism and Patterns" in Moral Particularism, ed. Hooker and Little (2000), 79-99. Because criterial particularism endorses unhedged moral principles and, thereby, at least some structure (or patterns) at the non-moral level, it appears immune to this argument.
Even more specifically, for the relevant non-moral features to rationally guide us, must they be expressible as sufficient and necessary conditions or is it enough that they can be expressed as (a variety of) sufficient conditions? The very idea of criterial particularism, not to mention nondefinitive analytic naturalism, depends on the latter possibility, the possibility that rational guidance can be had from non-moral features expressible only as sufficient conditions for a moral term’s application. While the possibility of such guidance seems coherent enough – a legal analogue supports it\textsuperscript{20} – it has seldom, if ever, been explicitly recognized as a possibility for understanding how we use moral terms. For many philosophers, then, something appears to foreclose the possibility of nondefinitive rational guidance, the very possibility upon which criterial particularism depends.

The problem with the idea of nondefinitive rational guidance lies not with the guidance provided by various non-moral sufficiency conditions, at least not \textit{per se}. These conditions simply tell us when we must apply some thick moral term on pain of revealing conceptual incompetence with the term. So far, so good. The problem relates instead to the claim that all the sufficiency conditions together for one moral term do not generate a necessary condition of the term’s application. Or, what comes to the same thing, that we can only look to criterial cases – not definitions – for guidance in applying thick moral terms. Now these claims that criterial particularism involves admittedly do imply that the guidance provided by criterial cases and the sufficiency conditions they yield fails to determine a uniquely correct answer in every case as to whether the relevant moral term applies. In short, the distinctive moral semantics that informs criterial

\footnote{However, I must confess that avoiding the force of the Canberra argument was not a prime mover in my development of criterial particularism, only a more-or-less happy coincidence lately recognized.}

\footnote{I discuss this legal analogue in Chapter Five.}
particularism takes moral terms to have gaps.\textsuperscript{21} This is no small matter for \textit{inter alia} it means scaling back the ambitions of traditional moral theory.\textsuperscript{22} But – and here lies the problem with criterial particularism and its idea of nondefinitive rational guidance – these distinctive moral-semantic claims also supposedly imply that when uniquely determinative guidance runs out in certain cases (i.e. the hard ones), then all rational guidance whatsoever runs out, leaving it arbitrary in such cases how/whether we apply the relevant moral term(s). Ultimately, then, I take up Wittgenstein’s rule-following considerations and develop an alternative account of moral reasoning in Chapter Five to show that this supposed implication does not follow.

Turning to Wittgenstein’s rule-following considerations helps in two ways. First, they illuminate why the (false) inference from gappiness to arbitrariness can seem so compelling. Briefly, it is because we often assume an over-inflated picture of what genuine rule-following – genuine rational guidance – must involve, to wit, that it must involve rules that cover all possible cases of a term’s application independently of the reactions and responses involved in our learning and applying terms.\textsuperscript{23} But second,

\textsuperscript{21} Talk of ‘hard cases’ can also express the same point. That is, according to criterial particularism, the application of moral terms admits hard cases in the sense of occasions on which, apart from ordinary vagueness, the question of whether a term applies has no uniquely correct answer.

\textsuperscript{22} It also means that, in cases where no uniquely correct answer is determined, supervenience fails to hold.

\textsuperscript{23} This idea that rational guidance requires comprehensive rules formulable independently of our reactions and responses follows from what Alice Crary calls the “abstraction requirement” and the related idea that “the subjective (i.e. perceptual and affective) endowments we draw on in thinking and talking about how things are have an essential tendency to distort our view of reality.” Thus: “[I]t is … only by abstracting from such endowments that we can assure ourselves …” that how we apply some term is responsive to the way things are objectively and, correspondingly, guided by reason. Alice Crary, \textit{Beyond Moral Judgment} (2007) at 20-21.
properly understood, the rule-following considerations also show how we can (and arguably must) understand rules and rule-following differently.\(^{24}\) Understood in this less-inflated, more “realistic” way, the gappy moral semantics of criterial particularism implies not that our use of moral terms must be arbitrary but that such use can indeed be rationally guided.\(^{25}\)

Crary goes on to reject the abstraction requirement and urge what she calls a “wider conception of objectivity” according to which “our concepts … [are] essentially integrated into … customs or practices that instill us with the relevant sensibilities,” that is, the “sensibilities that we acquire in mastering … concepts[s].” Crary, Beyond Moral Judgment at 24-25.

Of the two main arguments that Crary gives against the abstraction requirement (and, by implication, in favor of the wider conception of objectivity with its view of concepts as essentially integrated with our “subjective” reactions and responses), the first makes use of Wittgenstein’s rule-following considerations much like the argument I give in Chapter Five urging similar conclusions. Her second argument against the abstraction requirement depends on a critique of the idea of literal sentence-meaning that she finds in J.L. Austin. Though it seems interesting and provocative, I hazard no opinion about the strength of this second argument. I merely observe that, to the extent it does have merit, it would provide independent grounds for some of the conclusions I urge in Chapter Five.

\(^{24}\) I include the ‘must’ because the picture of a comprehensive rule formulable – at least in principle – independently of the reactions and responses involved in our learning and applying some term is arguably a consoling fantasy, a mere \textit{dues ex machina}. Cora Diamond, "Realism and the Realistic Spirit" in The Realistic Spirit (1991), 39-72 at 39-42, 64-70. However, \textit{pace} Kripke, I argue in Chapter Five that skepticism about this consoling picture of rules and rule-following does not entail skepticism about rules and rule-following \textit{per se}. Granted, the case for this point is controversial and, given the response that others have received (e.g. Cora Diamond, John McDowell, Barry Stroud, and arguably Wittgenstein himself) when arguing roughly the same point, I do not expect to convince everyone. Indeed, my main goal in much of Chapters Three to Five, especially the sections on rule-following, might be better taken as trying to work out and set before my reader a more developed account of criterial (or Wittgensteinian) moral semantics than has hitherto been offered. Perhaps, with a more fully worked out account, some of the view’s merits may become apparent that might otherwise have been overlooked. Consider, for instance, the new way of addressing Moore’s open question argument discussed in Chapters Four and Five. In addition, though I do not discuss this issue much explicitly – another topic for future research – a criterial account of moral semantics arguably can provide a more realistic (as opposed to a consoling) account of moral agreement and disagreement.

\(^{25}\) For the relevant sense of ‘realistic’, see the essay by Cora Diamond cited in the previous footnote.
Interestingly, however, the rule-following sections in the *Investigations* focus on concepts such as ‘Add 2’ that, unlike moral concepts, do not admit hard cases, that is, cases where uniquely determinative guidance runs out.\(^2\) Therefore, even if (*per impossibile!*!) one were thoroughly convinced by Wittgenstein’s rule-following considerations that criterial particularism’s gappy moral semantics does not spell arbitrariness in *all* cases, one might still wonder how it explains reason’s role in hard cases. Indeed, if criterial particularism cannot say more about how reason guides our use of moral terms in hard cases – especially since, *ex hypothesi*, such guidance cannot be the more familiar, uniquely determinative kind – then, with some legitimacy, we might worry that arbitrariness creeps back in through the gaps that hard cases represent. But this is where Chapter Five’s alternative account of moral reasoning enters the picture. I intend the account developed in the last half of Chapter Five to, first and foremost, eliminate this basis for worrying about arbitrariness.\(^2\) Put another way, criterial particularism with its gappy moral semantics needs to show how, even if various non-moral conditions do not uniquely determine (i.e. tell us either ‘Yes’ or ‘No’) whether a moral term applies in certain cases, these same conditions nevertheless rationally constrain how we think about and decide whether the term applies in such cases. And the alternative account of moral reasoning developed in Chapter Five meets precisely this need.

\(^2\) In contrast to moral concepts which are ‘essentially contested’, Susan Hurley labels the concepts upon which Wittgenstein focuses ‘uncontestable’. Hurley, "Objectivity and Disagreement" at 81.

\(^2\) As I discuss in Chapter Five, the alternative account of moral reasoning developed there serves criterial particularism in two other important ways beyond the primary one I emphasize here. Also see the discussion in the Introduction above related to footnote 5.
CHAPTER ONE

Cognitive Pro Attitudes: An Explication and Defense of Anti-Humeanism

Imagine two people, Alice and Bob, both expert swimmers watching as a young child trips and falls into a pool apparently at risk of drowning. Alice sees the child and is immediately moved to help whereas Bob also sees the child but is utterly unmoved to do anything except keep sipping his pina colada. What is the difference between Alice and Bob? It seems natural to say that Alice and Bob see their situation differently. And this is exactly what the anti-Humeans I defend – neo-Aristotelian anti-Humeans\(^1\) – do say.

\(^1\) Contemporary anti-Humeans are either neo-Aristotelian or neo-Kantian. Both hold that some kind of thinking is internally related to motivation. For neo-Aristotelians, this thinking is experiential/quasi-perceptual; for neo-Kantians, it is the thinking constitutive of being rational. As the metaphors of “seeing things aright” and “moral vision” suggest, I focus upon and defend the neo-Aristotelian variety. Unless otherwise noted, references to anti-Humeans or anti-Humeanism should be understood accordingly.


For anti-Humeans, Alice and Bob’s difference in motivation amounts to a difference in their thinking, in how their situation appears to them. The question is whether such anti-Humean talk is just metaphorical. For many philosophers and social scientists, it is. According to an influential way of conceptualizing motivation – often called Humean, rightly or wrongly, in honor of David Hume – the difference between Alice and Bob is purely affective. Logically independent of any thoughts they have of their situation, Alice simply has a desire to help while Bob does not.²

In defense of anti-Humeans, I briefly review their basic position. Then, taking up the Humean orthodoxy, I examine why motivation is held to be logically independent of cognition, why motivationally-efficacious desires are supposed to logically independent of thought. One recently influential answer has it that only a logically independent desire can have telic direction of fit. After discussing (and rejecting) an interpretation of anti-Humeanism that has tempted some, I grant that motivating states must possess telic direction of fit but argue that not all states possessing telic direction of fit are logically independent desires. To this end, I distinguish cognitive from noncognitive pro attitudes, and defend the possibility of the former. Both cognitive and noncognitive pro attitudes possess telic direction of fit, but with the former the pro attitude is an essential aspect of a certain kind of thinking. Thus, by conceiving motivation in terms of cognitive pro attitudes, anti-Humeans can reject the demand to analyze motivation into an exclusively affective (as opposed to cognitive) state. Finally, I briefly suggest that the Humean

² Note that I do not intend ‘motivation’ in what Al Mele refers to as the success sense. That is, I take it that being moved to \( \phi \) does not entail that one actually \( \phi \)’s. Alfred R. Mele, "Motivation: Essentially Motivation- Constituting Attitudes," *Philosophical Review*, 104 (1995): 387-423 at 387.
demand to segregate motivation from cognition must be rejected on the grounds that only by integrating them – as in cognitive pro attitudes – can we rationally explain action.

1. Anti-Humeans are distinguished by their combination of realism about the deliverances of practical reason and internalism about motivation. In other words, they hold (i) that what I have reason to do is an objective question and (ii) if I grasp what I have reason to do (say φ) in a certain way, then I will be motivated to φ in virtue of this understanding. Since opponents of anti-Humeanism adhere to Humean psychology – the view that motivation involves logically distinct cognitive and motivational states – they have to choose between realism about practical reason and internalism about motivation. In this chapter, I take for granted the objectivity about reasons presupposed by anti-Humeans and show how, given this objectivity, anti-Humeans can still be internalists. On Stephen Darwall’s helpful map of the different varieties of internalism about motivation, anti-Humeans appear as those who believe that actual consciousness of or cognitive contact with reason facts necessarily motivates. But what sort of

---

3 I discuss the objectivity of moral concepts in Chapter Five.
4 Michael Smith has laid out these meta-ethical positions and their interrelationships in terms of a trilemma which he calls “the moral problem.” Michael Smith, *The Moral Problem* (1994) at 12-13. See also McNaughton, *Moral Vision* at 23.

Darwall distinguishes between rational intuition and sensibility versions of non-constitutive existence internalism, and he classifies (neo-Aristotelian) anti-Humeans as sensibility theorists. Darwall’s scheme can mislead. On one hand, it suggests that anti-Humeans are latter-day moral sense theorists, the historical opponents of rational intuitionists. On the other hand, it suggests that anti-Humeans and rational intuitionists differ more than they do. In fact, anti-Humeanism tries to combine the virtues of both
understanding could this “consciousness” or “cognitive contact” be such that it necessarily motivates?

2. Perhaps justifiably, Humeans do not recognize any real challenge to their psychology in the metaphorical talk of anti-Humeans. “What has to be explored,” Mark Platts observes, “is the ground for the persistent manœuvre [by Humeans] of reading a desire into any apparent counterexample ...”6 Platts continues:

The ground for that manœuvre must have at least this structure: there is some condition upon a motivating specification of a reason for action which is met only by the introduction of reference to desire ...7

In other words, if Humeans are justified in reading a desire into every apparent counterexample to their psychology, there must be something about motivating reasons that requires the presence of a desire. This suggests two questions:

(i) What is required to have a motivating reason?, and
(ii) What is it about desires that satisfy these requirements?

To answer these questions, we can begin with Michael Smith’s defense of Humean psychology.8

On Smith’s view, there is a condition upon motivating reasons that only desires can fulfill. Having a motivating reason involves having a goal. Only desires can fulfill this condition because having a goal means being in a state with which the world must fit rational intuitionism and moral sense theories, thus resembling both views in certain respects. In this chapter, I defend the possibility of such a combination.

6 Platts, "Moral reality and the end of desire" at 74. Since I want to present the Humeans’ case for their psychology on their own terms, ‘desire’ in this passage and elsewhere should be understood as desire in a logically independent sense unless stated otherwise.

7 Ibid at 74.

8 Smith, The Moral Problem at 92-129.
and only desires have this direction of fit. Indeed, according to Smith, ‘desire’ just means a state with which the world must fit. He summarizes his argument:

(1) Having a motivating reason is, *inter alia*, having a goal

(2) Having a goal is being in a state with which the world must fit, and

(3) Being in a state with which the world must fit is desiring.9

Thus he concludes that the various states with which anti-Humeans would ostensibly replace desires must really be desires themselves – that is, states with which the world must fit – or else they cannot embody having a goal and cannot motivate our actions.

The teleological nature of action explanation and, hence, Premise (1) should be uncontroversial. Action explanations are unabashedly teleological. We come to understand why someone acted as she did by learning what she was trying to achieve. Since we cite motivating reasons to explain an agent’s actions, they must reveal her end, purpose, or goal.10 Consequently, we should turn to the notion of direction of fit that figures in Premises (2) and (3).

Direction of fit is a property of mental states that specifies one way in which mental states relate to their contents.11 Direction of fit comes in two varieties: thetic and

---

10 However, Premise (1) is uncontroversial only if we do not read too much into the meaning of ‘teleological’. As Richard Norman has correctly stressed, “[A]lthough we may allow that the primary kind of explanation of human action is teleological, this in effect only means that such explanations are in terms of the agent’s reasons, and does not entail that those reasons are themselves ‘teleological’.” Richard Norman, *Reasons for Actions: A Critique of Utilitarian Rationality* (1971) at 89.
11 The *de re/de dicto* distinction as applied to propositional attitudes captures another difference in how mental states relate to their contents, one which is also important for the view defended in this chapter.
telic. In Smith’s terminology, there are states which must fit the world and states with which the world must fit. Belief has been considered a paradigm of the former; desire a paradigm of the latter. As Platts describes the difference:

Beliefs aim at the true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit with the world, not vice versa. Desires aim at realisation, and their realisation is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing in the desire, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa.

In other words, beliefs exhibit thetic direction of fit in the sense that when we consider our beliefs vis-a-vis the world, we start with the world and see whether our beliefs measure up. By contrast, desires exhibit telic direction of fit in the sense that when we consider our desires vis-a-vis the world, we start with our desires and see what has to be done for the world to measure up.

The thought behind Premise (2) should now be clear. For a mental state to embody having a goal, it could not be a state with an exclusively thetic direction of fit. A goal is not something -- like a belief -- that one gives up if the world does not correspond to it. Just the opposite, the whole point of a goal is to provide a benchmark for changing the world. Having a goal is something that, by its very nature, persists when the world does not measure up. Therefore, having a goal can only be embodied by states with telic direction of fit. Thus it appears that the anti-Humean position must be wrong, at least so

---


long as we accept Premise (3), that is, Smith’s definition of ‘desire’ as a state with telic direction of fit. Despite how things appear, I argue that Smith’s argument moves too quickly and that, in fact, Premise (3) is false. But first I consider (and reject) one way that anti-Humeans might try to avoid the force of Smith’s argument.

3. Smith may save the Humean dogma but, according to some, he does so only by emptying it of content.\(^{14}\) Let pro attitudes be those mental states with telic direction of fit. Focusing upon the broad, pro attitude sense of ‘desire’, Smith renders the Humean dogma trivially true but, in so doing, passes over the phenomenological sense of ‘desire’ that really matters. If, contra Smith, we understand the Humean dogma in terms of the phenomenological sense of ‘desire’, anti-Humeanism is vindicated. Or so goes this initial response to Smith.

Recently, Fred Schueler has developed this line of argument.\(^{15}\) First, he distinguishes:

[T]wo senses of the term “desire”: On the one side [pro attitudes] ... the philosophers’ sense ... in which desires are so to speak automatically tied to actions because the term “desire” is understood so broadly as to apply to whatever moves someone to act. On the other side [desires proper] ... the more ordinary sense in which one can do things one has no desire to do, that is, the sense in which one can reflect on one’s own desires, try to figure out what one wants, compare one’s own desires with the desires of others or the requirements of morals, the law, etiquette or prudence, and in the end, perhaps, even decide that some desires one has, even very strong ones, shouldn’t be acted on at all. (1)

Thus, on Schueler’s view, the Humean dogma “suffer[s] from a deep ambiguity in terms such as ‘desire,’ want,’ and their cognates.” (6) Given his characterization of pro

\(^{14}\) Mark Platts suggests such an objection when he writes, “Nothing but muddle (and boredom) comes from treating desire as a mental catch-all.” Ibid at 294.

\(^{15}\) Accordingly, the numbers in parentheses in the main text refer to G.F. Schueler, *Desire: Its Role in Practical Reason and the Explanation of Action* (1995).
attitudes and desires proper, desires proper represent the contentful sense of ‘desire’ and, when interpreted as being about desires proper, the Humean dogma is false. Schueler concludes that stating the Humean dogma using his terms “greatly changes [it], and for the worse.” (6)

This general sort of argument hinges upon the characterization of pro attitudes and desires proper. Pro attitudes are supposed to represent “the sense of ‘desire’ in which it is analytic that if I do something intentionally then I want (or have a desire) to do whatever it was I was trying to achieve.” (34) Pro attitudes are conceptually related to actions because they refer to whatever state moves us to act. Desires proper, on the other hand, are just a subset of the states that move us to act. They “include such things as cravings, urges, wishes, hopes, yens, and the like.” (35) Desires proper, unlike pro attitudes, are supposed to be real existences that we can reflect upon in deliberations and decide whether to be moved to action by them.

This last contrast reveals why the Humean dogma is supposedly best understood in terms of desires proper and why, when understood in this way, the dogma comes out false. Take the latter point first. If the Humean dogma is understood as being about desires proper, then desires are not essential to the explanation of action. For some actions could be explainable in terms of other considerations that we reflect upon in deliberation besides desires proper, say, moral and political beliefs. To use Schueler’s example, I might attend a meeting at my child’s school not because I feel some desire to attend – indeed, if I did what I felt, I would stay home – but because I believe it to be my responsibility to my child and my community.
But shouldn’t we be charitable and interpret the Humean dogma as being about pro attitudes? Consider what Schueler says about pro attitudes:

> [T]he sense of “desire” that ... I have been referring to is merely a kind of placeholder sense ... If I do something intentionally, then it follows that I wanted to do it. But we should not think that there is some psychological state called a “desire” or “want” thus referred to. There is just whatever moved me to do what I did. (34)

Later, after quoting Don Locke’s argument that pro attitudes cannot cause behavior since they are not real existences, Schueler concedes:

> That an agent ... had a pro attitude toward performing the act ... tells us nothing about the role of the pro attitude in explaining the action. In particular, it doesn’t tell us that the pro attitude (even if we assume there is something to be referred to in this way) caused the action ... (37) (emphasis in original)

But Schueler adds:

> [I]t might still turn out that all the (perhaps extremely various) sorts of “things” that lead people to act, and hence justify speaking of pro attitudes toward some goals, are causes of the actions to which they lead. (38)

These passages explain what Schueler means by referring to pro attitudes as mere placeholders and, correspondingly, why pro attitudes are disfavored for understanding the Humean dogma. Pro attitudes are not supposed to pick out one kind of thing. In other words, the states that move people to act (and justify ascribing pro attitudes) are quite heterogeneous. These various states supposedly have nothing in common except the fact that they move people to act. But then ascribing pro attitudes conveys no new information. And it is their empty formality that explains why reading the Humean dogma as about pro attitudes is disfavored. Reading the Humean dogma as about pro attitudes reduces it to a tautology. Therefore, on what appears to be the only non-trivial
(and hence acceptable) interpretation – the desire proper interpretation – the Humean dogma is false.

Is this general sort of argument sound? No. While it is certainly true, as Schueler observes, that desires proper are not essential to the explanation of action, the banality of this truth mitigates against reading the Humean dogma in terms of desires proper. To read the Humean dogma in this way would allow anti-Humeans too cheap a victory. More importantly, Humeans can plausibly claim that pro attitudes are not mere placeholders. On Smith’s view, as we have seen, pro attitudes have something in common besides the empty acknowledgment that something has moved an agent to act. Pro attitudes possess telic direction of fit. However, the question remains whether allowing that pro attitudes move us to act suffices to make Humean psychology true.

4. The two possible answers to this question frame the Humean/anti-Humean debate as it is best understood. Recognizing that pro attitudes move us to act, Humeans believe, suffices to save their psychology. Anti-Humeans deny this. Their disagreement on this point should be understood as a more fundamental disagreement. Although Humeans and anti-Humeans agree that pro attitudes move us to act, they differ regarding the nature of pro attitudes. Acknowledging that pro attitudes move us to act says nothing about how pro attitudes are logically related to states possessing thetic direction of fit. Given that states with thetic direction of fit are cognitive states and that Humean psychology requires a distinct, noncognitive state for motivation, Humeans hold that the motivation constitutive of pro attitudes is logically independent of states with thetic direction of fit. But anti-Humeans need not accept this. Indeed, on their view, pro attitudes are an essential aspect of a certain sort of cognition.
Call these two conceptions noncognitive pro attitudes (NCPA’s) and cognitive pro attitudes (CPA’s), respectively. They should be understood as follows:16

an agent $A$ has a NCPA $=Df$

$A$ is disposed to $\phi$ in certain conditions $C$

(except where ‘certain conditions $C’ = ‘certain conditions $C^*$ in virtue of $A$’s perceiving, or its appearing to $A$, that she has reason to $\phi$)17

while

an agent $A$ has a CPA $=Df$

(i) $A$ is moved to $\phi$ in some particular situation $\chi$ in virtue of $A$’s perceiving, or its appearing to $A$, that she has reason to $\phi$ in $\chi$;

or,

(ii) $A$ is disposed to $\phi$ in certain conditions $C$ in virtue of $A$’s perceiving, or its appearing to $A$, that she has reason to $\phi$.18

---

16 I use the term ‘reason’ to mean an all-things-considered reason or, as some would say, a conclusive reason. To avoid confusion in this context, I use ‘consideration’ or ‘feature’ instead of ‘prima facie reason’.

17 Having a NCPA is not just a matter of being disposed to do something in certain conditions for the following reason. If anti-Humeans are correct and, necessarily, we are moved to act by CPA’s, it would be analytically true – in virtue of the thinking involved in CPA’s – that each of us is disposed to do whatever she has or appears to have reason to do whenever she perceives or it appears to her that she has reason to do something. Far from noncognitive, this disposition logically depends on a certain kind of cognition. To exclude this kind of disposition, the definition of NCPA’s includes the second clause beginning with ‘it is possible that’.

18 I do not say much about the second disjunct in the definition of CPA’s because everything I have to say about the first applies to the second as well. We need the second disjunct because, of course, we are not always moved to act by what we have reason to do. Rather, we are sometimes – perhaps often – taken in by appearances, by what it merely appears that we have reason to do. Indeed, we may be faced with particular situations in which more than one action appears to be what we have reason to do. While I do not develop this suggestion here, I believe that anti-Humeans can account for weakness of will and related phenomena in terms of such conflicting appearances.

A disjunctive account also enables anti-Humeans to claim that, when an agent’s situation and what she has reason to do appears to her veridically, what moves her to act
In contrast to NCPA’s, CPA’s are thoughts that essentially move us to act in a certain way.

Humean psychology is best understood as about noncognitive pro attitudes. That is, according to Humean psychology, noncognitive pro attitudes are necessary for the explanation of action. Anti-Humeans distinguish themselves by denying this. By contrast, they claim that cognitive pro attitudes are at least sufficient for the explanation of action. Whether anti-Humeans go further and claim that CPA’s are necessary for the explanation of action is a question that I abstain from addressing here.

We should now be able to see why Smith’s argument in defense of Humean psychology moves too quickly. Recall that he argues:

(1) Having a motivating reason is, inter alia, having a goal

(2) Having a goal is being in a state with which the world must fit, and

(3) Being in a state with which the world must fit is desiring.

Smith’s argument neglects the fact that being in a state with which the world must fit could be desiring in either of two senses: having a noncognitive pro attitude or having a cognitive pro attitude. Once Premise (3) is revised to reflect this distinction, Smith’s argument offers no support for Humean psychology over the anti-Humean alternative.

is nothing short of the fact itself, namely, the fact that she has reason to do such-and-such. One reason we might want to make such a claim would be to defend practical reasons as external in the sense that Bernard Williams argues against. Bernard Williams, "Internal and external reasons" in Moral Luck (1981), 101-113. For an analysis of Williams’ argument that suggests the need for such a claim, see Dancy, "Why There Is Really No Such Thing As The Theory of Motivation."

5. Most Humeans, including Smith, are unlikely to give up so easily. Fundamentally, they object that CPA’s are incoherent and, hence, Smith’s argument is sound. Right away, it must be conceded that, given certain widely-held views about mental states and their individuation, the very idea of CPA’s – of logically bridging the gap between cognition and motivation and uniting them in one state of mind – can seem problematic. Simon Blackburn puts the problem facing anti-Humeans this way:

‘Representation’ suggests one direction of fit with the world, while ‘motivational state’ suggests the other . . . [I]t is not just a question of inflating a term like ‘representation’ to cover ethical commitment, but of showing that one has a right to do so given the different functional essence of the state expressed . . .

In other words, cognition and motivation have different functional roles corresponding to their different directions of fit. Cognitive states present (or re-present) the world to their bearer as being the case that $p$ whereas motivational states move their bearer to change the world so that it is the case that $p$. The difference in direction of fit amounts to a difference in how these states counterfactually depend upon the fact that $not-p$: cognitive states that $p$ tend to go out of existence while motivational states that $p$ tend to endure. But nothing about this intuitive functionalism rules out a mental state that, by its very nature, plays both a cognitive and a motivational role. Indeed, consistent with the preceding analysis, CPA’s tend to be absent if it is not the case that their bearer has reason to do something in her present situation and, presupposing that their bearer does

---

have reason to do something in her present situation, CPA’s tend to be present if it is not yet the case that their bearer has done that thing.\(^{20}\)

Still, a theoretically more ambitious functionalism would rule out CPA’s, and presumably something like this stronger view animates the concerns of Blackburn and others. This view – nonnormative functionalism we might call it – holds that mental states stand in merely causal relations to their respective contents. Eschewing all resources with which one might individuate mental states except causal relations, nonnormative functionalism would individuate mental states according to the causal relation that determines a mental state’s content and direction of fit: thus, different causal relation, different mental state. Being what they are -- that is, states with different directions of fit -- any possible pair of cognitive and motivational states will necessarily involve two different causal relations. Therefore, according to nonnormative functionalism, any possible pair of cognitive and motivational states will necessarily be distinct mental states, leaving no logical space for CPA’s.

For CPA’s to be possible, then, nonnormative functionalism must be false. Cognition and motivation must stand in something other than causal relations to their respective contents and each other. While showing that nonnormative functionalism is false goes well beyond the scope of this chapter, such a showing might begin with the inability of nonnormative functionalism to make room for various kinds of thinking that

\(^{20}\) Michael Smith’s skepticism regarding the coherence of CPA’s is unwarranted because it rests on the dubious view that a state with both directions of fit must stand in the thetic relation and the telic relation to the very same propositional content. Smith, *The Moral Problem* at 118. Certainly, if this were true, CPA’s would be incoherent. But, as Maggie Little has observed, nothing supports – much less requires – this dubious view. Margaret Olivia Little, "Virtue as Knowledge: Objections from the Philosophy of Mind," *Nous*, 31 (1997): 59-79 at 63-64. Moreover, the definition of CPA’s should make it clear that they are directed thetically towards one content and telically towards another.
depend not merely causally but also logically upon the existence of their respective objects, for example, singular thoughts. With this suggestion in mind, I explore the kind of thinking and motivation that anti-Humeans need in order for cognition to essentially involve motivation as envisioned by CPA’s. Exploring these things should not only undermine confidence in nonnormative functionalism and show CPA’s to be plausible but also reveal some interesting features of the anti-Humean position.

6. My explanation and defense of CPA’s begins with the motivation they involve. Compare NCPA’s first. NCPA’s involve an agent’s being moved in virtue of her possessing a bare disposition. In terms of another way of individuating motivational states – apart from how they logically relate to cognitive states – NCPA’s involve being moved de dicto. In other words, when we ascribe to an agent A the disposition to φ in certain conditions C (where ‘φ’ and ‘C’ are specified de dicto), we commit ourselves to nothing more than a causal relation between A’s being in C and A’s φ-ing (or between A’s being in C and A’s having a favorable response towards φ-ing). Thus, for example, the fact that whenever someone eats a large midday meal, we regularly find him wanting to nap suffices for his possessing an NCPA towards napping. Accordingly, being moved de dicto is not essentially a response to any particular situation and, without more, need not be regarded as a conscious, thinking response. Indeed, just as a child’s skin might be

21 I say ‘bare disposition’ by way of contrast. For while the anti-Humean picture of CPA’s certainly grants that an agent can be disposed to do various things, the dispositions to act that the anti-Humean picture allows do not hold ‘barely’ (or simply, or in virtue of nothing else) but only in virtue of an agent’s perceiving, or its appearing to her, that she has reason to act accordingly. See also note 16.

22 As Jaegwon Kim puts it, “. . . a mere causal connection isn’t enough; causal connections can be cognitively blind . . .” Jaegwon Kim, "Perception and Reference Without Causality," Journal of Philosophy, 74 (1977): 606-620 at 615.
disposed to burn in the summer sun without the child understanding what this disposition involves, so even an agent with greater conceptual sophistication and, correspondingly, greater capacity for understanding could possess an NCPA towards $\phi$-ing in $C$ without understanding (or being able to specify) what she is disposed to do or the conditions in which she is disposed to do it. Simply because one bears a disposition that can be specified *de dicto* – such as an NCPA – does not necessarily place the bearer in any priviledged position vis-à-vis others when it comes to understanding (or specifying) the disposition. Rather, because such dispositions are merely causal phenomena, third party observers may be just as well, or better, placed than their bearers to understand (or specify) these dispositions. Therefore, since an agent’s being moved *de dicto* does not necessarily presuppose any occurrent understanding by the agent of what she is disposed to do – i.e. since being moved *de dicto* does not essentially involve cognition – CPA’s must involve another way of being moved.

Unlike NCPA’s, CPA’s involve motivation conceived of as a conscious, thinking response by an agent to what she has reason to do. As we have seen, however, if someone responds to a situation merely because it has certain features specifiable *de dicto* and not because it is the particular situation it is, the response involved could just be a matter of a causal relation holding. Therefore, to ensure that the motivation involved in having a CPA is essentially a conscious, thinking response, the agent having a CPA must be responding to some particular situation. In other words, CPA’s involve being moved *de re*. Being moved *de re* ensures a conscious, thinking response because only by

---

24 Strictly speaking, the motivation involved in CPA’s is a response by an agent to what she has *or appears to have* reason to do. But since everything I have to say about the first disjunct in the definition of CPA’s applies to the second, I omit the ‘or appears to have’ out of convenience.
consciousness engaging with her surroundings can someone keep track of a particular and
the ability to keep track of a particular is necessary for responding essentially to it.25

There might seem to be a problem in the requirement that, to have a CPA, one
must be moved essentially in response to a particular situation while one is primarily
supposed to be moved in response to what one has reason to do. These two demands can
appear inconsistent. But, given particularism, they are not. According to particularism,
what someone has reason to do depends essentially on her particular situation.26 This
helps explain why anti-Humeans are particularists. For if particularism is true, when
someone responds to what she has reason to do, she is responding essentially to a
particular situation.27 So, to restate what we have learned about CPA’s: they involve

25 Gareth Evans, *The Varieties of Reference* (1982) at 145-150. The conception of *de re*
attitudes involved in CPA’s and articulated by Gareth Evans is more demanding than the
purely causal conception advocated by many, including David Lewis. David Lewis,
On this more demanding conception, thinking of a particular is neither a matter of merely
causal “acquaintance” nor of unmediated Russellian acquaintance. (For a defense of the
latter as against the former, see Kim, "Perception and Reference Without Causality,"
especially 618-620.) Instead, thinking of a particular involves thinking of the thing, the
referent, under a particular mode of presentation that enables us to keep track of it. That
is, for Evans, demonstratives have sense. The same goes for indexicals. For Evans’
account of thinking about one’s own particular situation as ‘here’ and thinking about
oneself as ‘I’, see Evans, *The Varieties of Reference* at 151-170, 205-270. Also see note
26.

26 Particularism also holds that what an agent appears to have reason to do depends
essentially on her particular situation. So, again, everything I have to say about the first
disjunct applies to the second.

27 I defend particularism in Chapters Three to Five. Here I merely suggest how and why
anti-Humeans rely on it. I will add this: one line of argument for particularism begins
with the holistic way in which the (potentially) reason-giving features of a situation
interact. In other words, the weight (if any) that some feature of a situation carries in
determining what an agent has reason to do in that situation depends upon what other
features are present. This holism means that what an agent has reason to do in some
situation does not depend upon certain features of the situation taken in isolation and
motivation conceived of as a conscious, thinking response by an agent to what she has reason to do where what she has reason to do depends essentially on her particular situation.

7. Turning to the cognition involved in CPA’s, the question arises: what kind of thinking essentially involves motivation? We already have a partial answer. Thinking about what one has reason to do – where this thinking refers essentially to one’s particular situation – is supposed to essentially involve motivation. Together with what we know about the motivation involved, this yields a necessary condition for having a CPA. The thinking involved must essentially be about the same particular situation to which the motivation involved is essentially responding. Otherwise, the motivation need not accompany the thinking. But this is not yet a sufficient condition for having a CPA. We still lack an explanation for why thinking about what one has reason to do necessarily involves motivation.

The explanation can be stated simply. The thinking involved in CPA’s is not only de re but also first-personal. Specifically, it is an agent’s first-person understanding of capable of being repeated in other situations. Rather, it depends upon the situation itself and all of its features. Dancy, *Moral Reasons* at 60-62.

28 Like the thinking involved in understanding indexicals that Evans and others refer to as ‘egocentric’. Evans, *The Varieties of Reference* at 151-157. They claim that, to properly understand an indexical such as ‘I’, someone must know more than just the rule that fixes its reference. In addition, a special kind of thinking is involved – first-personal or egocentric thinking – that constitutes its sense. To illustrate this point, consider an example courtesy of John Campbell and Star Trek. Campbell asks us to imagine a people (like the Borg on *Star Trek*) who:

[T]hough intelligent and possessed of a shared language, [they] nonetheless have no use for the first person. They have no such sign in their language. And . . . they also fail rudimentary behavioral tests for self-consciousness. We can say to them that they ought to consider introducing the term ‘Id’ into their language, which should be governed by the rule that any token of it refers to whoever produced it. And they might indeed fall in with this proposal. They might
certain ‘I’-thoughts – thoughts ascribing to herself the property of having reason to act -- that necessarily involves motivation. An example may help. Imagine a security guard regaining consciousness after having dozed off. (To be fair to security guards, perhaps he was drugged.) Still half asleep, he glances at the closed-circuit television that displays his security post and the surrounding terrain from a vantage point slightly above and behind his post. His glance discloses a man inside a glass booth being approached threateningly by a group of heavily armed men, and he thinks that man has reason to call for help. However, because he is still half asleep, our security guard fails to realize that he himself is that man. Consequently, the thought he would express as ‘I have reason to call for help’ does not occur to him and, correspondingly, he lacks any motivation to call for help. Later, when our security guard sees armed men outside his post, he grasps in the first-person that he himself has reason to call for help and, without need of any further thought, is motivated accordingly. This example shows that someone’s thinking de re – or at least quasi-de re\(^{29}\) – of herself, her situation, and what she has reason to do – need not involve motivation. Rather, for someone’s thinking to necessarily involve motivation, she must grasp in the first-person that she herself has reason to act.

Regarding this crucial notion of a first-person understanding of ‘I’-thoughts, consider again our security guard. But now suppose when he sees that man on the

---

introduce the term and stipulate that it is governed by that rule. But this will not render them self-conscious. The modes of self-knowledge that we have might still be a mystery to them. John Campbell, *Past, Space, and Self* (1994) at 111. Again, this more demanding view of Evans and Campbell contrasts with the view of those such as David Kaplan David Kaplan, "Demonstratives" in *Themes from Kaplan*, ed. Almog, Perry and Wettstein (1989), 481-563.

\(^{29}\) I say ‘quasi-de re’ because thinking of someone on television as that man is, strictly speaking, a hybrid of purely demonstrative and purely descriptive thinking. Evans, *The Varieties of Reference* at 150-151.
television, he realizes that he himself is *that man* and, therefore, he accepts the thought that he himself has reason to call for help. Has he grasped this thought in the first person? No. In contrast to his first-person grasp of this thought when he sees armed men outside his post, our security guard cannot grasp this thought in the first-person by seeing himself on the television. ‘I’-thoughts consist in self-ascribing some property, but this can be done in two very different ways. One way is immune to error through misidentification; the other is not. Having a first-person grasp of an ‘I’-thought means self-ascribing a property in the way that is immune to error through misidentification.30 Returning to our security guard, his self-ascription of ‘having reason to call for help’ on the basis of seeing himself on the television is not immune to error through misidentification. As we have seen, his self-ascription of this property on this basis requires a logically independent, self-identifying belief, and this belief could be false. Achieved in this way, our security guard’s understanding cannot be first-personal, at least not in the required sense.

But compare our security guard’s understanding when he sees armed men outside his post. In this case, his understanding that he himself has reason to call for help is first-personal because it is perceptual. Properly understood, perceptual thinking essentially involves a perceiver’s grasp of her particular situation as ‘my situation here and now’, and her thinking of her particular situation in this way identifies her as the occupant of a

30 For the idea of immunity to error through misidentification, see Ibid and Sydney Shoemaker, "Self-Reference and Self-Awareness" in *Self-Knowledge*, ed. Cassam (1994), 80-93. Anti-Humeans need to side with Evans, contra Shoemaker, that it is not just with self-ascriptions of mental properties that we enjoy this immunity to error. We can also self-ascribe non-mental properties in this way. Evans, *The Varieties of Reference* at 216-217. This is important because, for anti-Humeans, ‘having reason to φ in χ’ is a non-mental property (or, at least, a wide/external mental property).
unique spatio-temporal perspective on the world. In other words, a kind of self-identification is built right into perceptual thinking. Thus, any property that can be grasped perceptually can be self-ascribed in the way that is immune to error through misidentification. And so it is that our security guard, based on his perceptual grasp of his particular situation, grasps in the first person that he himself has reason to call for help.

Knowing what we do now – that grasping a fact about oneself in the first-person means self-ascribing a property in a certain, special way – gives us the resources to explain why the thinking involved in CPA’s necessarily involves motivation. To paraphrase Stephen Darwall, what anti-Humeans seek is the kind of understanding of practical reasons that will necessarily grip us. Like Darwall, anti-Humeans want to discredit the kind of skepticism expressed by questions like ‘Why should I do that?’.

Unlike Darwall, they discredit this kind of skepticism – motivational skepticism – by logically excluding it (given, of course, the right kind of understanding). Consider an analogy. Someone grasps in the first-person that she herself is standing in front of the Sears Tower. This means that her self-ascription of ‘standing in front of the Sears Tower’ is immune to error through misidentification and it literally makes no sense for her to ask ‘But am I standing in front of the Sears Tower?’ Similarly, then, if someone grasps in the first-person that she herself should do something, it literally makes no sense for her to ask ‘But should I do that?’.

Thus, anti-Humeans discredit motivational

skepticism by drawing our attention to a special kind of understanding – first-person understanding – that logically excludes such skepticism, and this is the kind of understanding involved in CPA’s.

8. To this point, my explanation and defense of CPA’s has taken for granted the objectivity of reasons that anti-Humeans claim. However, by saying something about one aspect of CPA’s that has so far escaped comment, at least one worry about objectivity can be defused. Besides, I should say something about this aspect of CPA’s anyway. But first the worry.

Recall Darwall’s map of the different varieties of internalism. The worry arises when we juxtapose the internalism of CPA’s with what Darwall calls judgment internalism.33 The worry is that the internalism of CPA’s (according to which the perception of reason facts necessarily motivates) collapses into judgment internalism (according to which sincere assent to a reason judgment is necessarily connected to motivation).34 If this worry were well-founded, it would suggest that CPA’s should be understood along projectivist lines as merely quasi-cognitive.35 However, this worry only arises if perceiving is presumed to be a matter of believing.36 While perceiving often goes together with believing since the former is often our basis for the latter, they

33 Darwall, "Internalism and Agency". Again, my characterization of the different internalisms is a broader one in terms of not just moral reasons but all-things-considered reasons.
34 Jamie Dreier raised this worry in a seminar of his on Darwall and internalism.
36 In the discussion that follows, I do not mention being appeared to because, unless otherwise noted, everything I say about perceiving goes for being appeared to as well.
are distinct. This distinction is reflected in the aspect of CPA’s that has so far escaped comment.

CPA’s implicitly recognize the distinction between perception and belief because, necessarily, an agent is moved to act if she perceives that she has reason to act and does not disbelieve this, regardless of whether she believes it. Consider our security guard again but now suppose that when he sees armed men outside his post (who are, in fact, enemy soldiers), he is not sure what to believe. Let’s say, he received notice that a group of friendly soldiers would be having night training exercises in his sector. He knows that the men he sees could be the friendly soldiers he was told about. Then again, they could be enemy soldiers as, in fact, they are. But until they get a little closer, he cannot tell for sure. So, while he neither believes nor disbelieves the fact that he has reason to call for help, it makes sense that he nevertheless feels some motivation to call for help given what he sees. And this is just what CPA’s would have us say about this scenario, thus belying the worry that, even on their own terms, CPA’s really reflect judgment internalism.

9. Until now, I have primarily been on the defensive, explaining what CPA’s must involve to be possible. I have not tried to justify the anti-Humean claim that, necessarily, CPA’s move us to act. While the entire positive case for the anti-Humean alternative to Humean psychology cannot be presented here, I offer a glimpse of one element in such a case.

37 As I said, the same thing goes for being appeared to as for perceiving. Just imagine the same scenario involving our security guard except for the fact that, in this case, the armed men outside his post are the friendly soldiers he was told about.
Both Humeans and anti-Humeans agree that whatever moves us to act, it must be a psychological state. John McDowell, for example, acknowledges that explaining actions in terms of reasons means crediting the agent with some psychological state. However, not just any psychological state will do. As McDowell points out, for a psychological state to explain an agent’s action, it must “reveal the favourable light in which the agent saw his projected action.” This is a kind of intelligibility requirement. For a psychological state to explain why an agent does something (or attempts to do something), citing the state must make it rationally intelligible why the agent was motivated to do that thing. This intelligibility requirement yields an argument in favor of anti-Humeanism.

Let ‘ϕ-ing’ stand for ‘doing something’. The argument goes: for any agent whose ϕ-ing (or attempted ϕ-ing) we are trying to explain, CPA’s do, while NCPA’s do not, make it rationally intelligible why the agent was motivated to ϕ in the first place. Citing an agent's CPA makes this rationally intelligible because a CPA essentially refers to the fact that she had or appeared to have reason to ϕ by way of her perceiving or being appeared to accordingly. For example, referring to an agent’s perception -- say, that she has reason to avoid being mauled by a bear -- serves to make her motivation -- to avoid being mauled by a bear -- rationally intelligible. By contrast, citing an agent’s NCPA

38 Humeans have expressed concern about anti-Humeans on this score. See, for example, Smith, The Moral Problem at 96 and Derek Parfit, "Reasons and Motivation I," Proceedings of the Aristotelian Society, Supplementary Volume, 71 (1997): 99-130 at 114, fn. 28. The concern is that anti-Humeans confuse normative and motivating reasons where these two sorts of reasons are supposed to be distinct and only motivating reasons are supposed to be psychologically real. But this concern simply begs the question against anti-Humeans. Indeed, as I have arguably shown, normative reasons (i.e. reason facts) necessarily do motivate when grasped in the first person.

39 McDowell, "Are Moral Requirements Hypothetical Imperatives?," at 15.
towards φ-ing does not make her motivation to φ rationally intelligible because there is nothing more to a NCPA towards φ-ing than the logically independent (and merely causal) disposition to φ, and this disposition is the very thing that needs to be made rationally intelligible.

Taking our cue from Phillipa Foot, we might query whether its appearing to someone that she has reason to φ suffices to make her motivation to φ rationally intelligible. The concern is that, unless there are objective constraints on what an agent has reason to do, it can appear to someone that she has reason to do anything, thus robbing CPA’s of any rational intelligibility they might contribute. For instance, if what we have reason to do were not a matter of fact in some sense, then it could appear to someone that, say, she has reason to drink a can of paint -- without anything needed in the way of special circumstances to explain its appearing so. Thus, if we were trying to explain why someone did such a thing, referring to how obligatory drinking a can of paint appeared to her would not make her drinking it (rationally) intelligible. For, without reasons having some objective constraints, its appearing to our paint drinker that she had reason to drink the paint provides no more (rational) intelligibility than her logically independent (and merely causal) disposition to φ, which is to say: none at all.

---

41 The can of paint example is from Donald Davidson, "Actions, Reasons and Causes" in Essays on Actions and Events (1980), 3-19, though Davidson and I disagree about its significance. Davidson might argue that NCPA’s can play a role in rational explanations if only under a different description. Arguably, such a move fails for reasons that I discuss in Chapter Two.
The anti-Humean response is simple if not clear. CPA’s presuppose that there are facts about what each of us have reason to do. Perceiving is a factive mental state.42 Without reasons being in some sense a matter of fact, my talk here of an agent’s perceiving (as opposed to quasi-believing) what she has reason to do would not, strictly speaking, make sense. Furthermore, how things appear is arguably parasitic on how things are. So, when anti-Humeans explain someone’s action in terms of its merely appearing to her that she had reason to act, their explanations still provide rational intelligibility. Of course, issues remain regarding the objectivity presupposed by anti-Humeans that I have so far taken for granted. I return to these issues in Chapter Five.

42 On factive mental states including perceiving, see Williamson, "Is Knowing a State of Mind?," at 551-558.
CHAPTER TWO

Having Your (Layered) Cake and Eating It Too:
How to Save Nonreductivism from Trouble with Mental Causation

In a well-known paper, Donald Davidson holds that particular events such as my fleeing a bear may be explained in different sorts of ways according to the level of description employed.¹ Because it combines physicalism with a kind of antireductionism about the mental, Davidson’s approach to mental causation has proven widely attractive, especially in its realist guise. But Davidson and his realist interpreters face grave difficulties. Moreover, it is often thought that, if Davidson’s approach fails, nonreductive physicalism itself fails for the same reasons. I take issue with this. While I agree that a nonreductive physicalist has no room for mental (qua mental) causation if, following Davidson, she accepts the token-identity of mental and physical events, a nonreductive physicalist should not and need not accept token-identity.² For in the words of John Haugeland:

¹ Donald Davidson, "Mental Events" in Essays on Actions and Events (1980), 207-225.
² Others who reject token-identity include Jennifer Hornsby, "Physicalism, Events and Part-Whole Relations" in Simple Mindedness: In Defense of Naive Naturalism in the Philosophy of Mind (1997), 46-62 and John Haugeland, "Weak Supervenience" in Having Thought: Essays in the metaphysics of mind (1998), 89-107. In contrast to the view developed here, Hornsby does not seem terribly concerned about whether she remains a physicalist after rejecting token-identity, though see Jennifer Hornsby, Simple Mindedness: In Defense of Naive Naturalism in the Philosophy of Mind (1997). Haugeland is more concerned but, as we shall see, accepts a position that fails to respect (some of) the basic intuition(s) behind physicalism. Moreover, neither Hornsby nor Haugeland explicitly link their rejection of token-identity to giving the mental (qua mental) a causal role.
The world does not come metaphysically individuated, any more than it comes metaphysically categorized, prior to and independent of any specific descriptive resources – which is not to say that these individuals (and categories) aren’t perfectly genuine and objective … It is only to say that the individuals we can discuss in one way of talking need not be identified with the individuals we can discuss in another way of talking, even if the latter way of talking is somehow “basic and comprehensive”. That is, we need another way of accounting for the basicness and comprehensiveness of physics than the [token-]identity theory.3

With this point in mind, I develop an alternative version of nonreductive physicalism that does not involve token-identity. In short, my alternative envisions the world as layered, each layer (or level) populated by distinct substances -- e.g. persons -- and their characteristic properties -- having reason to do x, doing x, etc. -- with substances at one level composed of substances from lower levels as persons are composed of a body, cells, molecules and so on.4 After explaining why my alternative deserves the title of physicalism and defusing a fundamental objection based in mereology, I show how my alternative allows for mental (qua mental) causation.

3 Haugeland, "Weak Supervenience" at 103. This passage points to a tension in Davidson’s work. For his idea that, without earning the right to speak of things in different schemes as type-identical, we can nevertheless regard such things as token-identical reflects David Wiggins’ substratum myth. David Wiggins, Sameness and Substance (1980) at 139-140. But Wiggins’ substratum myth is essentially the scheme/content distinction that Davidson famously criticizes in Donald Davidson, "On the Very Idea of a Conceptual Scheme" in Inquiries into Truth and Interpretation (1984), 183-198.

In his own defense, Davidson might claim that various schemes can be about the same world (i.e. commensurable) only if token-identity holds. Adrian Cussins, in fact, suggests an argument for token-identity along these lines. Adrian Cussins, "The Limitations of Pluralism" in Reduction, Explanation, and Realism, ed. Charles and Lennon (1992), 179-223 at 185-192. However, this argument and the related defense of Davidson fails since, I maintain, the things we talk about in various schemes can be composition-related (and ipso facto part of the same world) without being token-identical.

4 I take a substance to be anything that exists through time and therefore, as the kind of thing it is, possesses unique persistence and identity conditions. This neo-Aristotelian conception of substance can be found in Wiggins, Sameness and Substance at 57-74.
1. Davidson’s approach to mental causation can be seen as an attempt to show how reasons for action can also be causes. Notice first that, according to Davidson, if my having reason to act causes my acting, then at least some mental events causally interact with physical events. For my having reason to act is typically thought of as mental while my acting is typically thought of as physical. However, the causal interaction of mental and physical is problematic given the strict nomological character of causality and the anomalism of the mental. The former holds that “where there is causality, there must be a [strict] law.” But the latter holds that the mental is governed by its own constitutive ideal of rationality and cannot, therefore, enter into strict laws. Consequently, we face a trilemma: either we give up on the characteristic normative aspect of reasons by reducing them to something that can figure in strict laws, or we give up on the idea that causality implies strict laws, or we give up on the idea that reasons can be causes.

Davidson does not intend to relinquish any of these claims. He purports to reconcile them. To do this, he sharply distinguishes events as bare particulars from the different ways we have of describing such particulars. Where there is a causal relation – as there seems to be between my having reason to act and my acting – it is supposed to obtain simply in virtue of the particular events involved and regardless of the descriptions we use to think and talk about them. So, the events picked out by ‘my having reason to act’ and ‘my acting’ may be causally related even if the descriptions we use to pick them

---

5 Consistent with Davidson’s approach, I take events to be particulars that cause and are caused by other such particulars.
6 Davidson, "Mental Events" at 208.
7 As noted earlier, Davidson’s sharp distinction of events from their descriptions is curious given his famous criticism of the scheme/content distinction. Davidson, "On the Very Idea of a Conceptual Scheme".
out cannot be brought under a strict law. Put another way, allowing that my having reason to act and my acting are causally related in no way undermines the anomalism of the mental because the anomalism of the mental concerns not the events themselves but only one way of describing them. Furthermore, though one scheme for explaining and describing causally related events does not involve strict laws, there must – according to the nomological character of causality – be a scheme that does. Physical theory provides it. Accordingly, Davidson reconciles these three claims – mental-to-physical causation, the nomological character of causality, and the anomalism of the mental – as follows. He grants that my having reason to act and my acting are anomalous according to our commonsense psychological ways of describing them but holds that, since they are causally related, they must fall under strict laws when described in terms of physical theory.

On one hand, Davidson’s position is antireductionist. Given the anomalism of the mental, there can be no strict psychophysical laws. On the other hand, his position is arguably physicalist. It reflects the primacy of the physical in the sense that every event with a mental description also has a physical description but not vice versa. This claim – that every event with a mental description also has a physical description -- is known as token-identity physicalism and, along with his antireductionism, forms the core of his position. Despite this attractive combination, Davidson’s approach to mental causation is

---

8 Jaegwon Kim maintains that Davidson’s claim of mental-physical supervenience conflicts with his antireductionism. E.g. section IV of Jaegwon Kim, "The myth of nonreductive materialism" in Supervenience and Mind: Selected Philosophical Essays (1993), 265-284. I have not said anything about supervenience in my account of Davidson primarily because supervenience, by itself, does not add much. Simon Blackburn, Spreading the Word (1984) at 186. Moreover, if Kim is right, it only buys Davidson trouble. Though on the latter point, I agree with [Seager, 1991 #357] that
unnecessarily revisionist. For him, “[T]he mental is not an ontological but conceptual
category... To say of an event ... that it is mental is simply to say that we can describe it in
a certain vocabulary...”9 In contrast to this nominalism, I believe we should try to respect
“common sense natural realism” according to which “objects and properties ... talked
about in whatever discourse are real, unless it has been specifically argued otherwise.”10
Of course, perhaps the main argument for mental irrealism concerns the troubles that
nonreductive realists have in finding a causal role for the mental.11 No doubt, for this
reason, irrealism can appear attractive. Still, we need not abandon nonreductive realism
about the mental due to these troubles, or so I shall argue.

2. One can accept the core of Davidson’s view -- antireductionism about the
mental, the token-identity of the mental with the physical – and be a realist about the
mental. Nevertheless, nonreductive realists who accept token-identity have trouble with
mental causation. Given what these realists accept and two independently plausible
premises – the causal closure of the physical and the stricture against systematic
overdetermination – there is no room for the mental (qua mental) to make a causal
difference. Thus, epiphenomenalism is the price that nonreductive realists about the
mental pay for accepting token-identity.

supervenience need not mean trouble for Davidson or nonreductive physicalism
generally. I return to this issue in section 6.
9 [Davidson, 1987 #262] at 46. Davidson emphasizes that laws are also linguistic.
Davidson, "Mental Events".
10 J Van Brakel, "Interdiscourse or Supervenience Relations: The Primacy of the
Manifest Image," Synthese, 106 (1996): 253-297 at 261. Also see Huw Price,
11 Some argue for moral irrealism on similar grounds. Gilbert Harman, for example,
writes: “[T]here does not seem to be any way in which the actual rightness or wrongness
of a given situation can have any effect ...” Gilbert Harman, The Nature of Morality
(1977) at 8.
Let us take a closer look at why epiphenomenalism results from combining nonreductive realism with token-identity.\(^\text{12}\) Recall our putative example of mental causation: my having reason to act causes my acting. On a realist construal of Davidson’s position, an event with the mental property of being my reason to act causes an event with the physical property of being my action. But given token-identity, it follows that my having reason to act also has some physical property. And now the question arises: In virtue of what property -- mental or physical -- does my having reason to act cause my acting? To constitute mental (\textit{qua} mental) causation, my having reason to act must cause my acting in virtue of the former event’s having the mental property that it does, namely, the property of being my reason to act. However, the causal closure of the physical and the stricture against systematic overdetermination leave mental properties of events with no causal role to play. The causal closure of the physical holds that every physical event has a (complete) physical cause. In other words, for every event \(e\) with a physical property \(P_1\), there is another event \(c\) with a physical property \(P_2\) that causes \(e\) in virtue of \(c\)’s being \(P_2\). This means that, if nothing else, my acting is caused by my having reason to act in virtue of the physical property that this latter event possesses. By itself, the causal closure of the physical does not exclude the possibility that my acting is also caused by my having reason to act in virtue of the mental property that it possesses. Overdetermination can happen. But to support a causal role for the mental, the envisioned kind of overdetermination would have to be systematic -- not just a fluke -- and our other,\(^\text{12}\) 

independently plausible premise rules this out. The stricture against systematic
overdetermination holds that, typically, events have only one (complete) cause. That is,
typically, events stand in causal relations in virtue of just one sort of property that they
possess. As we have seen, my acting is caused by my having reason to act in virtue of the
latter event’s possessing some physical property. Thus, in the typical case, there is no
room for my acting to be caused by my having reason to act in virtue of its possessing the
mental property that it does. There is no room, in other words, for a mental (qua mental)
cause of my acting.

One might object to the causal exclusion argument just mooted by questioning
either the causal closure of the physical or the stricture against systematic
overdetermination. However, neither line of objection is promising. Take systematic
overdetermination. Any model of causal relations that makes it a feature of some generic
class of events such as actions can hardly be plausible. To paraphrase Stephen Schiffer,
such a model of causal relations would make God out to be an incredibly bad engineer.\textsuperscript{13}
Systematic overdetermination is metaphysically extravagant – hence, the stricture against
it.

Now consider the causal closure of the physical. Understood as an empirical
claim, the weirdness of quantum physics might be thought to refute it.\textsuperscript{14} But the causal
closure of the physical is best understood as a conceptual claim whose truth is a necessary
condition of any scientific study of the physical. There are limits, in other words, on how
weird quantum physics can get and still count as science. In this chapter, I cannot
adequately defend this view of causal closure, but a simple analogy brings out the thought
\begin{footnotes}
\textsuperscript{14} Michael Silberstein suggested this to me in conversation.
\end{footnotes}
behind it. Like a physicist who wants to understand what makes electrons spin, my mechanic wants to figure out why my car won’t start. But it makes sense for my mechanic to try to fix my car only if my car is a closed system (at least to the extent of the immediate cause of its failure to start). For if the immediate cause of my car’s failure to start has nothing to do with my car, then I have no more reason to take my car to a mechanic than a witch doctor. For similar reasons, a physicist’s investigation (qua physicist) of electron spin makes sense only if such fundamental goings-on are part of a closed system.

For a more promising strategy against the causal exclusion argument – and the one that I pursue here – nonreductive realists about the mental should reject token-identity. Indeed, to avoid Davidson’s irrealism and the epiphenomenalism of his realist interpreters, a commitment to antireductionism about the mental requires nothing less. I now turn to the case against the token-identity of the mental with the physical.

3. First of all, rejecting token-identity means taking issue with one of Davidson’s three claims – mental-to-physical causation, the nomological character of causality, and the anomalous of the mental -- that token-identity serves to reconcile. For if we grant these three claims as premises and the further premise that every mental event stands in some causal relation, it follows that every mental event is token-identical with some physical event. So, conversely, if we reject token-identity, we have to reject one of these premises. But here’s the rub: Davidsonians will maintain that each of these premises is eminently plausible and, therefore, so is token identity.

While it may appear that the premises yielding token-identity are all eminently plausible, this is not true. In particular, the nomological character of causality – the claim
that where there is causality, there is strict law – is dubious. Consider the sorts of causal
relations that we are familiar with as part of our everyday knowledge of the world – e.g.
throwing a baseball at a glass window tends to break the window, scratching the paint on a
new car tends to upset the car’s owner, etc. Given the relative imprecision that the relata
of everyday causal relations involve – e.g. how hard was the ball thrown? how thick is the
glass in the window? -- everyday causal relations cannot be captured by strict laws.
Rather, any attempt to formulate them (as I have done in my examples) must be hedged –
either implicitly or explicitly – by a ceteris paribus clause or its equivalent. Granted, there
may be some causal relations expressible in terms of strict laws. And if any causal
relations are so expressible, it is the causal relations that hold between physical events.
But if such causal relations exist, they appear to comprise an entirely different subject
matter from the causal relations with which we are familiar.\textsuperscript{15} Certainly, without more,
we have no reason to believe that everyday causal relations are really causal relations
expressible in terms of strict laws. Thus, the Davidsonian argument does not compel
acceptance of token-identity.

Just as we lack reason to believe that everyday causal relations are really causal
relations expressible in terms of strict laws, so we lack reason to believe that
commonsense psychology (with its talk of my having reason to act) concerns the same
particular events as fundamental physics. Put rhetorically, “why assume,” with advocates
of token-identity, “that we can superimpose these two accounts and find that there are

\textsuperscript{15} Others who stress the distinction between our everyday notion of cause and the more
technical notion of a strict cause include John Haugeland, "Phenomenal Causes," \textit{The
Explanation" in \textit{Essays on Davidson: Actions and Events}, ed. Vermazen and Hintikka
discrete events that each [account] describes in its own way?" If we assume anything, it should be the opposite, namely: that commonsense psychology and fundamental physics cannot be superimposed to reveal the same particular events. In other words, we should assume that mental and physical events are not token-identical. To see why this is the correct assumption, we need to consider what establishing token-identity would require.

Imagine that some ambitious scientists set out to demonstrate the token-identity of mental and physical events. With this goal in mind, they invent an amazing device: a helmet with a brain-imaging system that allows the scientists, back in their lab, to monitor (and record) the physical events going on inside the brain of the person wearing the helmet as this person goes about her daily business. The helmet also has a miniature video camera and microphone that lets the scientists see and hear the world from the helmet wearer’s perspective. In addition, the microphone allows the scientist to hear the helmet wearer’s sincere, contemporaneous reports of her occurrent thoughts and feelings. Two features of the information that the helmet yields deserve mention. First, the brain-imaging system gives the scientists a view of the causally related chain of physical events going on inside the helmet wearer’s brain while the audio and video transmissions give the scientists a view of the helmet wearer’s mental life. Second, the two chains of events revealed by the helmet happen at the same time. Since the recording of these events preserves their synchronization, the scientists can review the chain of mental events experienced by the helmet wearer, choose one to focus on, and then examine the corresponding chain of physical events to see which of these events was happening at the same time as the relevant mental event. With all this information, it might seem that our

scientists can now establish the token-identity of mental and physical events. But can they? I think not.

For one reason, consider the problem raised by the following scenario.17 Suppose the scientists’ helmet-wearing subject is walking down the street when she happens upon a local café. At the same time as she smells the aroma of freshly ground coffee, she also sees patrons of the café leaving with cups of coffee. The former perception excites a desire in her for coffee while the latter gives rise to a belief about how she can satisfy her desire. Thus, in terms of the subject’s mental life, two distinct mental events – a desire for coffee and a means-end belief – occur at the same time. Focusing on these two simultaneous mental events, the question is: can the scientists say which bunch of physical events is the desire and which bunch is the belief? If token-identity holds, they should be able to. But they cannot because the physical events that unfold over the same expanse of space and time in which these two mental events occur comprise a formless jumble, a blooming, buzzing confusion (to appropriate James’ famous phrase). Therefore, no basis is available for distinguishing one part of this jumble as the desire and another part as the belief. Moreover, such a basis seems unavailable in principle. There do not appear to be any further facts that, if known, would allow the scientists to make the required identifications. In light of the absence, in principle, of the decisive information, any claim that the envisioned identifications nevertheless hold -- that mental events are token-identical with physical events -- deserves to be treated with skepticism.

17 For similar arguments, see Jennifer Hornsby, "Which physical events are mental events?" in Simple Mindedness: In Defense of Naive Naturalism in the Philosophy of Mind (1997), 63-77 and especially Haugeland, "Weak Supervenience".
One way of short-circuiting this skepticism should be mentioned. I have in mind the view, held by Quine and now Davidson, that events are nothing but regions of space-time. On this view, the scientists in my example establish the token-identity of some (so-called) mental event with some (so-called) physical event just by specifying (via their brain-imaging helmet) the expanse of space-time associated with the (so-called) mental event. There is no need of any further finding. For, given what events are on this view, there is really only one such thing here – that is, one region of space-time – variously described in mental and physical terms. This view of events, in other words, builds token-identity right in. And this is the problem. Since we have reason to be skeptical about token-identity, any view of events that simply assumes it is unsatisfactory. If token-identity is to have our assent, it must be earned.

4. While the preceding line of argument against token-identity seems right as far as it goes, some may be reluctant to give up token-identity due to an apparent problem with the alternative. The alternative to token-identity involves accepting the distinctive ontology of commonsense (including commonsense psychology) at face value rather than reinterpreting it as being about the same particular events as fundamental physics. Of course, suitably idealized, we should also accept the ontology of fundamental physics. But this equanimity appears to pose a problem. For if mental and physical events are not even token-identical, can we still be physicalists? (At this point, it’s reasonable to ask why we should want to be physicalists – more on this shortly.) So consider physicalism. At a minimum, it requires that there be some asymmetry between the mental and the physical, that the former depend on the latter in some sense. In turn, some have thought
that this dependence requires token-identity. Therefore, giving up token-identity appears to mean giving up physicalism as well.

We should notice one thing right away about this supposed problem. Without at least some intuitive grasp of the relevant sense of dependence, we cannot begin to say whether one must accept token-identity to be a physicalist. The crucial question concerns this sense of dependence. We can imagine a whole range of different degrees (or strengths) of dependence and corresponding versions of physicalism. But we should focus on the sense of dependence that captures the legitimate intuitive appeal of physicalism. On one hand, this means that the sense of dependence involved cannot be so strong as to exclude the distinct reality of the mental. Any sense of dependence that strong would clash with our commonsense realist view of the mental. Respect for commonsense thus explains my failure to discuss reductive, type-identity physicalism. In light of commonsense, type-identity physicalism involves a sense of dependence that is simply too strong to be plausible. On the other hand – at the end of the spectrum where physicalism distinguishes itself from dualism – the sense of dependence involved cannot be so weak as to make relations between the mental and the physical inexplicable. We want to be able to explain, for example, why I feel pain when part of my body suffers tissue damage.

Honing in on the legitimate intuition behind physicalism, these two competing considerations suggest why token-identity physicalism has proven attractive. In its realist version, token-identity physicalism respects commonsense enough to recognize the

---

18 As John Foster observes, “[W]e have ... a natural inclination ... that would lead us to reject any conceptual analysis of mental statements in non-mentalist terms or any metaphysical reduction of mental facts to non-mentalist facts ...” John Foster, The Immaterial Self: A defence of the Cartesian dualist conception of the mind (1991) at 206.
distinct reality of the mental while it also provides, via token-identity, an explanation of mental-physical relations. Still, I have argued that token-identity physicalism is not sufficiently respectful of commonsense. In particular, token-identity physicalism cannot allow for the causal autonomy of the mental. The problem facing us if we reject token-identity can now be stated more pointedly. Is there a sense in which the mental depends on the physical that is weaker than (i.e. does not involve) token-identity but nevertheless strong enough to make sense of mental-physical relations? The answer, I maintain, is ‘Yes’. We can be physicalists in the legitimate, intuitively appealing sense even if we reject token-identity. However, making good on this claim is not as easy as some philosophers have thought.

Consider John Haugeland who shares my belief that the legitimate, intuitively appealing sense of physicalism ought to be available sans token-identity. For Haugeland, the legitimate intuition behind physicalism is that “[F]ixing’ the physical fixes everything, or that nothing could have been otherwise without something physical

---

19 At this point, someone might be tempted to say: ‘So much the worse for commonsense. Token-identity physicalism reconciles our commonsense intuitions about mind and matter as much as we have a right to expect. Put bluntly, commonsense is incoherent so we cannot expect to do any better.’ While this sort of skepticism may appear justified to some, I am not ready to concede the incoherence of our commonsense scheme. Indeed, you might say that the burden of this chapter is to show how the different demands that commonsense places on our conceptions of mind and matter can be reconciled.

Token-identity physicalism also comes under attack from the opposite direction, that is, from type-identity physicalism. For if token-identity physicalism cannot allow a causal role for the mental, can it allow for the distinct reality of the mental? This is another form of the pressure that I mentioned earlier. While before it was an inference from the causal impotence of the mental to its irreality, here it is an inference from the causal impotence of the mental to its type-identity with something not causally impotent. Again, there is a parallel move in the moral realism debate. Nicholas L. Sturgeon, "Moral Explanations" in Essays on Moral Realism, ed. Sayre-McCord (1988), 229-255. 20 Haugeland, "Weak Supervenience". Page numbers in the main text refer to this essay.
having been otherwise.” (95) He stresses that this intuition – or, the sense of dependence it reflects – can be captured by a supervenience relation that does not presuppose token-identity. Thus, for Haugeland, the legitimate appeal of physicalism can be expressed as:

Set A of properties *globally supervenes* on set B of properties just in case there are no two [possible] worlds indiscernible in B and yet discernible in A (where A are mental properties and B are physical properties).21

Conspicuously absent from global supervenience is any reference to the individuals bearing mental and physical properties. This is no coincidence. About the intuition that Haugeland regards as the legitimate appeal of physicalism, he writes:

There is no reason whatsoever to tie this intuition down to any particular domain of individuals; that is, expressing the intuition need not at all be confined to the form: ‘For each individual, it could not . . . ’ The idea that fixing the physical fixes everything need not be broken down on an individual-by-individual basis – it can be, so to speak, ‘world-wide’. (95-6)

This passage, and Haugeland’s position generally, suggests two points. First, in Haugeland’s view, mental-physical supervenience cannot hold on an individual-by-individual basis if it is to avoid token-identity. Second, and more importantly, Haugeland believes that mental-physical supervenience need not hold on an individual-by-individual basis if it is to capture the legitimate, intuitive appeal of physicalism.

21 Haugeland calls the supervenience relation that he proposes ‘weak supervenience’. Following more common usage, I call it global supervenience. This formulation of global supervenience is (GS1) from Jaegwon Kim, "Supervenience for Multiple Domains" in *Supervenience and Mind* (1993), 109-130 at 117. (GS1) is equivalent to Haugeland’s formulation in terms of the truths of two languages since, in Jaegwon Kim’s words, “we may assume that all the truths holding in a world are fixed, in the standard ways, by the ‘singular truths’ that hold in it, that is, the distribution of properties over the individuals.” Kim, "Supervenience for Multiple Domains" at 118.

Common philosophical usage also takes ‘strong supervenience’ and ‘weak supervenience’ to mean supervenience relations that hold across possible worlds and within such a world, respectively.
Haugeland is wrong on both counts. To see why he is wrong on the first count, we need only consult Jaegwon Kim’s account of supervenience across possible worlds (i.e. strong supervenience) for multiple, coordinated domains. Each domain $D_1$ and $D_2$ of individuals has its own distinctive set of properties $A$ and $B$, respectively, and each individual $x$ in $D_1$ is related by $R$ to some individual(s) in $D_2$. This corresponding set of individuals in $D_2$ is $R|x$, or what Kim calls the image of $x$ under $R$. With this as background, we can state Kim’s schema (MSS):

\[
\{A, D_1\} \text{ strongly supervenes on } \{B, D_2\} \text{ relative to relation } R \text{ just in case for any } x \text{ and } y \text{ in } D_1 \text{ and any worlds } w_1 \text{ and } w_2, \text{ if } R|x \text{ in } w_1 \text{ is } B\text{-indiscernible from } R|y \text{ in } w_2, x \text{ in } w_1 \text{ is } A\text{-indiscernible from } y \text{ in } w_2. \]

Let the members of $D_1$ and $D_2$ be persons and bodies, respectively, with certain mental properties comprising set $A$, the properties exemplified by persons, and certain physical properties – e.g. being damaged in such and such a part – comprising set $B$, the properties exemplified by bodies. Furthermore, let $R$ be the one-to-one relation between a particular person and her body, that is, the composition relation. Interpreted in this

---

22 Kim, "Supervenience for Multiple Domains" at 124.

23 I would limit the mental properties in $A$ to those such as feeling pain because, on my view, feeling pain is a special (i.e. first-person) way in which a person is aware of her own body and what’s happening in various parts of it. For a recent discussion of pain as bodily awareness, see M.G.F. Martin, "Bodily Awareness: A Sense of Ownership" in The Body and The Self, ed. Bermudez, Marcel and Eilan (1995), 267-289.

24 This one-to-one relation between supposedly distinct individuals might make some wonder how we explain this correspondence. Without token-identity, MSS can seem to reflect a miraculous coincidence. Kim, "Supervenience for Multiple Domains" at 126-27. But the one-to-one relation between persons and bodies is explained by the sorts of things
way, MSS delivers mental-physical supervenience that holds on an individual-by-individual basis but does not presuppose the identity of the individuals involved.

MSS is also what we need to express the legitimate, intuitive appeal of physicalism. Certainly, contra Haugeland, the legitimate intuition behind physicalism, cannot be captured by mental-physical supervenience that is merely world-wide. Rather, mental-physical supervenience must hold on an individual-by-individual basis, at least when we are talking about mental properties like feeling pain. Otherwise, two possible worlds that differed in some trivial, far removed physical characteristic could differ in any or all of their mental characteristics, that is, so far as the mental is concerned vis-a-vis the physical, it would be anything goes. For we do not really tie down the mental in any significant sense if we only tie it down on a global basis.25 As an example, consider two worlds in which my hand suffers tissue damage, say, it’s badly burnt. Moreover, virtually all other physical characteristics of these two worlds are the same. Nevertheless, in one world, the Hope diamond is composed of $x$ carbon atoms while, in the other, it is composed of $x + 1$ carbon atoms. Consistent with Haugeland’s world-wide sense of dependence, I may feel pain when my hand is burnt in one world but not the other. Indeed, in the other virtually identical world, no one may feel (my) pain -- not even Bill Clinton! But whether I feel pain should depend upon whether my hand is burnt and what’s happening to the rest of my body -- not upon trivial, far removed differences in the rest of the world. This is a matter of commonsense -- it’s the legitimate intuitive appeal of that persons and bodies are, namely, things that stand in the composition relation to each other (and composition is not identity). For this view of persons and their bodies, see Wiggins, *Sameness and Substance* at 163-164.

25 The same general point is made by Jaegwon Kim, ""Strong" and "Global" Supervenience Revisited" in *Supervenience and Mind* (1993), 79-91 at 85-86.
physicalism -- and Haugeland’s world-wide supervenience is too weak to capture it. MSS, by contrast, does capture this intuition. For if my hand is burnt in two worlds and my body in these worlds is indiscernible in all other (physical) respects then if I feel pain in one world, I feel pain in the other world – regardless of what’s going on in these worlds apart from my body.

Haugeland’s willingness to accept global supervenience and his consequent failure to capture the legitimate intuitive appeal of physicalism raise an interesting question. Why does he have a blindspot for MSS generally and, in particular, for the interpretation of MSS that I am suggesting in terms of persons, bodies, and the composition relation. The answer, I think, must be that another argument for token-identity lurks in the background and appears to foreclose the more reasonable kind of physicalism (sans token-identity) that I am recommending.

5. Interestingly, Haugeland accepts most of my layered ontology. He agrees that ontological pluralism goes hand-in-hand with explanatory pluralism. Recall that he holds:

The world does not come metaphysically individuated, any more than it comes metaphysically categorized, prior to and independent of any specific descriptive resources . . . [T]he individuals we can discuss in one way of talking need not be identified with the individuals we can discuss in another way of talking . . . (103)

In other words, given the different ways we have of singling out substances and making sense of the world, the substances being talked about in our various schemes should be taken at face value – that is, as different things – rather than reinterpreted as the same things differently described. So Haugeland apparently accepts my picture of the world as populated by distinct kinds of substances with their own characteristic properties. Presumably, he also accepts that such substances and properties are interrelated since he
believes that, at least globally, one such level (the mental) supervenes on another (the fundamental physical). Nevertheless, Haugeland fails to notice that even if composition relates these various things – as seems plausible – supervenience can still be secured on an individual-by-individual basis without committing ourselves to token-identity.

To see why Haugeland might miss this possibility, we need to be clear about composition. Substances picked out by one scheme at one level – e.g. persons – are composed of different sets of lower-level substances – bodies, cells, molecules, and so on – according to the lower-level scheme involved. Moreover, every higher-level substance is ultimately composed of some set of fundamental physical substances. Besides underwriting a stronger, more intuitively appealing supervenience relation, composition also explains the mental’s dependence on the physical. For if everything is composition-related and all the physical stuff in the world ceased to exist, then so would everything else including persons and their mental properties. Similarly on an individual-by-individual basis, if someone’s body ceased to exist, then so would that person. The problem with this? For someone like Haugeland who wants to avoid token-identity, mereology can appear to block characterizing interlevel relations in terms of composition.

The two principles of mereology that, taken together, appear to block my alternative version of nonreductive physicalism are what we might call (E) for the existence of compositions and (U) for the uniqueness of compositions. Consider (E) first. (E) holds:

\[(x_1)\ldots(x_N)(zy)(y \text{ is composed of } x_1\ldots x_N)\]

where ‘\(x_1\)’…‘\(x_N\)’ and ‘\(y\)’ range over things of one kind. Applied to, say, the particles at the fundamental physical level that compose a person at a time, (E) implies that this set of
fundamental physical particles is itself a fundamental physical thing. Thus, at any time during a person’s existence, there is a fundamental physical thing arrived at via (E) that is composed of the very same particles as compose the person at that time. But now consider (U) which holds:

$$(x_1)...(x_N)(zy)\{(y \text{ is composed of } x_1...x_N) \& (z)[(z \text{ is composed of } x_1...x_N) \rightarrow (y = z)]\}$$.  

Of the two supposedly distinct things that we have been discussing – a person and a fundamental physical thing arrived at via composition – (U) implies that there is really only one thing here, call it what you will. Persons are really just fundamental physical things arrived at via composition. Thus, mental events such as my having reason to do $x$ must be token-identical with some fundamental physical event because ‘my’ refers not to a distinct subject for mental properties – i.e., a person as commonly understood – but to some composition of basic physical stuff. Consider, however, that when a person dies, she ceases to exist while her body does not. More generally, whatever example we use, a person and the body which composes that person do not share all the same properties. Therefore, given Leibniz’ Law, (U) cannot be true and the path lies open for my alternative (layered) version of nonreductive physicalism.

6. Now let us see how a layered ontology, each layer (or level) with its own distinct substances and its own characteristic properties, makes room for mental \textit{qua} mental causation in the physical world.\textsuperscript{26} Mental causation occurs when some thought

\textsuperscript{26} Hereafter, I shall simply speak of mental causation. However, when I so speak, I should be understood as referring to mental \textit{qua} mental causation unless otherwise noted.
(or combination of thoughts) that a person is thinking causes him to do something.

Take my fleeing a bear. I do this because I want very much that I am not mauled by a bear and I believe that my fleeing will help bring this about. Emphasizing ‘because’ here signals something distinctive about mental causation, namely, that it is an instance of not merely causality but rationality as well. Put another way, the connections between a person’s thoughts and her actions are not blind causal connections but causal connections mediated by thoughts. This means, for one, that when we say what made a person do what she did – that is, when we causally explain her actions – we appeal to some thought she had at the time that also makes what she did rationally intelligible. More broadly, the fact that mental causation is rational causation points to the causal relations in which persons and their characteristic properties figure, namely, the causal relations picked out by commonsense psychology. For commonsense psychology is in the business of rationalizing our actions.

One crucial element, then, in the layered ontology being defended here: not only do the counterfactual-sustaining generalizations of commonsense psychology reflect the causal relations we want to accommodate – the causal connections constitutive of mental causation – but the items involved in commonsense psychology – persons with their

27 The psychological state of a person that causes her to act is usually taken to be an amalgam, namely, a desire that ψ and a belief that by φ-ing, ψ will ensue. The view that I defend in Chapter One, by contrast, claims that cognitive pro attitudes are the psychological states that cause us to act. To avoid distraction and complication, I shall speak only in terms of the more familiar view of the relevant psychological state as a belief-desire pair.

28 Thus, the awareness of a moral property that can cause us to act is not an unmediated acquaintance with the fact that some action possesses this property. The fact itself does not move or cause us to act. Rather, it only does so when grasped under a mode of presentation, namely, the noninferential first-person mode of presentation (as I argued in Chapter One).
thoughts and actions – occupy their own level in such an ontology. Commonsense psychology is neither cognitive science, neurophysiology, neurobiology nor neurochemistry. Nor – least of all – is it physics or any other other causal-explanatory scheme. Rather, the counterfactual-sustaining generalizations definitive of commonsense psychology and other causal-explanatory schemes each relate distinctive sorts of individuals having their own characteristic properties. The significance of such a layered ontology with mental causation occurring at one of its levels has been noticed by Jaegwon Kim. He writes:

[T]he layered model may appear to hold promise as an elegant way of averting violation of the causal closure of the physical: causal interactions could perhaps be confined within each level, in a way that respected the autonomy and closedness of the causal processes at the fundamental physical level. In particular, on the nonreductivist version of the layered model, it may be possible to view causal chains at a given level, like the [individuals and] properties distinctive of that level, as forming an autonomous and self-contained realm immune to causal intrusions from neighboring levels.

Indeed, this is precisely the idea behind the view being proposed here to make room for mental causation in the physical world. Call this idea the causal autonomy and containment thesis (or “CAC”). Paraphrasing Kim, CAC holds that: The causal chains of events at a given level – like the individuals and properties distinctive of that level from which these events are composed – form an autonomous and self-contained realm in that they are not causally related to events at other levels. Consequently, nonreductive physicalists maintain that if CAC holds – as it should when the layered model is properly

---

29 This brings to mind Bishop Butler’s famous dictum “Everything is what it is, and not another thing,” used by G.E. Moore for his inscription to Principia Ethica. Of course, I am talking about not just individual things but whole causal networks of things.

30 Kim, "The nonreductivist's troubles with mental causation" at 338.
understood – then they can accommodate mental causation as a kind of causal chain unique to one level, the level associated with commonsense psychology.

Despite the attractions of the layered model for nonreductive physicalists, Kim rejects attempts to resolve their troubles with mental causation by appeal to such a model. I propose, then, to examine Kim’s objections to the layered model and offer a response. My response should better elucidate how I understand the layered model and how, pace Kim, it allows nonreductive physicalists to happily make room for mental causation.

So why is Kim bearish about the layered model? In brief, Kim argues that nonreductive physicalists who try to avail themselves of the layered model find themselves committed to downward causation and downward causation leads back to the same unattractive set of options that constituted their trouble with mental causation in the beginning. In response, I claim that nonreductive physicalists who avail themselves of the layered model need not be committed to downward causation and, therefore, can happily accommodate mental causation.

Before considering Kim’s fundamental objection, two other ways that the layered model might commit nonreductive physicalists to downward causation should be addressed. One such way would be if the layered model were conceived of as layers of distinct sorts of properties instantiated by just one sort of basic individuals. On this less populated version of the model, any mental event (i.e. the instantiation of a mental property by an individual) would be token-identical to some physical event (i.e. the instantiation of some physical property by the same individual). Thus, conceiving of the layered model in this sparser way – with distinct sorts of properties but not distinct sorts of individuals at each level – builds token-identity right into the model. And token-
identity – as we have seen – leads to downward causation which, in turn, leads to all the unhappy alternatives that nonreductive physicalists need to avoid. Fortunately, this is not the only way to conceive of the layered model. The layered model can also be thought of as more densely populated, that is, as involving both distinct sorts of properties and distinct sorts of individuals at each level. Conceived in this fuller way, the layered model does not guarantee token-identity, thereby protecting nonreductive physicalists from a too-quick commitment to downward causation and its concomitant problems.

Interestingly, while his fundamental objection does not depend on construing the layered model in the sparse way unfavorable to nonreductive physicalists, Kim nevertheless presents the model in these terms. That is, he is only willing to grant distinct status to the properties at each level, not to the individuals at each level. One wonders about Kim’s ontological parsimony here. As I have just argued in the previous paragraph and as I have suggested throughout this chapter, the failure to think of the layered model with distinct individuals at each (or at least some) of its levels hamstrings nonreductive physicalists from the outset in their efforts to accommodate mental causation. The motivation appears to be a concern that nonreductive physicalists not transmogrify into some form of dualism. Kim tells us that:

The basic ontological thesis of nonreductive physicalism confers on the physical a certain kind of primacy: all concrete entities are physical – there are no nonphysical particulars, no souls, no Cartesian mental substances, and no ‘vital principles’ or ‘entelechies.’”

Fine and good. But nonreductive physicalists can admit distinct sorts of individuals as well as distinct sorts of properties at each level of their layered ontology without quarreling with Kim’s physicalist scruples. On Kim’s view, in other words, it appears

31 Ibid at 339-340.
that admitting distinct sorts of individuals at each (or some) of the levels in the layered model rates as admitting “nonphysical particulars” into one’s ontology, tantamount to dualism. But concerns about dualism at this point are misplaced. For admitting distinct sorts of individuals at different levels of one’s ontology need not commit one to “nonphysical particulars.”

Part of the problem, I suspect, relates to ambiguities in the meaning of the term ‘physical’. Indeed, Kim acknowledges that, “There appears to be no generally accepted account of exactly what it means to say that something is ‘physical.’”32 The crucial point is that, on the richer version of the layered model that I propose for nonreductive physicalism, the distinct sorts of individuals introduced at different levels of the model are ‘physical’ in the relevant senses of that term. An individual can be said to be ‘physical’ if:

(i) it occupies a determinate space and time;

(ii) it figures in the causal relations picked out by a truly fundamental physics;

(iii) it figures in the causal-explanatory scheme of a truly fundamental physics;

(iv) it is composed out of individuals that figure in the causal relations picked out by a truly fundamental physics; or,

(v) it can truly be described in terms drawn from the causal-explanatory scheme of a truly fundamental physics.

These are five legitimate and important senses of the term ‘physical’. Now take the distinct sorts of individuals – persons – that would instantiate behavior-producing thoughts on a richer version of the layered model. Persons are ‘physical’ in all except the

32 Ibid at 340.
second and third senses. But we can expect no more from any other higher-order entity.

For instance, the higher-order entities that would instantiate mental properties on the thinned-down version of the layered model – mereological aggregates on Kim’s view – are not ‘physical’ in the second and third senses either. Therefore, this particular source of concern about dualism need not keep nonreductive physicalists from availing themselves of a richer version of the layered model and thereby avoiding token-identity and a too-quick commitment to downward causation.

Nonreductive physicalists, however, may find themselves committed to downward causation in another way that relates to how the layered model is understood. Recall my earlier example of mental causation: a desire not to be mauled by a bear and a belief that a bear is in the vicinity produces flight behavior. Thus far, we have discussed the left-hand side of the mental-causal relation but said little about the right-hand side. How are we to understand this side? That is, how are we to think of the behavior that results from our thinking certain thoughts? Two ways of understanding behavior can be gleaned from the literature.  

The first thinks of behavior as everyday actions, things that people do. Taking out the trash, shopping for groceries, checking e-mail, and so on: this is behavior in the first sense. Distinctively, behavior in this first sense possesses intentionality. It is consciously directed towards some goal; it has an object. By contrast, on the second understanding, behavior is mere bodily motion, nothing but muscle contractions and limb movements. Narrowly conceived, behavior in this second sense

33 See, example, the Editors’ Introduction to Norman Care and Charles Landesman, eds., *Readings in the Theory of Action* (1968) at xiv. See also Kim, "Dretske on how reasons explain behavior" at 292-297 for a discussion of Dretske’s “dual-explanandum strategy,” an approach to mental causation that, like mine, also depends on distinguishing these two senses of behavior.
includes immediate bodily responses to stimuli such as wincing from pain and pulling one’s hand back to avoid further pain. More broadly conceived, it includes all the complex bodily motion involved in our everyday actions understood as nothing but bodily motion (i.e. stripped of intentionality). 34

Problems arise for nonreductive physicalists when discussions of mental causation assume that the right-hand side of the mental-causal relation should be understood as behavior *qua* bodily motion. For behavior *qua* bodily motion is correctly thought to be located at a lower level in the layered model than the intentional, mental happenings that figure in the left-hand side of the mental-causal relation. Consequently, if the right-hand side is conceived as behavior *qua* bodily motion, then the mental-causal relation exemplifies downward causation on its face and efforts to accommodate mental causation by appeal to the layered model do not even get off the ground. To avoid this false start, nonreductive physicalists need to insist that we do not forget their layered ontology when it comes to the right-hand side of the mental-causal relation. Accordingly, the intentional, mental happenings on the left-hand side of the relation are to be thought of as producing behavior in the intentional sense of everyday actions, not behavior in the sense of mere bodily motion. We have, then, a relation where a person’s having certain propositionally contentful thoughts results in this person’s doing some intentional action. Both relata involve the same sort of thing – a person – instantiating properties drawn from the same causal-explanatory scheme – commonsense psychology.

34 Given these two very distinct senses of ‘behavior’, the nonreductive physicalism that I endorse with its emphasis on behavior in the first, everyday sense need have no complaint about scientific studies of behavior in second sense (except to the extent that scientists involved in such studies or their promoters in the media confuse the two senses of behavior).
In other words, the events related by mental causation are supposed to occur at the same level in a layered ontology, thereby avoiding any (obvious) commitment to downward causation and the troubles it involves.

We can now consider Kim’s main objection to the layered model as a way for nonreductive physicalists to accommodate mental causation. He correctly observes that:

Most nonreductive physicalists want to go beyond the claim that mental properties are had by physical systems [such as persons]. . . . The main idea here is that, in spite of their irreducibility, mental properties are in some robust sense dependent on or determined by physical-biological properties. . . . [That is,] [t]he irreducibility claim is a negative thesis; nonreductive physicalists want a positive account of the relationship between the two sets of properties.35

In other words, it is not enough that persons – the distinct sort of things that possess mental properties – are ‘physical’ in the sense that they are made of the same fundamental physical things as everything else. Nor is it enough that the physicality of persons in this sense guarantees the ontological primacy of fundamental physical things.36 Apart from these claims, most nonreductive physicalists also want to make some claim – some claim beyond irreducibility – about how the mental properties possessed by a person relate to the physical-biological properties possessed by her body. And for many nonreductive physicalists, Kim points out, “[T]he catchwords are ‘dependence’ and ‘determination.’” 37 That is, many nonreductive physicalists want to claim not only that a person’s thinking some thought depends on her body’s being in some state or other but also that her body’s being in a certain state causally determines

35 Kim, "The nonreductivist's troubles with mental causation" at 340-41.
36 I mean ontological primacy in the sense that if the world were somehow emptied of all fundamental physical things, it would be emptied of all persons (and everything else) as well.
37 Kim, "The nonreductivist's troubles with mental causation" at 341.
her thinking a certain thought. Typically, as Kim notes, these two claims – the
dependence claim and the determination claim – have been captured by supervenience
and physical realization, respectively. But this is where the troubles begin for
nonreductive physicalists. Specifically, Kim argues that their commitment to physical
realization entails a commitment to downward causation and thereby dooms their attempt
to accommodate mental causation by appealing to the layered model.

Let us take a closer look at Kim’s argument. Kim maintains, correctly I think,
that all plausible forms of nonreductive physicalism are committed to physical
realization. In the most generic terms, physical realization involves the claim that:

Whenever a person instantiates a mental property \( M \), her body has some
physical property \( P \) such that her body’s having \( P \) realizes her having \( M \).

But what does it mean for someone’s thinking a thought to be “realized” by her body’s
having some physical-biological property? Kim proposes what he calls the Physical
Realization Thesis (or “PRT”). PRT holds:

When some physical-biological property \( P \) of a person \( S \)’s body is said to
“realize” \( S \)’s possession of some mental property \( M \), \( P \) must specify a
microstructural property of \( S \)’s body that provides a causal mechanism for
\( S \)’s possession of \( M \).\(^{38}\)

Crucially, as Kim understands PRT, the physical realizer of \( S \)’s possession of \( M \) – i.e. the
microstructural property of \( S \)’s body specified by \( P \) – refers to a causal mechanism that
not only is, in some sense, ‘for’ \( S \)’s possession of \( M \) but also is itself causally sufficient

\(^{38}\) Again, this is a paraphrase modified to reflect the richer version of the layered model
but still intended to capture the essence of Kim’s Physical Realization Thesis. In Kim’s
own words, PRT reads:

When \( P \) is said to “realize” \( m \) in system \( S \), \( P \) must specify a
microstructural property of \( S \) that provides a causal mechanism for the
implementation of \( M \) in \( S \).

Ibid at 343.
for S’s possession of M. I must emphasize that Kim’s understanding of PRT is in no way idiosyncratic. Otherwise, his argument would lack the force and interest that it has.

With this as background, here is Kim’s argument reconstructed as follows:

1. S’s having some mental property M causes S to have another mental property M*.
2. S’s having M* is “physically realized” in Kim’s sense of PRT.
3. S’s having M* is caused by her body’s having some physical-biological property P*.
   (From Kim’s sense of PRT, 2)
4. S’s having M* has two independent causes, M and P*. (From 1, 3)
5. But S’s having M* cannot have two independent causes.

6. Either the assumption that there is mental causation (as in premise 1) or the assumption that mental properties are physically realized in Kim’s sense of PRT (as in premise 2) is false.

7. That mental properties are physically realized in Kim’s sense of PRT (i.e. premise 2) cannot be false.

8. That there is mental causation (i.e. premise 1) must be false. In other words, there is no mental (qua mental) causation.

On my reconstruction, then, Kim offers us a reductio. Specifically, he reduces nonreductive physicalism to the absurdity that a person’s having some mental property is systematically overdetermined. Certainly Kim is correct to hold – for reasons discussed above in section 5 – that systematic overdetermination is too high a price to pay for mental causation even if systematic overdetermination is not, strictly speaking, incoherent. Thus, Kim is also right to conclude that one of the two premises upon which his argument is based must be false since, taken together, they imply systematic overdetermination.

Controversy arises over Kim’s view as to which of these two premises must be false. Kim apparently believes (as reflected in premise 7) that the commitment to his
understanding of PRT is on surer footing than the commitment to mental causation. But unlike Kim’s interpretation of PRT which is loaded with philosophical baggage, the commitment to mental causation is pre-theoretical. It is a datum from which we begin when we do philosophy. Without claiming that such datum can never be overturned by philosophical reflection, surely we ought to inspect our baggage before doing so. Consequently, we might just as reasonably conclude from Kim’s reductio that PRT – at least as Kim interprets it – needs to be rethought.

Moreover, even apart from his reductio, we should question Kim’s interpretation of PRT. On his understanding of PRT, the physical realizer of a person’s having some mental property is also supposed to be a sufficient cause of her having this mental property, regardless of the kind of mental property involved. However, among the mental properties that persons possess, we commonly distinguish the purely qualitative from the propositionally contentful. And if we accept content externalism for propositionally contentful mental properties – as seems plausible – then PRT in Kim’s sense fails to hold with respect to contentful mental properties of this sort. For granting that a person’s body instantiates some physical-biological property $P$ when the person has some propositionally contentful mental property $M$, there is no $P$ – given content externalism – such that her body’s having $P$ is causally sufficient for her having $M$. That is, according to content externalism, the thoughts that a person has causally depend not just upon what is going on within her body (say, in her brain) but also upon what is going on in the world around her. Therefore, as it stands, Kim’s interpretation of PRT does not hold.
One obvious amendment would restrict Kim’s interpretation of PRT to purely qualitative mental properties. Such a restriction would arguably make Kim’s interpretation of PRT true. For at least with respect to purely qualitative mental properties, it is plausible to think that the physical realizer of a person’s instantiating such a property is also a sufficient cause of her doing so. Take, as examples, someone’s feeling pain or someone’s feeling depressed. In both cases, our belief that there exists a pharmacological remedy – say, aspirin and Prozac, respectively – depends upon our belief that there is an underlying (i.e. physical-biological) etiology with which the said remedy interferes in some way. Assuming, then, that we limit Kim’s interpretation of PRT to purely qualitative mental properties (and thereby save it), does this support Kim’s charge that nonreductive physicalists cannot accommodate mental causation? On one hand, the proposed amendment would mean that Kim’s argument goes through with respect to purely qualitative mental properties. Thus, a person’s possessing a purely qualitative mental property does not possess causal powers qua such an event. To this extent, Kim’s case against nonreductive physicalism is bolstered. On the other hand, nonreductive physicalists can maintain that the epiphenomenal status of a person’s feeling pain is not something they need to rebut. Rather, the kind of mental causation that nonreductive physicalists need to accommodate involves a person’s having propositionally contentful mental properties such as believing and desiring. For it is the possession of propositionally contentful mental properties that is involved in the mental causation to which we are so deeply committed.

Some might object to my claim that nonreductive physicalists need only worry about the causal powers of events involving propositionally contentful mental properties.
We are also – according to this complaint – pre-theoretically committed to causal powers for events involving purely qualitative mental properties. For instance, a person’s feeling pain might cause her to take two aspirin. Presumably we want to acknowledge this, but acknowledging this does not commit us to causal powers for events involving purely qualitative mental properties. *Qua* these properties, events involving them are epiphenomenal. Strictly speaking, the only kind of events that produce actions within the scope of commonsense psychology are events that involve propositionally contentful mental properties such as believing and desiring. Thus, when someone asks ‘Why did Ally take two aspirin?’ and we respond ‘Because she was feeling pain’, our response should be seen as elliptical. We elide what is really doing the causal work, namely, Ally’s desiring that she get some relief for her pain and her believing that she will achieve this by taking two aspirin. Granted, Ally’s feeling pain figures in the complex event that causes her to take the aspirin. That it does represents the truth in our initial, elliptical explanation of her aspirin-taking. But notice how Ally’s feeling pain figures in the causal story: it is embedded within the scope of the relevant propositional attitude(s) – in this case, a belief and a desire – and it is her having these propositionally contentful mental properties that causes her to take the aspirin, not her feeling pain in and of itself. Moreover, this point can be generalized. Whenever someone instantiates some purely qualitative mental property, such an event never enjoys causal powers in and of itself but only insofar as it figures within the scope of an appropriate set of propositional attitudes had by the person.

But what about all the expressive behavior (e.g. crying, wincing, etc.) and avoidance behavior (e.g. pulling one’s hand back) that appears to causally follow upon
pain? Surely this is something to which we are pre-theoretically committed. Nonreductive physicalists, however, need not abandon this commitment. On the layered view I propose here, they can acknowledge causal connections between instances of pain and these sorts of events provided they confine these connections to the bodily level as opposed to the personal level. So, it is not my feeling pain that causes me to jerk back my hand. Rather, it is my body’s being in pain (i.e. my body’s possessing some physical-biological property such as tissue damage) that causes a reflexive response by my body such as pulling my hand back. Notice that locating causal relations involving pain at the bodily level means that such relations are not mediated by any thought but instead are brutely causal. That is, in Kantian terms, such relations are not rational but (merely) empirical. This, I think, is a virtue of the view. It highlights the significant difference between, on one hand, my doing something in the everyday, intentional sense because I have certain beliefs and desires and, on the other hand, my body’s responding in a reflexive, automatic way because part of my body suffers tissue damage.

My doing something in the everyday, intentional sense because I have certain beliefs and desires: this is mental causation, strictly speaking. It is also a phenomenon that the layered model can happily accommodate provided we remember several key points. First, as with every genuinely distinct level in the layered model, the personal (or commonsense psychological) level involves not only distinct properties but also distinct individuals. This ensures that we do not build token-identity into the layered model and thereby violate CAC. Second, the behavior produced by mental causes is taken at face-value – i.e. understood as behavior in the sense of everyday, intentional actions – rather than reduced to complexes of mere bodily motion. This also ensures the integrity of the
layered model and CAC. Finally, only some mental events – i.e. those comprised of persons instantiating contentful mental properties – are supposed to be causally potent in their own right. By contrast, events, comprised of persons instantiating purely qualitative mental properties are acknowledged to be epiphenomenal. In this way, since the usual understanding of physical realization – one which involves causal determination from below – is not true of contentful mental properties, the integrity of the layered model and CAC are again protected. Thus interpreted, we can say that the layered model not only appears to be but is “an elegant way of averting violation of the causal closure of the physical” and making room for mental causation.

7. However, before nonreductive physicalists can have their layered cake and eat it too – that is, before we can legitimately claim that the layered model allows nonreductive physicalists to happily accommodate mental causation – some remaining concerns about dualism must be addressed. For it is not enough to explain how the layered model, by locating persons with their thoughts and actions in a causally autonomous and self-contained realm, can make room for mental (qua mental) causation. We must also show that the mental realm, while causally autonomous and self-contained, is nevertheless tied to the physical in a sufficiently strong sense. Otherwise, we may have found a place for mental causation, but it will not be in a physical world. Put another way, it needs to be demonstrated – not just asserted – that the nonreductivist position I have put forward also deserves the title physicalist.

The required demonstration involves addressing two remaining concerns about dualism. They are not unrelated. First, discussing Kim’s critique in the previous section, we saw that for the layered model to secure the mental’s causal autonomy and
containment and thus make room for mental (*qua* mental) causation, a common view of what it means for the mental to be physically realized must be rejected. But since a commitment to physical realization is a *sine qua non* of the nonreductive physicalism being advocated here, the concern arises as to whether nonreductivists (such as the author) who would accommodate mental causation by rejecting this common view of physical realization can still legitimately claim to be physicalists. Kim puts the charge directly: “I challenge these nonreductivists . . . to state an alternative principle on just how the causal powers of a realized property are connected with those of its realization base; or explain, if no such connections are envisioned, the significance of the talk of realization.”

Physical realization, on the view being rejected, involves not only the dependence of the mental on the physical but also the causal determination of the mental by the physical. This causal determination from below is the problem. It clearly violates the mental’s causal autonomy and containment. So, as Kim’s challenge suggests, the concern is really this: assuming we must accept physical realization in some sense (in order to be physicalists) but we cannot accept the sense that involves causal determination from below, is there another sense in which the mental is physically realized that satisfies the demands of physicalism?

Before proceeding, we should be clear about why physical realization is a *sine qua non* of the nonreductive physicalism being advocated here. For given the focus on supervenience in our discussion of Haugeland’s failure to capture the legitimate, intuitive appeal of physicalism with only a global version of this notion, there may be a temptation to think ‘Why do we need physical realization at all? Why isn’t supervenience alone

---

39 Ibid at 355.
sufficient to secure our physicalist credentials?’. To understand the importance of physical realization and why this tempting thought does not help, consider some aspects of supervenience that have only been hinted at until now. At best, supervenience only serves to bring out the implications of some dependence relation that is supposed to hold between the relevant domains anyway. By itself, the fact that one domain supervenes on another merely shows that the latter co-varies with the former. Without some explanation for this co-variance, it could just be a miraculous coincidence that the mental supervenes on the physical – not something that sits well with physicalism. Therefore, physicalism requires an explanation for why supervenience holds and the fact that the relevant domains stand in some dependence relation independent of supervenience is typically thought to provide it. Physical realization looms large for the nonreductivists being considered here because they, unlike physicalists who explain supervenience by appeal to some form of the identity relation, reject any such appeal given their embrace of the layered model in all its richness. Therefore, if these nonreductivists are to explain supervenience and earn the title of physicalist, they have nowhere else to look but to the relation of physical realization.

Nonreductivists who accept the layered model in all its richness and, accordingly, reject even the token-identity of the mental with the physical can still appeal to the composition relation to yield a sense in which the mental is physically realized. But does such an appeal yield a sense of physical realization in which the mental depends on the physical strongly enough to capture the legitimate, intuitive appeal of physicalism? The second remaining concern about dualism is that the answer to this question is ‘No’.

40 I established the availability to these nonreductivists of an appeal to composition in section 5.
Indeed, given content externalism and the fact that mental causation involves persons instantiating contentful mental properties, the worry is that appealing to composition fails to yield any sense in which the causally potent aspects of the mental depend on the physical, let alone one strong enough to support physicalism.

That persons are each composed of their own respective human body (i.e. a body with certain essential physical-biological properties) – call this the fact of composition. The fact of composition implies that the mental is physically realized at least in the sense that if a person is thinking a thought or doing an act then there is a body – her body – with some corresponding physical-biological property. In other words, the fact of composition conjoined with the plausible idea that a human body never goes “naked” (i.e. always instantiates some physical-biological property or other) sets up a de facto correspondence between instances of persons thinking thoughts or doing things and instances of their respective bodies possessing some physical-biological property. But notice that based on the fact of composition alone, this correspondence need not be one-to-one. In other words, consistent with the fact of composition, what someone is thinking or doing may change without any change in the physical-biological properties of her body. For even so, it would still be true that whenever someone thought or did something, her body would possess some physical-biological property or another. Consequently, physical realization as derived from the fact of composition does not even establish co-variance, to say nothing of the dependence required for physicalism.

Physical realization in this sense means mere embodiment.

---

41 That is, the correspondence between instances of persons thinking thoughts or doing things and instances of their respective bodies possessing some physical-biological property need not be one-to-one. By contrast, the correspondence between persons and bodies must (in an important sense of ‘must’) be one-to-one.
Moreover, content externalism makes it difficult to see how physical realization in the sense of mere embodiment can be supplemented to secure dependence of the mental on the physical-biological. For, as we observed earlier, if content externalism is true, there is no physical-biological property of a person’s body/brain that causally necessitates her thinking any specific thought. So, taken together, the two concerns about dualism that we have been discussing possess some force.

Two possible responses quickly come to mind. Neither of them, however, prove helpful. The first and perhaps most obvious stems from the recognition that, without further specification, content externalism does not exclude the existence of narrowly contentful mental properties. The guiding idea behind narrow content is this. For all contentful mental properties the possession of which depends on environmental and/or social factors external to a person’s head, there are supposed to exist corresponding content properties that depend entirely on factors internal to the relevant person’s head. By contrast, these latter mental properties are what I call narrowly contentful. Significantly for our purposes, a person’s instantiating a narrowly contentful mental property – unlike her instantiating other contentful mental properties – would be causally necessitated by her body/brain’s instantiating some physical-biological property, at least as narrow content is standardly conceived. Thus, espousing narrow content entitles one to claim that not only qualitative but also contentful mental properties are physically realized in Kim’s sense, a sense clearly strong enough for physicalism.

However, for the nonreductivist position I am defending, espousing narrow content would mean abandoning what makes this position interesting and important. The
whole point of this position is to accommodate mental *qua* mental causation by invoking
the layered model.

Accommodating mental *qua* mental causation involves recognizing the distinct reality of
persons, contentful thoughts of the sort that we pre-theoretically take persons to have,
events comprised of persons having such thoughts, and ultimately causal powers for such
events. All this the layered model enables us to do. But espousing narrow content as it is
standardly conceived would fatally undermine this effort. For if there are, in fact, narrow
contents corresponding to the contentful mental properties that we pre-theoretically take
persons to have, then mental events of persons having contentful mental properties – e.g.
the appropriate belief-desire combinations – would not possess causal powers in their
own right. Instead, any causal powers we might attribute to such mental events would,
strictly speaking, belong to events at the bodily level via something like Kim’s Causal
Inheritance Principle.42 The causally empowered events would be comprised of persons’
respective bodies instantiating whatever physical-biological properties realize the narrow
content properties involved. But then we would have given up trying to accommodate
mental *qua* mental causation. For this reason, the nonreductivists I am defending cannot
turn to narrow content to secure the mental’s dependence on the physical. They must
look elsewhere.

If appealing to narrow content would tie the mental to the physical too closely, the
second way that we might try to strengthen the mere embodiment sense of physical
realization errs in the opposite direction. The second way of strengthening the mental’s
tie to the physical involves dramatically expanding the base of properties with which the

42 Kim, "The nonreductivist's troubles with mental causation" at 355.
physical is characterized. Animating this second approach lies the thought that in order to say generally what properties of a person’s body determine what she (the person) is thinking, we need to look beyond the physical-biological properties of the person’s body since these properties are primarily non-relational. For given content externalism without narrow content, if any bodily properties determine what thoughts persons think, they will have to include a wide range of relational properties such as having been in, say, an H₂O environment. Accordingly, this approach secures the determination of the mental by the physical by embedding a person’s body in a particular world with a particular history, thereby including the gamut of relational properties in the physical base. In other words, by fixing the entire world and its history as it relates to a person’s body, this approach would fix what contentful mental properties the person has.

This “world-wide” approach for tying down contentful mental properties essentially tracks Haugeland’s approach in “Weak Supervenience” criticized in section 4. Not surprisingly, it suffers from the same general disability, namely, how loosely it ties the mental to the physical. Defenders of the world-wide approach, however, would claim that when it comes to contentful mental properties, a loose tie between the mental and the physical is no disability. With respect to purely qualitative properties such as pain, we complained that Haugeland’s world-wide approach allowed any sort of remote difference in the world to determine whether a person felt pain and not just the sort of local factors – e.g. whether the person’s body suffered tissue damage – that we intuitively regard as determinative. But given content externalism sans narrow content, such a criticism lacks force against the world-wide approach as applied to contentful mental properties. That

---

43 Even to the extent that the physical-biological properties of a person’s body are relational, the relational properties involved are too limited.
remote, non-local factors determine contentful mental properties is just what content externalism *sans* narrow content involves. Thus, defenders of the world-wide approach argue that if content externalism *sans* narrow content must be accepted – as it must by nonreductivists who want to protect the integrity of their layered model – then complaints about how loosely the world-wide approach ties contentful mental properties to the physical have no basis. For given our commitment to content externalism *sans* narrow content, the world-wide approach is the best we can do.

Still, whether we *need* to do better is the real question. The *ad hominem* defense of the world-wide approach just rehearsed should not mislead. Independently of what we think about content externalism *sans* narrow content, we need to know whether the looseness of the tie between the contentful mental and the physical afforded us by the world-wide approach is a problem. The answer is ‘Yes’. For just as we need to be able to explain why, with respect to a person suffering from severe pain, doctors inject morphine into *her* body, so we need to be able to explain why, with respect to a person suffering from violent, aggressive or otherwise “anti-social” thoughts, psychiatrists perform surgery on the frontal lobes of *his* brain. In other words, there is an important sense in which a person’s having certain contentful thoughts is determined locally just as her having purely qualitative thoughts is determined locally. So, when lobotomies were done in the past, they were not simply performed on anybody but rather on the frontal lobes of the person whose thoughts and behavior were at issue. Similarly, when someone suffers a severe brain injury, we reasonably expect that *she* will lose – at least temporarily – some degree of mental functioning including the ability to think certain contentful thoughts. The sort of local dependence reflected in these examples: this is the legitimate,
intuitive appeal of physicalism when it comes to contentful mental properties. But the world-wide approach can no more capture the local dependence of contentful mental properties than it could for purely qualitative mental properties. Consequently, we need a better (i.e. tighter) account of mental-physical relations than the world-wide approach provides: not only to earn our credentials as nonreductive physicalists but also (and more importantly) to ensure the basic plausibility of our view.

Given our commitment to content externalism sans narrow content, however, we might appear stuck with the unacceptably loose view of mental-physical relations reflected in the world-wide approach. Fortunately, consistent with espousing content externalism sans narrow content, we can do better. Consider the sort of local determination illustrated by my examples of lobotomy and brain injury. What my examples indicate – and all that we need to capture the legitimate, intuitive appeal of physicalism vis-à-vis content properties – is that a person’s capacity to possess a certain range of contentful mental properties is determined locally. This sort of local determination can be had without much trouble. We need only modify our idea of the mental property (the having of) which is supposed to be determined locally: from a person’s thinking some specific thought to a person’s possessing the capacity to think some range of thoughts.
CHAPTER THREE
Waking Particularism from Its Dogmatic Slumbers:
From Murdochian Description to Wittgensteinian Criteria & Back Again

1. Roughly speaking, moral particularism (or simply ‘particularism’ as I will call it) has been described by some of its adherents as the view that rejects all moral principles.\(^1\) While this way of thinking of particularism enjoys the virtue of simplicity, it obscures important differences among those who defend particularism. Accordingly, it helps to distinguish particularism about moral principles from both particularism about moral rules and particularism about moral reasons. Though it is common to distinguish one or two of these varieties of particularism, no one – to my knowledge – has explicitly distinguished all three.\(^2\) I now proceed to explain what each variety of particularism involves and how they relate to each other.

---


Particularism about moral rules focuses upon how people arrive at moral decisions. In the words of Roger Crisp, this form of particularism “claim[s] to …
point[…] to certain limits on the actual application of rules in moral reasoning, and … bas[e] [its] recommendation of an ideal of moral agency on [its] appreciation of those limits.” (24) ³ What are these limits? Pointing to McDowell and Martha Nussbaum as defenders of these limits, Crisp writes, “[The] complexity and unpredictability of human affairs are such that circumstances will arise in which … [any system of] rules [is] inappropriate for determining what we should do.” (27) Accordingly, the ideal moral agent will “be able to work out what to do on each occasion as it arises” only by possessing a sensitivity both to the morally relevant features present on such occasions and to how these features combine in the particular case to make a certain action right or wrong. (27) Particularism about moral rules, then, holds that “[M]oral rules are not sufficient on their own to provide moral guidance.” (25)

Of course, if particularism amounted to nothing more than this – skepticism about the sufficiency of moral rules for providing guidance in particular cases – it would not be very interesting. For “any plausible moral theorist,” as Crisp observes, “is likely to put weight on … the notion of a sensitivity to the morally salient features of situations.” (28) In other words, even if one thought that the nature of morality is such as to yield a straight-forward decision-procedure and even if such a procedure is otherwise desirable

³ Roger Crisp, "Particularizing Particularism" in Moral Particularism, ed. Hooker and Little (2000), 23-47. (The numbers in parentheses throughout this section refer to pages of this essay.) Given the reading of Wittgenstein’s rule-following considerations that I explicate and defend in Chapter Five, Crisp’s characterization of particularism strikes a false note, especially as to McDowell’s view. It is not rules per se that are the problem but a certain over-inflated conception of rules. Accordingly, Crisp gets particularism right only if he is understood as referring to rules in the problematic (over-inflated) sense.
to employ in one’s deliberations, one would still need to be sensitive to the morally relevant features of situations in order to apply it successfully.\(^4\) Thus, Crisp concludes, “The claim that practical generalizations [i.e. moral rules] run out may be accepted by ethical theorists of any stripe.” (28)

Despite this attempt at assimilation, particularism about moral rules is not as innocuous as it may appear. Indeed, focusing upon the limits that particularists would place on the use of moral rules in moral deliberation risks missing the point. What matters here is not so much the respective roles, if any, that moral rules and moral judgment should play in moral deliberation but rather the priority – the conceptual priority – of particular cases vis-à-vis moral rules. Ironically, Crisp mentions this aspect of particularism only in passing.\(^5\) While I do not intend to defend this priority of the particular now – I will later – I do want to highlight how this commitment distinguishes particularism. Traditionally, moral theories purport to give at least the nature if not the meaning of moral concepts. Accordingly, if one understands what the true moral theory has to say about moral concepts, one possesses some understanding of moral concepts even if one lacks the moral judgment (or sensitivity) to apply the moral theory in particular cases. Given its emphasis on the conceptual priority of the particular case, however, particularism denies the existence of such a purely theoretical form of moral understanding. Rather, for particularists (at least those who accept the conceptual priority of the particular case), whatever understanding of moral concepts one may possess must

---

\(^4\) Also responding to McDowell, Samuel Scheffler makes essentially the same point. See Samuel Scheffler, *Human Morality* (1992) at 39-43, especially 43. Still, I regard Scheffler’s efforts to blunt the force of McDowell’s particularism as no more successful than Crisp’s for the reason I give in the next paragraph.

\(^5\) Crisp, "Particularizing Particularism", at the top of 25 citing Nussbaum.
derive from understanding moral concepts in particular circumstances. Put another way, while traditional moral theorists regard moral rules as either expressing the nature of morality directly or deriving from a more general principle that expresses this instead, particularists (at least those who accept the conceptual priority of particular cases) regard moral rules as mere sketches or summaries of moral concepts. Clearly, then, particularism about moral rules need not be innocuous, pace Crisp.

I now turn to the other two varieties of particularism mentioned earlier. Both particularism about moral principles and particularism about moral reasons focus not upon those often derivative moral principles that supposedly figure in moral deliberation – i.e. moral rules – but upon more general moral principles that, according to one moral theory or another, are supposed to lay bare the nature, if not the meaning, of moral concepts. Or, in terms of the two traditional aims of moral theorizing identified by Mark

6 Of course, this does not mean that moral principles are useless or irrelevant. See Dancy, Moral Reasons at 67. Or consider Jay Garfield who writes, “On my view moral principles are not the objects of moral knowledge at all, but are at best dialectically useful summaries of that which we know.” Jay Garfield, "Particularity and Principle: The Structure of Moral Knowledge" in Moral Particularism, ed. Hooker and Little (2000), 178-204 at 198-199, n.40. Garfield continues:

To be sure, we often do resort to moral principles in our reasoning and disputation concerning difficult moral decisions. But this is only because we use principles in order to draw our attention to relevantly similar cases, and to particular dimensions of similarity, not because we believe that those principles have force of their own independent of examples. That is, we use principles in such cases as a bulwark against arbitrariness not by using them as universals under which to subsume the new case, but rather as guides to other cases with which to pair it in our consideration. I am hence urging a casuistical and not a Kantian account of the role played by moral principles. Ibid.

While I endorse this account of the role of moral principles, Garfield’s sole emphasis on the epistemic priority of particular cases threatens to neglect a more important kind of priority they possess, namely, conceptual/semantic priority. As promised, I will say more about this later.
Timmons\textsuperscript{7}, particularism about moral rules (at least in its innocuous, Crispian form) represents – as we have seen – skepticism about the first, practical aim of finding a moral decision-procedure while the latter two kinds of particularism represent skepticism about the second, theoretical aim of limning the underlying nature of morality. As Margaret Little picturesquely describes it, the theoretical aim of moral theorizing involves:

[T]he search for an architecture of morality. It is a search for the patterns that lie behind specific cases; more than that, it is the search for generalizations that hold not as mere accidents (as when we realize that Dora is usually cruel to Jack), but with a level of necessity, revealing the very nature or structure of the moral landscape.\textsuperscript{8}

Moreover, this search for universal generalizations – or principles – that hold with some level of necessity can focus on several different sorts of connections. Little continues:

Just as scientists try to parse out how the forces of physics interact systematically, moral theorists try to capture how moral considerations so identified (such as the requirements of justice and beneficence) are ordered in relation to each other. And again, just as scientists work to unearth laws linking, say, the property of temperature to the property of mean kinetic molecular energy, the job of moral theorists is to identify which natural properties make an action count as just or beneficent.\textsuperscript{9}

In other words, one sort of connection that moral theorists might try to establish would link combinations of \textit{pro tanto} moral considerations (e.g. that an action is beneficent and that it involves lying) to overall verdicts of right and wrong. But moral theorists also try to link natural – or better, non-moral – properties to \textit{pro tanto} moral considerations, showing us what underlying properties make actions just, cruel, beneficent, etc.

\textsuperscript{7} Timmons, \textit{Moral Theory: An Introduction} at 3-4.


\textsuperscript{9} Little, "Wittgensteinian Lessons on Moral Particularism" at 164.
A third sort of connection, though not mentioned by Little, may also be the object of moral theorizing. More ambitious moral theorists have sometimes tried to link non-moral properties directly to overall verdicts of right and wrong, effectively bypassing the two (intermediate) sorts of connections described by Little. Moral theorists are ambitious in this sense if they are monists about the fundamental features that are supposed to make actions right or wrong. That is, if a moral theorist claims “there is a single basic feature of actions that determines their deontic status,” she will search for the ‘architecture of morality’ in terms of this third sort of connection. By contrast, less theoretically ambitious pluralists direct their search for the architecture of morality towards the two sorts of connections envisaged by Little.

Compared to monists and pluralists, however:

Moral particularists have a very different picture of morality. Without questioning whether there are moral answers, they question whether those answers are, or need to be, backed by any architecture, whether they constitute pieces that fit together to form a structure with recoverable lawlike relations. In other words, unlike monistic and even pluralistic moral theorists, particularists deny that there are any of the lawlike connections sought after by these theorists. Thus, they deny the moral principles corresponding to each of the different sorts of connections we have been discussing. This is why I earlier said that particularism could be described as the view that rejects all moral principles. Nevertheless, for sake of clarity, what I have called ‘particularism about moral principles’ refers to the rejection of only two sorts of principles: those that would link some non-moral property to overall moral verdicts and those that would link combinations of pro tanto moral considerations to overall moral

---

10 Timmons, Moral Theory: An Introduction at 190.
11 Little, "Moral Generalities Revisited" at 278; Little, "Wittgensteinian Lessons on Moral Particularism" at 164.
particularism. ‘Particularism about moral reasons’ is reserved for the rejection of the remaining sort of principle, principles that would tie *pro tanto* moral considerations to underlying non-moral properties, respectively.

2. Reviewing the discussion to this point, I began with the idea that particularism is the view that rejects all moral principles. I then introduced three varieties of particularism – particularism about moral rules, particularism about moral principles, and particularism about moral reasons – each of which rejects moral principles in some form, notwithstanding the fact that only one is labeled ‘particularism about moral principles’. Now most debate concerning particularism has focused on its rejection of the sort of principles that would tie *pro tanto* moral considerations to underlying non-moral properties, that is, debate has centered on what I am calling ‘particularism about moral reasons’. This should not be surprising since the rejection of this sort of moral principle appears to be what is most distinctive about particularism. For even moral views that also reject other kinds of moral principles – moral pluralism, for example – do not go so far as to reject principles tying *pro tanto* moral considerations to underlying non-moral properties. Therefore, this kind of principle looms large in what follows.

However, the previous section also suggested an aspect of particularism that does not (or need not) involve rejecting moral principles at all and, thus, does not fit well into the “Anti-Principle” schema for understanding particularism. This is the idea that particular cases of right and wrong enjoy conceptual priority vis-à-vis moral rules, moral principles, and moral theory generally. Given this alternative idea of particularism, the words of J.L. Austin seem prescient when applied to the initial idea of particularism as the “doctrine” that rejects all moral principles. Austin writes:
One might well want to ask how seriously this doctrine is intended, just how strictly and literally the philosophers who propound it mean their words to be taken…. It is, as a matter of fact, not at all easy to answer, for, strange though the doctrine looks, we are sometimes told to take it easy—really it's just what we've all believed all along. (There's the bit where you say it and the bit where you take it back.)\textsuperscript{12}

We have now come to “the bit where you take it back,” where ‘it’ is the idea that particularism rejects all moral principles, especially principles that tie \textit{pro tanto} moral considerations to underlying non-moral properties. Crucially, however, particularists can take back their rejection of these moral principles in one of two ways. I have already alluded to the first. That is, if particularists insist upon the conceptual priority of particular cases vis-à-vis moral principles, then they have no need to reject such principles outright. Rather, they merely put these principles in their proper (subsidiary) place. However, the second way that particularists “take it back” has received more attention.\textsuperscript{13} Particularists following this path recognize principles tying \textit{pro tanto} moral considerations to underlying non-moral properties provided such principles are qualified with a \textit{ceteris paribus} clause. On this view, principles tying \textit{pro tanto} moral considerations to underlying non-moral properties are perfectly legitimate so long as they are interpreted as defeasible (rather than strict) generalizations.

As between these two ways in which particularists can retain some notion of moral principles – and which I shall refer to as the criterial view and the defeasibility

\textsuperscript{12} J.L. Austin, \textit{Sense and Sensibilia} (1962). Of course, Austin was referring to another doctrine, namely, indirect perceptual realism.

\textsuperscript{13} Mark Lance and Margaret Olivia Little, "Particularism and Anti-Theory (ms.)" ; Little, "Moral Generalities Revisited" ; Little, "Wittgensteinian Lessons on Moral Particularism" ; Margaret Olivia Little, "On Knowing the 'Why': Particularism and Moral Theory," \textit{Hastings Center Report}, 31 (2001): 32-40; Mark Lance and Margaret Olivia Little, "Mad Dogs and Englishmen: Moral Valence, Defeasibility, & Privileged Conditions (ms.)", (2002).
view, respectively – I will show that the first way (the criterial view) is superior to the second way (the defeasibility view) in several important respects. Not the least of these respects is the fact that only the criterial view can allow for and make sense of a connection between particularism and anti-Humeanism. However, I am getting ahead of myself. We first need to consider in more depth the unreconstructed view of particularism and what it involves besides the rejection of all moral principles. For both of the alternatives that we shall evaluate – that is, the defeasibility and criterial views – take their cue from the unreconstructed view. Thus, we need a better picture of the unreconstructed view to motivate not only the need to move beyond this view but also the debate between the alternatives.

3. Recalling the most distinctive aspect of particularism as seen through the lens of the “Anti-Principle” schema – that is, its rejection of principles that would tie pro tanto moral properties to underlying non-moral properties – consider what the leading “Anti-Principle” particularist has to say about such (Rossian) moral principles. Dancy writes:

… Ross’s principles do no more than specify a property as being one which counts generally in favour of (or against) any action that has it. This is a much more convincing account of what a moral principle says. And it gives a quite different picture of what a counter-example to a moral principle would look like. Instead of being an example where the principle tells us to do one thing and we think we ought to do the opposite (‘Do not steal’), it would be an example where, though the principle tells us that some feature counts in favour of any action that has it, we think it either makes no difference at all here or else that it does make a difference, but counts in the opposite direction.14

From this passage, we get a clearer picture of the moral principles that unreconstructed particularists deny outright. The moral principles at issue pick out certain properties that actions can have (e.g. the property of causing pain) and assert that, everywhere such a

14 Dancy, Moral Reasons at 95-96.
property atttaches to some action, it has the further pro tanto moral property of counting for (or against) said action.\textsuperscript{15} This passage also brings out what unreconstructed particularists like Dancy find problematic about Rossian moral pluralism, namely, its generalism. These principles, if true, would mean that the features of an action that we cite as reasons for or against doing the action are such that they count the same way – either in favor of doing the action or against it, whichever it is – on every occasion that an action possesses such a feature.

Though not apparent from the passage we have been discussing, it should also be added that the moral principles at issue are supposed to hold true with at least metaphysical necessity and, thus, the features we cite as reasons must count the same wherever they appear on pain of failing to be genuine reasons. As Dancy puts it:

Why does Ross suppose that there are general moral truths? … [He] simply takes it as obvious that there cannot be such a thing as a stubbornly particular reason. He supposes it metaphysically impossible for a property to make a difference here without making the same difference everywhere… Ross seems not to have noticed the possibility that reasons are holistic.\textsuperscript{16}

Thus, unreconstructed particularists like Dancy can grant that some reason-giving features may \textit{de facto} turn out to possess the same moral valence everywhere they appear. What they primarily object to is the idea of requiring, as Rossian moral principles do, that all reason-giving features must behave in this way in order to qualify as reasons. In other words, generalists like Ross fail to acknowledge both the possibility of “stubbornly particular reason[s]” and, what this means for Dancy, a holistic account of such reasons.

\textsuperscript{15} Instead of ‘pro tanto moral properties’, some speak of ‘moral valences’. This comes to the same thing. In other words, a non-moral feature of an action can either weigh in favor of doing the action (positive moral valence), weigh against doing it (negative moral valence), or make no contribution one way or the other (neutral moral valence). Talk of ‘moral polarities’ also plays the same role.

\textsuperscript{16} Dancy, \textit{Moral Reasons} at 104.
To explain unreconstructed particularism’s rejection of Rossian moral principles and its corresponding embrace of “stubbornly particular reason[s],” Dancy would have us look to what he calls “Holism in the Theory of Reasons.” He tells us:

The leading thought behind particularism is the thought that the behaviour of a reason (or of a consideration that serves as a reason) in a new case cannot be predicted from its behaviour elsewhere. The way in which the consideration functions here either will or at least may be affected by other considerations here present. So there is no ground for the hope that we can find out here how that consideration functions in general … nor for the hope that we can move in any smooth way to how it will function in a different case.17

Now Dancy makes two key points here. First, a consideration that serves as a reason in one case is “stubbornly particular” in the sense that we cannot predict how it will behave in other cases based on how it behaves in the case before us. And second, holism helps explain this particularistic nature of features that serve as reasons. As David McNaughton describes this holism, “We cannot judge the effect of the presence of any one feature in isolation from the effect of the others. Whether or not one particular property is morally relevant, and in what way, may depend on the precise nature of the other properties of the action.”18 Or, in Margaret Little’s words:

A set of features that in one context makes an action cruel can in another [context] carry no such import; the addition of another detail can change the meaning of the whole… Natural features do not always ground the same moral import, which then goes into the hopper to be weighed against whatever other independent factors happen to be present. The moral contribution they make on each occasion is holistically determined: it is itself dependent … on what other non-moral features are present or absent.19

17 Ibid at 60.
18 McNaughton, Moral Vision at 193.
19 Little, "Moral Generalities Revisited" at 280; Little, "Wittgensteinian Lessons on Moral Particularism" at 165.
Still, with these accounts in mind, holism about reasons may be best elucidated by an analogy with art appreciation. The non-moral features of an action contribute to the action’s rightness (or wrongness) just as the elements of a painting (e.g. brushstrokes, colors, shading, etc.) contribute to the picture’s beauty (or lack thereof). For example, diminishing the foreshortening of a character’s foot may be just what is needed to correct and balance one painting, but this same thing done to another painting would be irrelevant or make things worse. Indeed, with respect to a genuine work of art, no one would think to generalize about such a thing. To generalize would reveal a fundamental lack of understanding about (high) art. So it is, on the unreconstructed view, for the features that we cite as reasons for (or against) doing an action. Intentionally telling a falsehood in one situation may contribute to the wrongness of what one does, but in other situations the fact that one would be lying may be irrelevant to the moral status of one’s action (e.g. when playing the game Diplomacy) or actually contribute to its rightness (e.g. when confronted by the secret police looking for political dissidents to torture). It all depends upon the particular situation just as, in the artistic case, it all depends upon the particular work of art.

---

20 Little credits McNaughton with this analogy and uses it in both of her essays cited above. In my deployment of it, however, I alter some details, inspired by the artist in Anna Karenina, Mikhailov, who paints Anna’s portrait. This part of Tolstoy’s novel was brought to my attention by R.W. Beardsmore, Art and Morality (1971).

21 This requires some qualification as I suggest with use of ‘high’ in parentheses. Only works of art that are extraordinary in some way render generalizations pointless. By contrast, garden-variety works (e.g. Robert Ludlum novels, B-movie thrillers, etc.) readily support generalizations. See Michael Tanner, "Ethics and aesthetics are -- ?" in Art and Morality, ed. Bermudez and Gardner (2003), 19-36 at 27-28.
Our discussion of unreconstructed particularism is almost complete. We only need to examine what the acceptance of holism about reasons has been thought to imply for how moral judgment is conceived. In this connection, David McNaughton maintains:

[W]e have to judge each particular moral decision on its merits; we cannot appeal to general rules to make that decision for us… [I]n trying to find out which is the right action … [w]hat is required is the correct conception of the particular case at hand, with its unique set of properties. There is no substitute for a sensitive and detailed examination of each individual case.\(^{22}\)

And Dancy adds:

The primary focus of particularism is the particular case, not surprisingly. This means that one’s main duty, in moral judgement, is to look really closely at the case before one. Our first question is not ‘Which other cases does this one best resemble?’, but rather ‘What is the nature of the case before us?’ Of course, a comparison with other cases may help us to decide how things are here, just as a long experience of car engines may help us diagnose the fault this time. But this decision or diagnosis is still essentially particular.\(^{23}\)

In other words, as suggested by the art appreciation analogy just mooted and their rejection of moral principles, unreconstructed particularists would have moral judgment focus on the particular case at hand, trying to appreciate its full significance. Moreover, the emphasis on appreciation here is important. All moral views, not just unreconstructed particularism, would have us focus on the particular case when deciding what to do. Particularism, at least in its unreconstructed variety, is different because it would make certain resources off-limits to moral deliberators. Here again is Dancy:

[O]ur account of the person on whom we can rely to make sound moral judgements is not very long. Such a person is someone who gets it right case by case. To be so consistently successful, we need to have a broad range of sensitivities, so that no relevant feature escapes us, and we do not mistake its relevance either. But that is all there is to say on the matter. To have the relevant sensitivities just is to be able to get things right case by case.\(^{24}\)

\(^{22}\) McNaughton, *Moral Vision* at 190.
\(^{23}\) Dancy, *Moral Reasons* at 63.
\(^{24}\) Ibid at 64.
The idea here, familiar to readers of some of Martha Nussbaum’s work, is that moral deliberators need to be keen perceivers. Unreconstructed particularists, however, take this idea further and maintain that a keen moral-perceptual sensitivity is all that we need to be excellent moral thinkers/deliberators. We specifically do not need to rely upon principles and the inferences they would allow us. Rather, our direct perceptual awareness of the particular situation in which we find ourselves is supposed to suffice.

4. We are now in position to consider what about unreconstructed particularism makes some particularists want to “take back” its rejection of moral principles and certain related theses. One group of these particularists puts forward something I call the defeasibility view. These defeasibility theorists, let’s call them, have two major problems with unreconstructed particularism, one epistemological and one metaphysical. Margaret Little sets forth the epistemological complaint. She writes:

[T]he real stumbling block to accepting particularism lies not in suspicion of moral discernment, but in suspicion that moral discernment is the only thing particularism has up its epistemic sleeve. After all, its central claim is that there is no cashing out when inflicting pain, say, constitutes cruelty and when it constitutes kindness: how, then, short of seeing a case in all its richness, can we know what its [i.e. pain’s] moral status is?25

She continues:

[I]t [can] sound as though we must be agnostic about any moral situation not immediately before us—as though, until we are able to see or interpret a case for ourselves, there is nothing we can say. For on this picture [i.e. unreconstructed particularism], not only would there be no inferential moral knowledge, there would be no warranted moral presumptions to bring to the world.26

In other words, defeasibility theorists protest the removal from our epistemic quiver of (at least a certain kind of) moral principles and the inferences they would allow, a protest

25 Little, "Moral Generalities Revisited" at 293.
26 Ibid at 294.
aided by their claim that – pace unreconstructed particularism – keen moral vision cannot perform all necessary moral-epistemic tasks. While defeasibility theorists appear to leave discernment some moral-epistemic work to do, it is not clear what this is. Still, they apparently accept the legitimacy of moral perception in some sense. Their main point seems to be – instead of rejecting talk of moral discernment outright – that particularism will be false to how we actually think and reason about moral concerns if it does not include a genuine epistemic-justificatory role for moral principles and the inferences they allow.

For the other, metaphysical complaint raised by defeasibility theorists, we turn again to Little but this time writing with Mark Lance. They tell us:

[Unreconstructed particularism] … gets at something deep and correct. Truth be told, though, even those who affiliate with the position often have a lingering sense of unease that the position proves too much. The view seems to imply that lying, killing, and the infliction of pain – not to mention, on some accounts, cruelty and depravity – have no more intimate connection to wrongness than do truth-telling, healing, and the giving of pleasure. After all, each, against the right context, can have a positive, negative, or neutral moral import. But the morally wise person, one would have thought, is someone who understands that there is a deep difference in moral status between infliction of pain and shoelace color, even if both can, against the right narrative, be bad-making.

---

27 Though Little admits no suspicions about discernment _per se_, many philosophers likely do find talk of moral vision, etc. “a real stumbling block to accepting particularism.” Placing so much weight on moral perception can seem like warmed-over intuitionism, “… enough to make one see red …” “… for those … who learned [their] moral philosophy as a cautionary tale beginning with Moore’s mistakes …” Richard Norman, "Making Sense of Moral Realism," _Philosophical Investigations_, 20 (1997): 117-135 at 117. However, as does Norman, I argue later that things are not as bad for particularism as all this suggests, the strong emphasis on discernment included.

28 For an illustration of what is _not_ a genuine epistemic-justificatory role for moral principles, see the references and accompanying discussion in note 7.

29 ‘Mad Dogs and Englishmen’, p. 2: Need permission from L&L to cite. (Alternatively, check to see if this has published under another title.)
Put simply, there is “a deep difference in moral status” between considerations such as lying, killing, etc. on one hand and those such as shoelace color, stepping on cracks in a sidewalk, etc. on the other. Lying, killing, etc. are considerations that typically are – and, thus, we presume they are – morally relevant in the usual way (i.e. they weigh against an action) unless some special circumstances alter this usual relevance. Shoelace color, stepping on cracks in a sidewalk, etc. typically are – and, thus, we again presume they are – not morally relevant unless some special circumstances makes them so. Moreover, these differences in moral relevance do not strike us as mere statistical coincidences. Rather, we might take them to reflect something deeper about morality. Indeed, we might think: it is part of the nature of morality, apart from special circumstances, that certain non-moral properties are morally relevant (e.g. infliction of pain) and certain non-moral properties are not (e.g. shoelace color). The problem for unreconstructed particularism is that with its wholesale rejection of moral principles, including principles that would tie non-moral properties to pro tanto moral properties, it is not able to express this basic aspect of moral life and thought. This is the metaphysical objection raised by defeasibility theorists.

To address these dual complaints, defeasibility theorists take back the claim that principles tying non-moral properties to pro tanto moral properties should be rejected along with other moral principles. With a ceteris paribus proviso, they would retain moral principles of this kind. So, for example, defeasibility theorists maintain that, in normal circumstances (that is, everything else being equal), the fact that an action involves lying counts against the action. Likewise, in normal circumstances, that an action involves giving pleasure counts in favor of it. Finally, in normal circumstances,
the color of one’s shoelaces is irrelevant to the moral quality of what one does. I refer to principles of this form as defeasible moral principles. They are defeasible in that the connections they assert between non-moral features and *pro tanto* moral properties can be defeated by abnormal circumstances. For defeasibility theorists, accepting defeasible moral principles not only resolves their two complaints but also preserves what, on their view, is “deep and correct” about particularism. To see how this works, let us consider these desiderata and their relation to defeasible moral principles in reverse order.

First, what is “deep and correct” – and presumably, worth retaining – about unreconstructed particularism according to defeasibility theorists? It is:

> [T]he view that considerations’ reason-giving force is irreducibly context-dependent … that considerations carry their moral import only holistically: a consideration that in one context counts for an action, can in another count against it or be irrelevant, and all in a way that cannot be cashed out in finite helpful terms.  

In other words, for defeasibility theorists, what is “deep and correct” about particularism is holism about reasons. But the compatibility of defeasible moral principles and holism about reasons depends upon, in Jonathan Dancy’s words, “a matter of tricky detail.”

Dancy continues:

> [P]articularists need to decide … [whether] they want to say that a property brings nothing to an action, its contribution being determined only when it is present with others, or whether they only want to say that though there is a default position, a contribution which the property makes other things being equal, that contribution can be reversed or annulled by untoward circumstances.

Unreconstructed particularists opt for the first of these options as evidenced by their embrace of the art appreciation analogy. On their view, not even defeasible moral principles can be supported since considerations in isolation “bring[] nothing to an

---

30 ‘Mad Dogs and Englishmen’ 1: Need permission to cite.
action.” This is the way, as we have seen, that unreconstructed particularism is holistic. By contrast, defeasibility theorists are able to sanction defeasible moral principles and yet remain holists about reasons by pursuing the second option mentioned by Dancy. That is, they can support moral principles tying a non-moral feature to its respective pro tanto moral property, ceteris paribus, because non-moral features are held to possess a default moral status. (The moral status that a non-moral feature enjoys in normal circumstances determines its default status.) At the same time, despite their embrace of defeasible moral principles, defeasibility theorists are holists about reasons because the principles they urge can be defeated by “untoward circumstances” (i.e. valence-switching or valence-neutralizing circumstances) and such circumstances can always arise in a new context. Thus, it is the permanent possibility of such circumstances that forces these theorists to accept holism. That is, one must consider a situation in its totality to rule out the presence of “untoward circumstances” and, correspondingly, to be certain about how some non-moral feature actually matters in the particular case at hand.

Next, consider how the admission of defeasible moral principles resolves the metaphysical complaint. Defeasibility theorists objected that morality has a nature (or underlying structure) and that this is lost to us if, following unreconstructed particularism, we simply reject moral principles across the board. They agree that strict moral principles must be rejected due to the possibility of untoward circumstances. However, defeasible moral principles are consistent with the possibility of such circumstances and arguably reflect the nature and underlying structure of morality. For defeasibility theorists, the nature and underlying structure of morality is such that certain non-moral features – lying, killing, etc. on one hand, truth-telling, healing, etc. on the other hand –
possess a default moral status (i.e. either counting in favor of or against the actions they characterize). The fact that untoward circumstances may conspire to switch or simply neutralize the default moral status of these non-moral features need not undermine the role of default moral status as part of the fundamental nature of morality.

Finally, recall the epistemological complaint. For unreconstructed particularists, we need to withhold judgment about moral situations not immediately before us; we cannot say anything until we are able to directly confront a case in all its particularity. But surely we can justifiably say, even when hearing only second-hand of, e.g., a man who holds a young child’s head in the toilet and threatens to drown the child if it fails to stop wetting its pants, that what this man does counts against his overall action. Call this the abusive boyfriend example.\textsuperscript{31} Since, \textit{ex hypothesi}, we only hear of the abusive boyfriend second-hand, our claim that what he does counts against his overall action cannot be justified by our keen moral perception. Thus, some alternative means of reaching moral conclusions must be granted. But unreconstructed particularism fails to grant this. Its failure in this regard comprises the epistemological complaint. Defeasibility theorists remedy this problem by admitting defeasible moral principles that tie non-moral features (e.g. holding a young child’s head in the toilet and threatening to drown it for wetting its pants) to \textit{pro tanto} moral properties. The admission of such principles answers the epistemological complaint because, from such principles and unproblematic judgments that circumstances are normal, we can reach the desired conclusions (e.g. that what the man does counts against his overall action).

\textsuperscript{31} I owe the abusive boyfriend example to Cora Diamond in correspondence, though I have modified it slightly from her original version.
5. Even if we grant the complaints raised by defeasibility particularists about unreconstructed particularism, the alternative preferred by these theorists has problems of its own. In a recent essay entitled “Murdoch, Practical Reasoning, and Particularism,” Elijah Millgram elucidates a number of problems facing particularism generally. While defeasibility particularism had yet to appear on the scene when Millgram wrote his essay, the defeasibility alternative to unreconstructed particularism has the same goal as the Murdochian picture that Millgram proposes, namely, providing better philosophical underpinning for particularism. Thus, the defeasibility alternative and Millgram’s Murdochian picture compete for our affections. If Millgram’s Murdochian picture addresses problems facing particularism better than both the unreconstructed and the defeasibility varieties, then these problems serve as a basis for criticizing not only unreconstructed particularism but also the defeasibility alternative. As we have seen, those expounding the latter are most exercised by epistemological and metaphysical concerns. By contrast, Millgram focuses upon the implications of particularism for practical reasons and practical reasoning. Not coincidentally, it is here that defeasibility theorists face a comparative disadvantage.

However, before we consider practical reasoning and the problem for defeasibility theorists in this area, we should consider what Millgram has to say about particularism. As he understands it:

[T]he most compelling and broadly shared moment in particularism [is] a move rather than a claim. Particularists will point out (call this the defusing move) that while a given consideration may count as a moral reason on one occasion, say for doing such-and-such, the very same consideration is on another occasion no reason for doing such-and-such, or even a reason for precisely not doing such-

--- Millgram, "Murdoch, Practical Reasoning, and Particularism,". Page numbers in this section refer to this paper.
and-such…. In a typical execution of the defusing move, the original consideration has not been *overridden* by another stronger reason … Rather, the reason behaves differently – for instance, it has a different ‘valence’ – once the background has changed. (65)

Millgram adds:

Particularists are impressed by the apparently uniform availability of the defusing move; once you get the knack of it, you are likely to feel quite confident in your ability to produce a defusing circumstance for just about any putatively general consideration. And so both proponents and opponents produce remarks like the following: ‘The leading thought behind particularism is… that the behaviour of a reason… in a new case cannot be predicted from its behaviour elsewhere.’ (65)

This being said, particularists do not provide much in the way of theoretical backing for their position. Such backing as they do supply is, in Millgram’s words, “unsatisfyingly thin.” (65) So, while the defusing move can – at least initially – be quite persuasive, particularists face significant problems without a better theoretical home.

Millgram identifies three problems facing particularists. First, particularists emphasize context and holism in explicating their view of when something counts as a reason. Not only are such gestures empty in terms of their explanatory value but, more importantly:

[A]ppeals to context-dependence … usually function philosophically as conversation-stoppers ... They direct attention to surroundings that are arbitrarily various; variety of this kind resists accounts of how change in context effects a change in (here) the force of reasons; and so they ... leave one not knowing what to think about next. (66)

Second, particularism urges antinomian conclusions: “[M]ost radically, that morality has no place for rules; more modestly, that moral rules function merely as reminders, or as generalizations about the kinds of situations we tend to find ourselves in.” (65) To their skepticism about moral rules generally – and specifically to one way that particularists have tried to make room for such rules – Millgram objects that, “Assimilating rules of
morality to [reminders] ... makes it hard to see how the rule could be deployed as a premise in an argument, and we do use rules in this way.” (66) Of another way that particularists have tried to salvage a role for rules, he argues that, “[T]reating [rules] as summaries of rough local regularities … seems similarly not to match our practice.” (66)

Thus, about particularism’s antinomian view of moral rules, Millgram concludes:

[W]e have a family of cases of particularist reasoning in which the rules do real (and not merely summarizing) work, and so we need ... a way of understanding particularism that does not preclude rules from doing work of this kind. (67)

Finally, and most importantly for our immediate purposes, particularism prompts metaphysical objections since it has been explicated in terms of “the ‘shapelessness’ of the evaluative with respect to the descriptive.” (71). While Millgram mostly dismisses these objections, he allows that:

[T]hey can be understood to express a [legitimate] worry ... that the particularist [is] abjuring patterns, and it is the systematic way in which we navigate these patterns that gives content and body to our claim that what we have on hand is a reason. When the patterns are thought of as rules, the complaint ... is that we do not know how to makes reasons intelligible except through their generality. (67)

However, for Millgram:

[T]he demand that we spell out what we mean by saying that something is a reason seems ... too deep to be addressed by a specifically ethical theory ... it would … be best to find a version of [particularism] that would allow the question to be postponed. (67)

In other words, with the theoretical home that Millgram envisions, particularists will neither have to make sense of non-general reasons nor, correspondingly, to account for what we mean by saying that something is a reason. With the right theoretical home, particularists will be able to defer these questions like theorists of other persuasions.

33 For these objections, see Jackson, Pettit and Smith, "Ethical Particularism and Patterns".
This last problem identified by Millgram provides the critical wedge for my claim that defeasibility particularism has significant drawbacks of its own. It is not, however, that defeasibility theorists lack a response to Millgram’s final problem. As he advises, they eschew Dancy’s idea of moral reasons as “stubbornly particular” and can thereby sidestep the need to account for such things. But they achieve this goal by turning to a theoretical structure borrowed from the philosophy of science. Accordingly, on their view, particularists should conceive of moral reasons in terms of defeasible generalizations. By contrast, I agree with Millgram that particularists do better by looking to Iris Murdoch’s work in moral philosophy. To support this claim, I will show how, in certain crucial respects, Murdoch’s work yields a more attractive conception of moral reasons than the defeasibility conception adapted from work in philosophy of science.

As Iris Murdoch distinctively conceived of it, practical reasoning begins and ends at an earlier stage – or, better, offstage entirely – relative to where it takes places in other, more familiar conceptions of practical reasoning. Millgram puts the contrast like this:

Most standard ways of seeing the problem space take practical reasoning to proceed from a description of a decision situation, one that is treated as simply given, to a practical conclusion…. But to Murdoch's way of thinking, the hard part of practical reasoning is getting the description of your situation right in the first place. You have to come to see your circumstances the right way, or, equivalently, to apply the right set of specialized terms, or, perhaps again equivalently, to employ the right family of metaphors; once that is taken care of, it will be obvious what to do. (emphasis added) (69)

In other words, for Murdoch, the real work of practical reasoning centers on how we describe the practical situations we face and getting these (descriptions) right. Thus, on her view, the real work has already been completed when we reach the point where other, more familiar accounts would have us begin.
Murdoch’s view is arguably quite plausible. Recall the abusive boyfriend case. Once we determine that the goings on next door are best described as this sort of case – and not, say, as an unruly child finally getting a strong male role-model and authority figure in his life – what we should do becomes clear, to wit, call child protective services. Millgram himself offers other convincing examples in Murdoch’s support.\textsuperscript{34} Beyond arguing by example, however, Millgram also points out that:

\begin{quote}
\texttt{[G]etting the problem description right is pertinent … to theoretical reasoning as well. When they are in school, the tricky part is getting the logic or physics students to convert the story problem to the right set of formulae, and after they graduate, the even trickier part is getting them to convert the situation they are facing into the right story problem. (70)}
\end{quote}

Interestingly, Murdoch’s key idea – that getting the problem description right is crucial – applies to theoretical reasoning \textit{despite the fact that}, as Millgram perceptively notes:

\begin{quote}
\texttt{[W]hen [we] think about what it takes to get the representations right in a strictly theoretical or factual domain, it is natural to start with problems in which the goal is given, and to understand the correctness of the representation in terms of its usefulness in attaining the given goal. (70)}
\end{quote}

Therefore, if Murdoch’s insight is important even in theoretical or factual contexts where a goal can properly be taken as given, it can only possess greater importance in practical or evaluative contexts where “… setting the goal is \textit{part of} … [the] problem …” (70)

Now, as Millgram realizes, some will object that the supposedly crucial task of ‘getting the problem description right’ involves nothing more than describing the facts of

\textsuperscript{34} [i] “You might take someone to be aloof and distant, and so be rather standoffish yourself; once you come to see his manner as shy, it will be natural to much more open towards him.” (69)

[ii] “It is redescribing his employer as recklessly and criminally endangering its workers, neighbors and clients that leads the whistleblower to step forward.” (69)

[iii] “It is opening up the question of whether someone is really your friend – whether he \textit{could} really be your friend, given how he had been acting – that is the most important part of figuring out how to conduct one’s future relations with him.” (69)
one’s situation, something that clearly falls within the domain of theoretical reasoning and, thus, could not be the critical moment in any bit of practical reasoning. However, this objection misses the point that, for Murdoch, “[G]etting a description right is not normally a matter of getting the metaphysically right description.” (71) Rather, as Millgram helpfully relates:

The problem of getting the right description is to see things as they really are, but “truth,” “reality,” and their paronyms, are, in Murdoch's way of using them, not to be captured by the idea of accuracy, of a man in a lab coat checking that his measurements correspond to the dimensions of the objects on his workbench. Murdoch, best known as a novelist, thinks of truth by way of novelistic truthfulness: “Truth is not a simple or easy concept. Critical terminology imputes falsehood to an artist by using terms such as fantastic, sentimental, self-indulgent, banal, grotesque, tendentious, unclarified, wilfully obscure and so on.” (70-71)

In other words, when we engage in practical reasoning à la Murdoch and try to arrive at an appropriate description of our situation, we not only can but typically should draw upon all of the vocabulary at our disposal, including a full range of evaluative terms. To think otherwise – i.e. to impose, across the board, some metaphysically-inspired limit on the concepts that we can deploy in describing our practical problems – itself constitutes the adoption of a substantive, evaluative view. Granted, there are times when we really ought to limit our resources for describing practical problems to terms less apt to generate controversy.³⁵ One such occasion might be when people from diverse backgrounds need to reach a consensus.³⁶ But notice that, even in special circumstances, any limits on our

³⁵ While limiting ourselves to non-controversial terms can be – and, in common parlance, often is – described as ‘sticking to the facts’, such talk does not support a metaphysical distinction between factual and evaluative terms. Non-controversial terms and, thus, what it means to ‘stick to the facts’ are relative to the intended audience. There is no absolute sense of ‘non-controversial’.
³⁶ Here is another example. When trying cases before juries, lawyers are often told to ‘Stick to the facts.’ This injunction represents just one of many limits that are placed upon what lawyers can say to a jury.
descriptive resources are justified practically or morally (e.g. by the good that would be
served if the group reached a consensus). By contrast, for the kind of across-the-board
limit on our descriptive resources that the objection under consideration would impose,
there appears to be no plausible practical or moral justification. Or at least, as Millgram
says, “… I do not know what the argument for always [imposing such a limit] would be.”
(71)

To rehearse a little then, Murdoch holds that the process of arriving at the right
description of our situation constitutes practical reasoning and the description at which
we arrive yields a practical reason (i.e. makes it clear what we should do and natural for
this action to follow). But Millgram next suggests, “[P]articularist defusing moves can
be recast as Murdochian moves to improved descriptions, ones in which the defusing
features play a pivotal role.” (69) Accordingly, Millgram recommends that particularists
look to Murdoch for a theoretical home. As he puts it, “Murdoch’s way of framing the
defusing move avoids some of the excesses of antinomianism.” (72) For one, it avoids
appealing to context so as to put an arbitrary end to deliberation. On Murdoch’s view,
practical (or moral) thought never stops, the search for improved descriptions never ends.
For another, “[T]he Murdochian picture can allow substantive rules … to play a guiding
rather than merely heuristic role …”, thereby taking back a further excess encouraged by
(unreconstructed) particularism. (72) Last and most important, the Murdochian picture
“allow[s] a particularist to sidestep the objection that we cannot make sense of reasons
without allowing for their generality.” (72)

This last point is most important because, while defeasibility theorists have their
own way of avoiding Millgram’s final problem, Murdoch’s way offers particularists the
advantage of deliberative immediacy. ‘Deliberative immediacy’ refers to Murdoch’s view that:

[W]ith the right description in place, your practical reasoning is done: “true vision occasions right conduct” (66/353). You are to arm yourself with descriptions that in an actual choice situation will have direct practical import; in her characterization, “the agent... will be saying ‘This is A B C D’ (normative-descriptive words), and action will follow naturally”. (71-72)

From this aspect of Murdoch’s picture, Millgram argues:

[I]f accepting a description \( D \) shows the action \( a \) to be appropriate, without further intervening deliberative steps, then we have a rule that takes one from \( D \) to \( a \). And this ... allow[s] a particularist to sidestep the objection that we cannot make sense of reasons without allowing for their generality. For on the Murdochian picture, whenever \( D \) is the appropriate description, \( a \) properly follows. But (and this is why a particularist can allow himself that last claim) none of the work is being done by such rules, because one cannot determine whether \( D \) is the appropriate description of a given situation simply by ticking off the features mentioned in \( D \). (72)

Put another way, the descriptions that we arrive at via Murdochian practical reasoning and that generate reasons to act have the form of general rules. By adopting this picture, particularists need not face the problem of how to make sense of non-general reasons. At the same time, particularists need not worry about conceding the generality of their reasons because this concession is merely formal. The general rules that particularists concede by following Murdoch do no work. If these rules did any work, then – as Millgram suggests – the appropriateness of a description \( D \) could be determined “simply by ticking off the features mentioned in \( D \).” But Murdoch rejects precisely this way of thinking, the way of thinking according to which each person’s task of finding an appropriate description for her situation is so straightforward that she only needs to ‘tick

\[37\] Neither Millgram nor I believe the idea that “true vision occasions right conduct” can be accepted without qualification. And, to be fair to Murdoch, it is not clear that she intended this idea to be taken without qualification. So taken, it would represent a form of motivational internalism so strong as to be (almost?) a caricature.
off” the relevant features of her situation to arrive at such a description. Rather, it is often no easy matter for someone to describe her situation in such a way that not only can she fully and sincerely accept the description but also the description she gives does her situation justice. Indeed, on Murdoch’s view, it is here – in arriving at descriptions that meet these dual constraints – where the work of practical reasoning gets done.38

We have now reached the point where the advantage for particularists of looking to Murdoch (as opposed to the philosophy of science) can be seen. For once someone arrives at a description of her situation that satisfies Murdoch’s constraints, the action called for by this description will typically follow.39 No further deliberation is required. This deliberative immediacy proves attractive since it captures a quality that our practical reasoning often does, in fact, exhibit. By contrast, deliberation as conceived by defeasibility particularism lacks this feature. Defeasible generalizations drawn from the philosophy of science cannot capture this immediacy. Consider again the abusive boyfriend case. While defeasibility particularism constitutes an improvement over unreconstructed particularism which would not let us conclude anything without direct, first-hand knowledge of the case, defeasibility particularism only allows us to conclude that the kind of thing the (abusive) boyfriend does – i.e. holding a young child’s head in the toilet and threatening to drown the child if it fails to stop wetting its pants – generally makes an action bad or wrong. To actually reach a firm conclusion in a particular case, coming to realize that the abusive boyfriend description fits the situation is not enough. For defeasibility theorists – unlike Murdoch – our deliberative work is not yet completed.

38 Murdoch often emphasizes how the “big, fat ego” gets in the way of “true vision.” See especially the last two essays in Murdoch, The Sovereignty of Good.
39 “Typically” (or ‘naturally’) as when something follows in the natural course of events.
On their view, we still need to assure ourselves that the embedding context is normal, that no valence-neutralizers or valence-flippers lurk in the surrounding circumstances. Millgram’s advice to particularists, then, appears to be good advice. Both the Murdochian picture that Millgram recommends and the defeasibility alternative take back excesses of unreconstructed particularism. However, of these two alternatives, only the Murdochian view enables particularists to curb excesses while also retaining a picture of deliberation that reflects its usual immediacy.

Still, Millgram’s suggestion cannot be endorsed without reservation. He reads Murdoch’s picture as involving two related, further claims, both of which raise problems. The first claim made a brief appearance earlier. It holds that “the process of substituting better descriptions never ends; there is always more work to be done … coming closer to seeing things as they really are.” (72) The second holds that moral reasoning, as it progresses, leads “away from the shared public world and into private vision and, eventually, into mutual unintelligibility.” (73) Once we appreciate the problems that these two claims raise, it should be clear why Millgram’s recommendation – at least as it stands – does not offer particularists a viable theoretical home.

Taking the two claims and the problems they raise in reverse order, the problem with the second claim -- especially the progression towards “mutual unintelligibility” that it calls for – should be clear. Expanding upon the second claim, Millgram writes, “[A] successful deliberator will sooner or later reach the point where others simply cannot understand his reasons (or, what is the same thing, his characterizations of choice situations).” (73) But this makes nonsense of the idea that what such a “deliberator” has gotten hold of are reasons. Millgram does grant that, “[T]here are stronger and weaker
ways of reading [the second] claim.” (73) Accordingly, he goes on to suggest a reading of Murdoch according to which ‘mutual unintelligibility’ should be understood as mere ‘idiosyncrasy’. Thus softened, the second claim should be unobjectionable.40

The first claim is another story. As we saw, it holds that “the process of substituting better descriptions never ends.” Call this aspect of Murdoch’s view ‘deliberative interminability’. But deliberative interminability undermines precisely the feature of Murdoch’s view that makes it distinctively advantageous for particularists, namely, deliberative immediacy. Recall what this latter feature involves. For Murdoch, while it is not easy for someone to describe her own situation in such a way that both she can fully and sincerely accept the description and the description does justice to her situation, arriving at such a description means that the action it calls for should follow without any further deliberation. That is, getting the right description of the practical problem at hand puts an end to the question of what one should do. But here’s the rub. If the search for improved descriptions never ends, then deliberative immediacy – the very aspect of Murdoch’s view that I have been trumpeting – proves illusory. For if the search for improved descriptions never ends – if we can never be assured that we have got the right description – then the question of what to do never gets resolved and no action, only indecision, ever follows. Deliberative interminability trumps deliberative immediacy.

Granted, in contrast to my emphasis, Milgram does not trumpet the deliberative immediacy of Murdoch’s view as its distinctive advantage. Rather, he takes a more holistic approach, showing “how the different pieces of the [Murdochian] view fall into place when one is looking at the other pieces.” (71) While this general tack makes a

40 Or, perhaps I should say: unobjectionable according to the criterial version of particularism that I defend later in this Chapter and in Chapters Four and Five.
great deal of sense in approaching Murdoch, Millgram nevertheless does not appreciate the tension between some pieces of her view. Specifically, he neglects the tension between deliberative interminability and deliberative immediacy. While Millgram rightly thinks Murdoch’s view fruitful for particularists – indeed, he deserves commendation for bringing Murdoch to our attention again – the tension in Murdoch’s view needs to be resolved before her picture can give particularists the kind of help they need. To resolve the tension in Murdoch’s view, however, we need another way to conceive particularism, one inspired by Wittgenstein’s notion of a criterion.41

6. To briefly sum up the argument to this point, we have good reasons to be wary of unreconstructed particularism. We have also considered two alternatives, but both “have issues.” Defeasibility particularism unnecessarily complicates moral thought and deliberation, giving a false picture of these. The Murdochian alternative, while attractive in large part because it promises to do a better job on precisely this score, does not live up to its substantial promise. As related by Millgram, different aspects of this otherwise attractive picture – deliberative interminability and deliberative immediacy – conflict. Still, if a way could be found to make the parts of Murdoch’s picture cohere, it would be a boon to particularists. I mention this because the alternative that I propose not only has attractions of its own but also resolves the tensions in Murdoch’s picture. Thus, my alternative and Millgram’s Murdochian alternative, suitably modified, should ultimately be seen as not only compatible but complementary.

I propose that particularism be conceived of in terms of criteria. This alternative to unreconstructed particularism – call it ‘criterial particularism’ – does not reject moral

41 Since Wittgenstein arguably had a significant, albeit diffuse, influence upon Murdoch, it should not surprise that the latter can be helpfully read in light of the former.
principles outright but demotes them to a secondary role. Specifically, moral principles cede pride of place to particular cases when it comes to understanding morality and moral concepts. Michael Tanner helpfully puts the central idea like this:

In both morals and art it is possible to encounter phenomena which provide one with one’s concept of greatness or goodness, and which serve … as touchstones for judging works of art, or people, or ways of life (among other things). Once the touchstone has been encountered, a principle can no doubt be elicited from it. But the principle is not only validated by the touchstone, but may well be given its sense by it.

In other words, rather than learn the meaning of ‘right’, ‘wrong’, ‘kind’, ‘cruel’, and other moral terms by learning moral principles and the conditions in which they apply, criterial particularism holds that we grasp the meaning of moral terms by being exposed to “touchstones” of right, wrong, compassion, cruelty, injustice, etc. On this view, touchstones – or exemplary cases – constitute the criteria by exposure to which we come to learn the meaning of moral terms. The criterial element in ‘criterial particularism’ derives from the thought, à la Wittgenstein, that criteria reveal meaning (and are not just symptomatic of some independent phenomena); the particularistic element from the claim that moral criteria, the keys to moral meaning, are cases. Most fundamentally, then, criterial particularism is a view about the meanings of moral terms and how we come to understand them.

42 In this regard, the criterial alternative that I am proposing is similar to the Murdochian picture that Millgram recommends. Both allow general rules or principles.
43 Tanner, "Ethics and aesthetics are -- ?" at 33 (emphasis in original). Right after this passage, Tanner adds:

[J]udgements of aesthetic, and in some cases of moral, value must be based on first-hand experience of their objects not simply because one is no position to assert the presence of the requisite properties without the experience, but also because one is not capable of understanding the meaning of the terms which designate the properties without the experience.” Ibid.

As I suggested in the Introduction, the idea of a link between certain kinds of judgment and certain kinds of experiential response suggests the plausibility of the similar anti-Humean claim I defend in Chapter One. I say more about this in Chapter Five.
know them. In contrast, unreconstructed particularism and the defeasibility alternative both focus on the metaphysics of moral reasons and, with their different stands on these issues, lead to different epistemological consequences. Millgram’s Murdochian alternative differs in that it takes moral deliberation as its point of departure, not moral semantics.

Understanding criterial particularism and its distinctive semantic-cum-epistemological perspective requires that we address at least two questions, two questions that anyone vaguely familiar with Wittgenstein’s notion of a criterion will likely have. First, Wittgenstein most often writes about criteria for someone’s being in this or that psychological state, and most commentators have focused on this use of criteria. Accordingly, I should say something about my proposed extension of Wittgenstein’s notion of a criterion to moral contexts and what basis exists for talking of moral criteria. Most commentators also regard Wittgenstein as having identified discrete considerations (e.g. grimaces, cries) as criteria whereas criterial particularism takes cases to be criteria. Thus, the second question we need to address: what explains this difference in approach. Of course, anytime draws upon Wittgenstein and such pregnant ideas of his as his notion of a criterion, many questions will arise. But addressing the preceding two questions should suffice for our purposes.44

44 Ultimately, a third question may need to be addressed. It arises from the fact that Wittgenstein’s notion of a criterion was once thought to resolve skeptical disputes in favor of, say, knowledge about other minds. But this early view of anti-skeptical view of criteria has been strongly criticized, some (e.g. Stanley Cavell) even arguing that criteria should not be thought of as reasons at all. The third question, then, is this: can criterial particularism give these developments their due seeing how it not only takes criteria to be reasons (in some sense) but also regards them as offering a response (in some sense) to skepticism? However, since taking up this question would be too much of a digression, I save this task for elsewhere.
To address these questions, I survey the debate over Wittgenstein’s notion of a criterion and consider how criterial particularism relates to it. I turn first, then, to a paper by John McDowell, a paper that has been strangely neglected by particularists. (My reasons for thinking this neglect strange, and surprising, will become clear.) In this chapter, McDowell describes the philosophical predicament against which Wittgenstein’s notion of a criterion has typically been seen and interpreted:

The predicament is as follows. Judgements about other minds are, as a class, epistemologically problematic. Judgements about ‘behaviour’ and ‘bodily’ characteristics are, as a class, not epistemologically problematic.... The challenge is to explain how our unproblematic intake of 'behavioural' and 'bodily' information can adequately warrant our problematic judgements about other minds. As Wittgenstein’s notion of a criterion is widely regarded as some sort of response to this predicament, discussion of criteria take place in its shadow. Indeed, the central place that this predicament plays in thinking about criteria justifies extending Wittgenstein’s use of criteria from the context of our knowledge of other minds to the context of our knowledge of moral claims. For while Wittgenstein does on occasion talk of criteria for phenomena besides other minds (e.g. angina) and, thereby, suggests a broader relevance for his notion of a criterion, the best way to appreciate its broader relevance is to realize that the predicament animating it is also a generic one. As Judith Jarvis Thomson has observed:

45 To be clear, exegesis is not my goal in consulting the debate over Wittgenstein’s notion a criterion. That is, I am not concerned whether the notion of a criterion that I take away from Wittgenstein and the related philosophical literature represents the best of the competing interpretations on offer. Instead, apart from any question of fealty to Wittgenstein’s own views, I simply mine his work and the related literature for the notion of a criterion that supports a more attractive form of particularism.

[Scepticism about other minds] is only a special case of something quite general … only one variety of what might be called local scepticism; other familiar varieties are scepticism about the physical world, scepticism about the past, scepticism about the future, [moral scepticism], scepticism about causality, and so on and on.  

Thomson elaborates:

[T]he arguments for each of these varieties of local skepticism all have the same structure….

First draw your hearers’ attention to the crack in the universe of discourse between an interesting kind of sentence, the Ks, and a contrasting kind, the Cs. The Ks may be moral sentences; for them the Cs are factual sentences. The Ks may instead be sentences about the physical world; for them, the Cs are perceptual reports. The Ks may instead be sentences in the past tense; for them, the Cs are sentences in the present tense. And so on.

[Next] invite your hearers to accept the following Premise (i*): the crack in the universe of discourse between Ks and Cs is very deep, that is, no Ks are entailed by Cs. Then invite your hearers to accept the following premise (ii*): Ks are epistemologically parasitic on Cs, that is, the only way we have of finding out that a K is true is by finding out that certain Cs are true, and drawing the K as conclusion from them. Last, invite your hearers to accept the constraint on finding out that is expressed in Premise (iii) [below] … Scepticism now follows for each K-and-C pair.

In other words, Thomson would have us accept the following argument:

Premise (i*): No K sentences are entailed by C sentences.

Premise (ii*): The only way we have of finding out that a K sentence is true is by finding out that certain C sentences are true, and drawing the K sentence as a conclusion from them.

Premise (iii*): If we come to believe a sentence S by finding out that certain other sentences SS are true and drawing S as a conclusion from them, then our coming to believe S constitutes our finding out that S is true only if the SSs entail S.

Conclusion: It is not possible to find out about any K sentence that it is true.

Of course, the predicament with which we started and which animates Wittgenstein’s notion of a criterion fits this pattern, judgments about other minds comprising the K

---


48 Ibid at 70-71.
sentences and judgments about behavior and bodily characteristics the C sentences. And, as Thomson makes clear, the same holds if the Ks are moral sentences and the Cs are factual sentences. Certainly, then, if we can talk of criteria for claims about other minds, we can talk of criteria for moral claims since both forms of skepticism arise out of the same general predicament. This should answer the first question.

To address the second, we need to delve further into the interpretative debate. Taking the predicament already introduced as background, an early interpretation of Wittgenstein construed his notion of a criterion as a way of bridging the divide between epistemologically unproblematic C judgments (e.g. sentences about behavioral and bodily phenomena) and epistemologically problematic K judgments (e.g. sentences about other minds), enabling us to move from the former to the latter and thereby defeat the relevant form of skepticism. On this reading, criteria of \( p \) are easily graspable conditions (e.g. ‘___ is crying’) that would, if satisfied, necessarily justify \( p \) (e.g. ‘Bob is in pain’).

However, faced with the possibility of pretence as well as other phenomena, this reading of criteria cannot be sustained. Someone who pretends to be in pain by grimacing and crying can satisfy what are, on this view, criteria for being in pain without actually being in pain. Thus, we come to a fork in the interpretative road. Either this reading gets Wittgenstein’s notion of a criterion right but, on this view, criteria cannot do what we would like them to do (i.e. provide indefeasible, skeptic-proof reasons) or this reading misunderstands Wittgenstein’s notion of a criterion.

---

49 Moreover, if Wittgenstein’s notion of a criterion offers a response to skepticism about other minds, it should provide the basis for a similar sort of response to skepticism about moral claims. What sort of response I hope to write about elsewhere.
According to which of these two paths we take, we arrive at very different readings of Wittgenstein’s notion of a criterion. The former path accepts both the idea that criteria are certain conditions (e.g. ‘___ is crying’) taken independently of particular circumstances and what follows from this, namely, that criteria of \( p \), so conceived, cannot yield indefeasible support for \( p \). Or, put another way, criteria so conceived do not enable us to get out of the predicament we face and leap from C sentences to K sentences in a single bound, contrary to the expectations that some once had for Wittgensteinian criteria. Still, those who take the former path allow that criteria of \( p \) provide at least some support for \( p \). Specifically, on this view, criterial support is defeasible, defeasible by the particular circumstances in which a criterial condition is satisfied. Moreover, to the extent that there is an orthodox reading of Wittgenstein’s notion of a criterion, it is this: criteria are conditions viewed apart from their circumstances of satisfaction and, when satisfied, they provide only defeasible support.

By contrast, the latter path – the path not (so much) taken – rejects the idea that criteria must be conditions viewed independently of particular circumstances and, thus, regards the orthodox reading of Wittgenstein’s notion of a criterion as a misreading. Instead, following the latter path, whether some condition’s satisfaction (e.g. ‘Bob is crying’) constitutes a criterion of \( p \) (e.g. ‘Bob feels pain’) depends upon whether the particular circumstances in which the condition is satisfied are favorable. Only if the circumstances are favorable – i.e. do not involve pretence or other “defeaters” – does the condition’s satisfaction constitute a criterion of \( p \). Significantly, this alternative reading

\[50\] Using the definite article here is misleading since there are other interpretations, beside the one I discuss in this paragraph, that are consistent with rejecting the orthodox reading. For example, some interpretations eschew the idea that criteria are reasons at all.
has the consequence that criteria do indeed provide indefeasible support. Whereas, on
the orthodox view, a certain condition’s satisfaction always counts as a criterion while the
support it yields varies depending on the circumstances, the alternative reading makes
criterial status and criterial support fellow travelers. That is, not only criterial support but
criterial status itself fluctuates according to the circumstances. If circumstances are
favorable, then a certain condition’s satisfaction constitutes both a criterion of \( p \) and,
correspondingly, an indefeasible reason in support of \( p \). On the other hand, if
circumstances are not favorable, then the very same condition’s satisfaction constitutes
neither a criterion of \( p \) nor, correspondingly, any reason in support of \( p \). Thus, criterial
status and criterial support stand or fall together. Understood in this way, criteria provide
indefeasible reasons.

Criterial particularism, then, finds some inspiration in this last interpretation of
Wittgenstein’s notion of a criterion. Perhaps not surprisingly, John McDowell has been
this interpretation’s chief proponent. Moreover, McDowell’s claim that the support (or
reasons) constituted by criteria need not be defeasible prefigures much of the debate
between unreconstructed and defeasibility particularists.\(^{51}\) At one point, McDowell
identifies the key assumption behind the opposing view of criterial reasons as defeasible.
He writes:

\(^{51}\) This is why I claimed that McDowell’s essay – originally published as McDowell,
"Criteria, Defeasibility, and Knowledge," – has been strangely neglected by
particularists. While both McNaughton and Dancy describe general debts to McDowell’s
work in the opening remarks of their books on particularism – Moral Vision (1988) and
Moral Reasons (1993), respectively – neither book cites this essay of McDowell’s. Nor,
to my knowledge, do these authors cite this essay of McDowell’s in their other work on
particularism. Finally, to my knowledge, defeasibility particularists have not cited
McDowell on this issue either.
This argument [in favor of defeasibility] clearly rests on [an] assumption about the generality of criterial status: that if some condition (specified in a non-question-begging way) is a criterion for a claim in some circumstances, then it is a criterion in any.\textsuperscript{52}

But this assumption – the generality assumption, seen earlier in this chapter – is precisely the one for which Dancy criticizes Ross and whose rejection leads to unreconstructed particularism, at least in the hands of Dancy and McNaughton. Consider next an example McDowell offers to illustrate his position. He explains:

In a schematic picture of a face, it may be the curve of the mouth that makes it right to say the face is cheerful. In another picture the mouth may be represented by a perfect replica of the line that represents the mouth in the first picture, although the face is not cheerful. Do we need a relation of defeasible support in order to accommodate this possibility? Surely not. What is in question is the relation of ‘making it right to say’; it holds in the first case and not in the second. Since the relation does not hold in the second case, it cannot be understood in terms of entailment. But why suppose the only alternative is defeasible support?\textsuperscript{53}

This example is strikingly similar to McNaughton’s aesthetic analogy cited by Little and discussed earlier in the explication of unreconstructed particularism. These examples from both McDowell and McNaughton illustrate a familiar notion associated with particularism: that the normative (or reason-giving) force, if any, of a consideration, whether it be the curve of a mouth or a splash of red, is determined by the context in which it is embedded.

This familiar notion from both McDowell and unreconstructed particularism explains, in part, why criterial particularism understands criteria as cases rather than discrete considerations. Particular circumstances – ‘cases’ in one sense of the word – determine whether the curve of a mouth or a splash of red constitutes a criterion, not the consideration’s presence alone. Still, this only partially answers our second question

\textsuperscript{52} McDowell, "Criteria, Defeasibility, and Knowledge" at 379.
\textsuperscript{53} Ibid at 378.
because criterial particularism also holds that some criteria are cases in a second, more familiar sense. On their view, ‘cases’ in the sense of rich descriptions (e.g. the abusive boyfriend example) can also possess criterial status and the indefeasible support that this status involves. Thus, for criterial particularism, two sorts of ‘case’ provide criteria for moral claims. First, given favorable circumstances, we may know that someone is malevolent simply by the tone of her voice. Just imagine Dickens’ Miss Havisham and the tone of her voice. Given the right circumstances, the tone of someone’s voice can yield a criterial case of malevolence in the special sense of a discrete consideration embedded in a favorable context. We may also know that someone is malevolent simply from how she is described (assuming that the description is both accurate and sufficiently rich in detail). Here again, think of Miss Havisham. Dickens’ description of her yields a criterial case of malevolence in the ordinary sense in which a case is a rich description.

7. We have now answered two key questions that arise about criterial particularism and its appropriation of Wittgensteinian criteria. However, to recognize as criteria ‘cases’ in the colloquial sense of rich descriptions – and not just ‘cases’ in the special sense of discrete considerations embedded in favorable contexts – significantly distinguishes criterial particularism. Given the importance of this difference in assessing criterial particularism vis-à-vis other versions of particularism, it merits a closer look.

54 I say ‘more familiar’ because when we think of a case, we typically mean a rich description, neither an actual particular nor a completely determinate description (if such descriptions are even possible).

55 Of course, the tone of Miss Havisham’s voice comprises an important element in Dickens’ description of her as malevolent. But this only reinforces the basic idea that something discrete (e.g. the tone of someone’s voice) can yield a criterion in two ways: (i) via direct experience given the right circumstances or (ii) as part of a sufficiently rich description. The two need not be mutually exclusive, except in cases where the item at issue is fictional!
So what about recognizing ‘cases’ in the colloquial sense of rich descriptions significantly distinguishes criterial particularism? For one, without this more inclusive understanding of ‘cases’ and the corresponding broadening of what can count as criteria for moral claims, criterial particularism could be regarded as doing little more than reinterpreting unreconstructed particularism in different terms and thereby emphasizing different (i.e. semantic-cum-epistemological) aspects of essentially the same basic view. Not that such a project would lack value. However, in taking (the satisfaction of) certain rich descriptions to qualify as (the satisfaction of) criteria, criterial particularists go beyond anything to which unreconstructed particularism would be committed after being reinterpreted in terms of criteria. Most importantly, this expanded understanding of cases and criteria represents a significant improvement over all of the other versions of particularism that we have considered. For surely we can know certain moral claims to be true simply from knowing that certain rich descriptions are satisfied. Just recall the abusive boyfriend example: knowing that someone satisfies the description (or condition) at issue – holding a young child’s head in a toilet and threatening to drown the child if it fails to stop wetting its pants – is enough to know that what he does is cruel and unjust. Contra unreconstructed particularism, we need not wait on the circumstances of this description’s satisfaction. Regardless of the particular circumstances, this description’s satisfaction means cruelty and injustice are afoot. Indeed, failure to see cruelty and

---

56 Arguably, the main proponents of unreconstructed particularism – Dancy and McNaughton – have not emphasized, let alone developed, the semantic-cum-epistemological aspects of their view as much as they could have. For example, when confronted with seemingly unattractive epistemological consequences of their view, Dancy appears willing to abandon these parts of the view instead of exploring what could be done to make them more palatable. (I base this last claim on a conversation I had with Dancy about an early draft of Dancy, *Ethics Without Principles*. )
Injustice in this description’s satisfaction betokens a failure to grasp what it means for something to be ‘cruel’ and ‘unjust’.  

But can criterial particularists really accommodate ‘cases’ like the abusive boyfriend example together with what they imply? Criterial particularism takes its cue from McDowell’s understanding of criteria, and – as I was keen to point out earlier – close parallels exist between McDowell’s understanding of criteria and unreconstructed particularism. But this suggests a negative answer to my question. To see how an affirmative answer is possible, we need to reconsider McDowell’s understanding of criteria and especially how they are made or unmade by particular circumstances.

In arguing that criteria need not provide merely defeasible support, McDowell discusses a line of reasoning that I call “the argument from context-sensitivity.” While everyone can agree that a criterion’s reason-giving force – i.e. criterial support – depends upon context, the key question is how particular circumstances enter into the equation. “The argument from context-sensitivity” promises some illumination. This argument is supposed to show that criterial support is defeasible. McDowell glosses it as follows:

The idea that criterial knowledge depends on circumstances is obviously faithful to Wittgenstein; but this argument rests on an interpretation of that idea that is not obviously correct. [The] assumption is evidently this: if a condition is ever a criterion for a claim, by virtue of belonging to some type of condition that can be ascertained to obtain independently of establishing the claim, then any condition of that type constitutes a criterion for that claim, or one suitably related to it. Given that such a condition obtains, further circumstances determine whether the support it affords the claim is solid; if the further circumstances are unfavourable,

---

57 On this point at least, criterial particularism is similar to the views of Philippa Foot and G.E.M. Anscombe, not to mention Murdoch. These similarities become more apparent in Chapters Four and Five.
we still have, according to this view, a case of a criterion’s being satisfied, but the support it affords the claim is defeated.\textsuperscript{58}

Spelled out more fully, the argument goes like this:

P1: If a feature is a criterion of some claim $p$, then this feature possesses normative (or reason-giving) force vis-à-vis $p$.

P2: A feature that possesses normative (or reason-giving) force vis-à-vis $p$ does so in virtue of belonging to a type of feature that can be ascertained to obtain independently of establishing $p$.

C1: If a feature is a criterion of $p$, then this feature possesses normative (or reason-giving) force vis-à-vis $p$ in virtue of belonging to a type of feature that can be ascertained to obtain independently of establishing $p$. (P1, P2)

P3: If a feature possesses normative (or reason-giving) force vis-à-vis $p$ in virtue of belonging to a type of feature that can be ascertained to obtain independently of establishing $p$, then any feature that belongs to this type constitutes a criterion for $p$.

C2: If a feature is a criterion for $p$, then any feature that belongs to this type constitutes a criterion for $p$. (C1, P3)

P4: If a feature that belongs to a certain type constitutes a criterion of $p$, then the normative (or reason-giving) force that a criterion of $p$ possesses vis-à-vis $p$ also depends upon the broader circumstances in which the criterion is embedded.

C3: If a feature is a criterion for $p$, then the normative (or reason-giving) force that a criterion of $p$ possesses vis-à-vis $p$ also depends upon the broader circumstances in which the criterion is embedded. In other words, the normative (or reason-giving) force that a criterion of $p$ carries vis-à-vis $p$ depends upon the broader circumstances in which the criterion is embedded. (C2, P4)

P5: If the broader circumstances in which a criterion of $p$ is embedded are sometimes unfavorable, the normative (or reason-giving) force that a criterion of $p$ possesses vis-à-vis $p$ is sometimes undermined.

P6: The broader circumstances in which a criterion of $p$ is embedded are sometimes unfavorable.

\textsuperscript{58} McDowell finds this argument in favor of the defeasibility of criterial support in a paper of Gordon Baker’s entitled “Criteria: A New Foundation for Semantics.” See McDowell, "Criteria, Defeasibility, and Knowledge" at 376-77, including n.15.
C4: The normative (or reason-giving) force that a criterion of \( p \) carries vis-à-vis \( p \) is sometimes undermined. In other words, the normative (or reason-giving) force possessed by criteria is defeasible.

Perhaps the key premise here is P2: the now-familiar generality assumption rephrased. Criterial particularists reject this premise along with McDowell and unreconstructed particularists. Doing so allows for “stubbornly particular reasons.” As we have seen, if McDowell is correct, Wittgensteinian criteria are stubbornly particular reasons, that is, features whose normative force depends upon particular circumstances, and not upon their mere presence alone.\(^{59}\) P2 must be rejected because it dictates that a feature’s normative (or reason-giving) force depends upon something independent of particular circumstances (i.e. belonging to a certain type of feature), thereby excluding stubbornly particular reasons.

Now we consulted “the argument from context-sensitivity” to help determine whether criterial particularism could plausibly accommodate criteria that do not depend upon particular circumstances for their criterial status as in the abusive boyfriend case. Joining with unreconstructed particularism to reject P2 might seem to confirm our pessimism on this score. But such pessimism is a mistake. Rejecting P2 only means that criteria (or features with normative force) need not derive their criterial status (or normative force) from their generality, thus opening the door for some criteria to be – à la McDowell’s reading of Wittgenstein – stubbornly particular reasons. Entirely consistent with a denial of P2 is the recognition that other features (e.g. the behavior of the abusive boyfriend) are criteria because of the type of feature to which they belong and, thus, regardless of the particular circumstances.

\(^{59}\) Compare defeasibility particularists here. On their view, a feature’s presence taken by itself does possess normative (or reason-giving) force, albeit defeasible normative force.
Parallels with unreconstructed particularism suggest otherwise only because unreconstructed particularists do not merely reject P2. In addition, they also endorse something like P2’s contrary, to wit: any feature that possesses normative force vis-à-vis some claim p does so in virtue of the particular circumstances in which it is embedded. Thus, for unreconstructed particularists, criterial status and normative force vis-à-vis p cannot be a matter of belonging to a type of feature that can be ascertained independently of p. On this key point, there is no parallel. Criterial particularists merely reject P2; they do not endorse its contrary. Accordingly, they can accommodate the lessons of the abusive boyfriend example while unreconstructed particularists cannot.

Interestingly, defeasibility particularists fare no better in accommodating the lessons of the abusive boyfriend example – even though I originally introduced this example to back their epistemological complaint against unreconstructed particularism. While this example does indeed support their epistemological complaint, it proves too much, at least for them. For when someone holds a young child’s head in the toilet and threatens to drown the child if it fails to stop wetting its pants, this does not tell us that his actions are defeasibly cruel and unjust but that his actions are cruel and unjust, period. We need not examine the particular circumstances to see whether his actions have their moral significance neutralized or inverted. Some things that people do are simply cruel and unjust – indeed, simply wrong – and nothing further that we could learn about the circumstances of their actions can change these facts. What the abusive boyfriend does is this kind of thing.

This difference between defeasibility particularism and criterial particularism is also reflected in their different takes on “the argument from context-sensitivity.” While
the former account of particularism holds that the normative force possessed by reason-giving features is defeasible at best and, accordingly, accepts the premises of this argument, the latter rejects not only P2 – as we have seen – but also P4. The rejection of P4 highlights the difference over the abusive boyfriend example because P4 expresses the kind of holism that prevents defeasibility particularism from accommodating this example’s lessons. That is, defeasibility particularists believe that when a feature is the type of feature that constitutes a criterion for some moral claim (e.g. the type of behavior exhibited by the abusive boyfriend), the normative force of this feature also depends upon the broader circumstances in the sense that it is subject to being neutralized or reversed by these circumstances. But since criterial particularists are unencumbered by P4 and this kind of holism, they can recognize the full force of the abusive boyfriend example, namely, that the moral-normative force of some features – e.g. holding a young child’s head in the toilet and threatening to drown the child if it fails to stop wetting its pants – does not depend upon the particular circumstances.

In significant contrast, then, with what we might call ‘orthodox particularism’ (i.e. particularism in its unreconstructed and defeasibility varieties), criterial particularism does not accept holism in the theory of reasons in any general or principled sense. Recall Dancy’s claim to the effect that:

[P]articularists need to decide … [whether] they want to say that a property brings nothing to an action, its contribution being determined only when it is present with others, or whether they only want to say that though there is a default position, a contribution which the property makes other things being equal, that contribution can be reversed or annulled by untoward circumstances.

---

60 Criterial particularists also reject the next (and final) two premises, P5 and P6, as they presuppose P4.
Whereas unreconstructed particularists are holists in the first sense and defeasibility particularists are holists in the second sense, criterial particularists deny that Dancy’s alternatives exhaust their options. If they are holists, they are only *ad hoc* holists. Thus, criterial particularists agree with unreconstructed particularists that some features possess normative force only in virtue of particular circumstances, but they do not generalize this holism. Instead, criterial particularists recognize that some features possess normative force independently of particular circumstances, and no formula exists for saying which features depend on particular circumstances for their normative force and which do not.\(^6\)

Moreover, in contrast with defeasibility particularism, features that possess normative force independently of particular circumstances do so indefeasibly. Such features are not just *defeasibly* morally relevant but morally relevant, period. Crucially, these differences weigh in favor of criterial particularism. As the abusive boyfriend example shows, we want to allow not only that some features possess normative force specifically because of the type to which they belong – and, thus, independently of particular circumstances – but also that such features possess normative force indefeasibly. Criterial particularism can accommodate these possibilities whereas orthodox particularism cannot. That it can do this constitutes an important argument in its favor vis-à-vis orthodox particularism.

**8.** With respect to a particularism informed by Murdoch – at least as Millgram understands these – criterial particularism enjoys the same advantage. For as Millgram reads both particularism and Murdoch, the supposedly key aspect of the former – the defusing move – can be captured by the latter’s idea that moral thought and deliberation

\(^6\) That being said, a very rough distinction perhaps could be drawn between (i) more-or-less discrete considerations that constitute criteria only in favorable contexts and (ii) richly described cases that constitute criteria independently of any particular context.
amounts to the search for better and better descriptions of one’s situation. But, as I have just argued, it is precisely the defusing move – given the insistence upon its universal availability – that makes orthodox particularism untenable in light of cases like the abusive boyfriend. Granted, the defusing move figures centrally in many expositions of particularism. Still, Millgram errs by continuing to insist on its universal availability. This continued insistence not only makes his Murdochian particularism vulnerable to cases like those of the abusive boyfriend, but also leads him to distort Murdoch’s views. For, unlike Millgram, I doubt Murdoch would have much sympathy for the idea that the defusing move should always be available, at least not in the sense that orthodox particularists are committed to its being available. On the other hand, this incompatibility between orthodox particularism and Murdoch does not mean that the latter has nothing to offer particularism. Murdoch has a great deal to offer. On this general point, Millgram is quite right. However, seriously engaging her views would affect particularism in ways that run rather deeper than Millgram envisions. Here I shall only argue that reading Murdoch in light of criterial particularism – as opposed to orthodox particularism with its insistence on the defusing move’s universal availability – makes better sense of her views without undermining their relevance. Moreover, the benefits run both ways. Murdoch’s views, properly understood, helpfully inform criterial particularism’s picture of moral thought and reasoning, a topic to be taken up later. Back now to the topic at hand.

---

Interestingly, the defusing move does not figure in any of McDowell’s work that is actually cited by orthodox particularists. While McDowell suggests the availability of the defusing move in his essay on criteria – an essay that Dancy and McNaughton arguably could cite as an influence – even here McDowell does not insist upon the move’s uniform availability. In a further irony – given how Millgram would have us read Murdoch – it is McDowell who acknowledges a debt to Murdoch while orthodox particularists do not.
As already intimated, the first reason to read Murdoch as a criterial particularist arises from cases fit for compelling descriptions such as ‘holding a young child’s head in the toilet and threatening to drown the child if it fails to stop wetting its pants’. *Inter alia,* what such cases compel is our use of certain moral terms (e.g. ‘cruel’, ‘abusive’, and ‘unjust’ not to mention ‘wrong-making’ and flat-out ‘wrong’). Moreover, we do not take the implausible view that conduct satisfying such compelling descriptions is merely *provisionally* cruel or *defeasibly* wrong-making, but rather hold that this conduct is cruel and wrong-making, full stop. But reading into Murdoch – *per* orthodox particularism – the defusing move’s universal availability would saddle her with precisely the former, implausible view. For in Murdoch’s scheme, the defusing move’s universal availability would mean that otherwise fitting moral terms (e.g. ‘cruel’ or ‘wrong-making’) could – given the right extenuating circumstances – be displaced from our descriptions of such cases. Read in terms of criterial particularism, however, Murdoch would be free of this implausibility. Instead, in cases like this, Murdoch would be taken to hold that otherwise fitting moral terms could not be displaced – no matter what the extenuating circumstances – on pain of failing to understand the moral terms involved. Therefore, our first reason to read Murdoch in terms of criterial particularism – and not in terms of orthodox particularism à la Millgram – is to avoid saddling her with the implausible view that conduct such as the abusive boyfriend’s is merely *provisionally* cruel or *defeasibly* wrong-making.

However, the benefits of reading Murdoch in this way extend further. Earlier I identified a tension between deliberative interminability and deliberative immediacy. While this tension in Murdoch’s position looms anyway, reading her in terms of criterial
particularism arguably resolves it. By contrast, reading her in terms of orthodox particularism – with its insistence upon the defusing move’s universal availability – offers no help and, worse, makes the problem appear irreconcilable. Here’s the problem. On one hand, the search for how best to describe one’s situation supposedly never ends. Since better descriptions can always be found, one presumably should not be satisfied with whatever description one may have arrived at. On the other hand, once one has found the right description of one’s situation, this finding should lead immediately (i.e. without any further deliberation) to an action called for by this description. But if one can never be assured that one has found the right description of one’s situation (as deliberative interminability seems to imply), then the advertised connection between deliberation and action – deliberative immediacy, that is – would never seem to “kick in” (or be engaged). Thus, two main virtues of Murdoch’s view seem to be at odds.

But here criterial particularism can help. Remember it holds that, for various moral terms such as ‘cruel’, ‘unjust’, and so on, there are certain descriptions that constitute criteria for these terms and these criterial descriptions (or cases) partially determine the meaning of these terms. Read with this in mind, then, Murdoch need not follow orthodox particularists in holding that any time we might be called upon to describe a situation as, for example, ‘cruel’ or ‘wrong-making’, such descriptions can still be overthrown given the right extenuating circumstances. Rather, in light of criterial particularism, Murdoch can be understood as holding that, for some situations (e.g. the abusive boyfriend case), the right description must make use of certain terms (e.g. ‘cruel’ or ‘wrong-making’) on pain of evincing our failure to understand these terms. Accordingly, Murdoch would hold that, at least for some situations and the terms these
situations require in any description thereof, the search for improved descriptions does come to an end. Deliberative interminability need not be understood as interminability with respect to every possible situation. Situations that involve moral criteria constitute a limit on moral deliberation, a limit on our search for better and better descriptions. Furthermore, when moral deliberation is limited in this way (i.e. when moral criteria are present), deliberative immediacy comes into play. That is to say, when criterial situations require that we use certain terms to describe them and thereby limit moral deliberation as Murdoch conceives it, the advertised connection between deliberation and action will typically show itself, something it cannot do if we can never be properly satisfied with our descriptions. Thus, reading Murdoch in this way resolves the conflict between deliberative interminability and deliberative immediacy. Read in light of criterial particularism, her view can accommodate both insights.

Lastly, Murdoch’s view benefits from being read in light of criterial particularism because the consequent appeal to moral criteria helps both to make sense of her thought that “the movement of [moral] understanding is onward into increasing privacy” and to justify the weaker, more palatable stance regarding what this key thought involves (i.e. idiosyncrasy as opposed to mutual unintelligibility). As I observed earlier at the end of section five, the view that moral thought and reasoning involve a progression into mutual unintelligibility should, if possible, be avoided. And, while Millgram acknowledges that Murdoch’s thought can be interpreted in both a weaker sense (‘mere idiosyncrasy’) and a stronger sense (‘mutual unintelligibility’), he offers little basis for settling on the weaker sense apart from the undesirability of the stronger. Perhaps this is all we need. But
reading Murdoch in terms of criterial particularism provides a clear, independent basis for understanding her view in the weaker, more palatable way.

Notice first that the appeal to criteria by which criterial particularism would have us conceive of the meaning of moral terms fits nicely with Murdoch’s thought that moral understanding moves “onward into increasing privacy.” More than this, criteria frame Murdoch’s thought itself in an illuminating way. The key point here: moral criteria only partially determine the meaning of moral terms; they do not provide univocal guidance for every possible use of such a term. Rather, beyond cases that actually involve moral criteria and require that certain moral terms be used (on pain of evincing a failure to understand said terms), moral criteria leave (logical) space for different users to apply moral terms in innovative and, yes, idiosyncratic ways. Indeed, just as Murdoch would have it, the real work of moral thinking gets done – and moral thinking shows progress – not in criterial cases but in cases beyond the criterial, cases that require more than just the minimal grasp of moral terms reflected in our shared reactions and responses to criteria. The idea here is this. Given that most of us share the basic grasp of moral concepts reflected in our shared reactions and responses to criterial cases, we show progress in our moral thinking when we can find something true and interesting to say about the moral character of a situation that is not prototypical of any moral concept. That is, we show progress in our moral thinking when we recognize the possibility of extending moral concepts to situations that are not obviously covered by such concepts. Take the mundane situation of young women doing menial tasks in an office or laboratory somewhere:
These girls sing, like girls sing … all over the world …. They talk among themselves: they talk about the rationing, about their fiancés, about their homes, about the approaching holidays …

‘Are you going home on Sunday? I am not, traveling is so uncomfortable!’

‘I am going home for Christmas. Only two weeks and then it will be Christmas again; it hardly seems real, this year has gone by so quickly!’

Nothing morally remarkable here, or so it might seem. In any event, situations like this involve nothing so striking as moral criteria. But consider what Primo Levi makes of this ubiquitous scene. Seeing in it a dramatic contrast to his own situation, he writes (and the ellipses are his):

… This year has gone by so quickly. This time last year I was a free man: an outlaw but free, I had a name and a family, I had an eager and restless mind, an agile and healthy body. I used to think of many, far-away things: of my work, of the end of the war, of good and evil, of the nature of things and of the laws which govern human actions; also of the mountains, of singing and loving, of music, of poetry. I had an enormous, deep-rooted, foolish faith in the benevolence of fate; to kill and to die seemed extraneous literary things to me. My days were both cheerful and sad, but I regretted them equally, they were all full and positive; the future stood before me as a great treasure. Today the only thing left of the life of those days is what one needs to suffer hunger and cold; I am not even alive enough to know how to kill myself.

In these two passages, Levi demonstrates how moral thinking can develop by finding moral significance in situations that appear to hold no such significance. By juxtaposing the ubiquitous scene described in the first passage with the thoughts he shares with us in the second passage, the reader realizes how insensitivity, cruelty, and worse can lurk in the quotidian. Taken together, these passages reflect the sort of interesting extension of moral thought envisioned by criterial particularism and called for by Murdoch’s idea of

---

63 Primo Levi, *Survival in Auschwitz (originally: If This Is a Man)* (1996) at 143.
64 Ibid at 143-44.
moral thought as a progression “onward into increasing privacy.” The increasing privacy – or idiosyncrasy – comes from what Levi brings to the situation he is describing, who he is and what he has experienced, especially his forced transformation over the past year from a chemistry student in Italy to Haftling 174517 at Auschwitz. While Levi’s experience is, by definition, idiosyncratic and his reflections upon it intensely personal, these facts do not make his reflections unintelligible to us. Indeed, Levi arguably shows how progress in moral thought can, at least conceivably, be transmitted from one moral consciousness to another. For reading the memoir from which these passages come makes it difficult, if not impossible, for the reader to look at the quotidian in quite the same way ever again. Of course, this kind of innovative, idiosyncratic moral thinking abounds in literature, something that surely helps explain Murdoch’s attraction to it. Most importantly for our purposes, thinking of moral meaning as criterial particularism would have us do – i.e. in terms of criteria – helps elucidate why it makes sense for Murdoch to identify progress in moral thinking with a kind of linguistic innovation often found in literature (as well as elsewhere).

9. Having discussed the benefits of viewing Murdoch through the lens of criterial particularism as I propose – rather than through the lens of orthodox particularism and the defusing move as Millgram proposes – we can now take up the other side of the bargain

65 Interestingly, Levi does not emphasize any specific moral terms in these passages and, as a result, the moral terms for which these passages are supposed to suggest novel or idiosyncratic uses are not entirely clear. ‘Insensitivity’ and ‘cruelty’ are two possibilities. We should note, then, that the greater force of what Levi has to say here goes beyond uncovering novel deployments of moral terms such as these and relates to his efforts to articulate the dehumanization he experienced in Auschwitz. That moral thought need not involve the use of specifically moral terms has been defended by Cora Diamond, "Wittgenstein, mathematics, and ethics: Resisting the attractions of realism" in The Cambridge Companion to Wittgenstein, ed. Sluga and Stern (1996), 226-260 and, more recently, in Crary, Beyond Moral Judgment.
referred to above, that is, the benefits of having Murdoch’s views inform (criterial) particularism. These benefits, as I see them, are primarily two-fold. First, Murdoch’s approach to moral thought and deliberation shows (criterial) particularism the way to a plausible, attractive account of moral reasoning. Second, Murdoch’s approach, especially its feature that I have tagged ‘deliberative immediacy’, shows (criterial) particularism the way to making good on its claims to anti-Humeanism.

Few would dispute that particularism requires a more plausible, attractive account of moral reasoning. Its emphasis on holism and context-dependence led Millgram to worry that “appeals to context-dependence … usually function philosophically as conversation-stoppers …” Relatively, others have worried that particularism’s emphasis on moral vision amounts to a revival of the dogmatic, “take it or leave it” approach to moral justification associated with intuitionism. Therefore, pace Maggie Little, invoking moral vision along with art appreciation analogies does create for some “a real stumbling block to accepting particularism.” Now, because Murdoch herself gives a significant role to moral vision, one might wonder how drawing upon her view enables particularism to move beyond the unattractive picture of moral reasoning suggested by such talk. However, as Millgram generally appreciates even if he ultimately fails to get her right, Murdoch’s view involves a number of other, important elements that affect how her talk of moral vision must be understood.

To elucidate why invoking moral vision – whether by Murdoch or others – does not entail a dogmatic picture of moral justification and, correspondingly, how looking to her view can helpfully inform (criterial) particularism, we can begin with a recent paper

---

67 Little, "Moral Generalities Revisited" at 293.
by Richard Norman.\(^{68}\) Norman begins with a scenario in which a university administrator asks for the names of supposedly “dead wood” colleagues (i.e. non-publishers who are otherwise hard-working and dedicated) with an eye towards forcing them out. Of such a scenario, he declares, “I know that loyalty to my colleagues requires me to refuse. To supply the names would be an act of betrayal.” (121) Norman then observes that:

> [T]his is … the kind of thing some … have in mind when they talk about ‘moral vision’. To see that the action of naming names would be an act of betrayal, and to see this as the overridingly important fact of the situation, is what [has been called] the ‘perception of salience’ here. (Ibid.)

Certainly, from what we know of Murdoch, this appears to be what she has in mind by ‘moral vision’. For the ongoing search regarding how best to describe one’s situation that, on her view, constitutes moral deliberation can also be characterized as trying to see one’s situation in the right way (e.g. in terms of loyalty and betrayal). So Norman is on the right track. But he also points out something too often overlooked, namely, that:

> [W]e are not talking about literal ‘vision’ here. A blind person would be as capable as anyone else of seeing that my naming names would be an act of betrayal. ‘Vision’ is here a metaphor, but it is a metaphor which comes naturally in our description of the situation, and it seems … to be an appropriate metaphor. (121-122)

Of course, if Murdoch and others’ talk of ‘moral vision’ is metaphorical, this raises other questions. For one, why do visual metaphors arise so naturally when we talk of moral recognition (i.e. the recognition that certain moral concepts are in play)? More importantly, do visual metaphors really help us say something illuminating about moral recognition?

Taking up these questions, Norman provides some insight. He writes:

\(^{68}\) Norman, "Making Sense of Moral Realism," . Page numbers here refer to this paper.
To what kind of seeing are we comparing the process of moral recognition? It is not like seeing in the sense of the mere passive reception of data. What provides us with our metaphor is the kind of seeing which involves the active interpretation of our visual experience – seeing not as data reception but as pattern recognition: seeing that the building in the distance is a church, seeing that the flower is an Early Purple Orchid, seeing that, in this painting, the movements of the angel and the Virgin Mary perfectly balance one another. It is a matter of picking out the relevant features, and of recognising the formal relationships between them. It is to this kind of seeing that we can compare the recognition of the action as an act of betrayal; it is seeing a gestalt.

Moreover, echoing something that Millgram emphasized about Murdoch, Norman adds:

[I]n both the literal and the metaphorical cases of seeing, the act of recognition is non-inferential. We can’t just list the appropriate features, tick them off, and infer that the flower is an Early Purple Orchid. It is a matter of seeing that the flowers have just this particular shape, are just this particular shade of purple, that the leaves are spotted in just this particular pattern. The relation between the recognition and the more primitive data is one of assembling the data into a total picture. Likewise in the case of moral recognition. (122) (emphasis in original)

In other words, visual pattern recognition and moral recognition share at least two similarities: the gestalt nature of the recognition and the immediacy with which this recognition dawns upon us. Using visual metaphors to characterize moral thinking highlights these two similarities.

Visual metaphors also bring out a third similarity between moral recognition and visual pattern recognition. Just like visual pattern recognition, Norman reminds us that:

Moral recognition is not a matter of applying a universal principle to the particular case. We do not adopt the principle ‘Don’t be disloyal’ or ‘Don't betray your colleagues’ and then apply it to the particular case to generate the moral conclusion. What is crucial for moral awareness is not the adopting of the

---

69 Contrary to what some might think, the immediacy of our recognizing that, e.g., naming names would be an act of betrayal does not make dogmatism a fait accompli. For simply because a moral conclusion like this can dawn upon us immediately (i.e. without our needing to rely upon a traditional argument with premises to reach it) does not mean that this process of moral recognition (or dawning moral awareness) depends upon no reasoning whatsoever. Reasoning is involved, but – as Norman and others suggest – it manifests itself at a different point and in a different way. I discuss this at length in Chapter Five.
principle, but the recognising that [something] *is* an act of betrayal. Once you have recognised this, there does not seem to be any work left for the universal principle to do. (Ibid.)

But even granting the priority of the particular as suggested here by Norman – that is, even granting that we recognize moral generality *in* particular cases, not via principles – moral principles can still be formulated to cover particular cases of moral recognition, even principles with much greater specificity than ‘Don’t be disloyal’ or ‘Don’t betray your colleagues’. However, such a task risks missing the point that moral principles often only recapitulate conclusions that we have already reached through the hard work of figuring out what to make of – or how best to describe – the situations we face. Put another way, being aware of moral principles and the fact that they apply prospectively does not get one very far. For neither moral principles nor any other kind of principle apply themselves. Unlike navigational aids that tell you in a soothing voice when to turn in order to reach a certain destination, moral principles do not announce when they apply (except in cartoons). Rather, each individual must recognize, for him or herself, when various moral concepts are in play (and, accordingly, when various moral principles apply), and this recognition is not always easy or obvious. It is certainly not automatic.  

With this last point about the difficulty of recognizing what moral concepts are in play (and, correspondingly, what moral principles apply), we get closer to understanding why talk of moral vision does not commit Murdoch or anyone else to a dogmatic picture

---

70This last point – that moral recognition and, correspondingly, how best to describe a situation is often neither easy nor obvious – is a crucial corollary of Murdoch’s view. In her paper “Missing the Adventure: Reply to Martha Nussbaum,” Cora Diamond begins with a striking illustration of this point and how ignoring it can lead us seriously astray. William Frankena’s gross misreading of how Socrates argues against escape in the *Crito* provides the illustration. Diamond’s paper is reprinted in Cora Diamond, *The Realistic Spirit: Wittgenstein, Philosophy and the Mind* (1991) at 309-17.
of moral justification. Still, such talk does highlight the particularistic, non-inferential and gestalt aspects of moral recognition, and this can make dogmatism seem unavoidable. But even granting that moral recognition possesses these features, getting someone to “see” (or better, recognize) something morally important about what he proposes to do – e.g. that naming names counts as a betrayal – need not be dogmatic. For, as I discuss in the second part of Chapter Five, we (can) justify the claim involved – e.g. that naming names counts as betrayal – in the process of getting someone to “see” (or recognize) it. Norman is correct, then, to hold that visual metaphors need not lead to the dogmatism associated with intuitionism. Such metaphors actually elucidate moral thought. Properly understood, visual metaphors highlight a picture of moral thought and deliberation very much at one with Murdoch’s view, a view according to which the real work involves trying to figure out how best to describe (or ‘see’) the particular situations in which we find ourselves.
In Chapter Three, I surveyed several varieties of particularism and concluded that one stood out as superior. Specifically, I argued for so-called ‘criterial particularism’ over not only unreconstructed particularism but also defeasibility particularism. Apart from its ability to make sense of the truth in anti-Humaneanism – no small matter – criterial particularism enjoys the significant advantage (or so I claimed) of accommodating as criteria not just certain key features when embedded in the appropriate (“no defeaters lurking”) circumstances but also certain more-or-less richly described scenarios regardless of the embedding circumstances. I take this to be a significant advantage because, while both defeasibility particularism and criterial particularism can account for the moral relevance of stabbing as opposed to shoelace color and, correspondingly, for our ability to make at least tentative moral evaluations of situations without actually confronting them first-hand, only criterial particularism makes room for the immediacy of our knowledge that, in some situations richly described, what is going on is cruel and counts unqualifiedly against it. Or, as I have maintained, certain cases richly described (e.g. the abusive boyfriend) support claims of noninferential a priori moral knowledge if anything does. But only criterial particularism can accommodate this upshot of such compelling cases.
However, what I am trumpeting as a significant advantage may raise some eyebrows. For criterial particularism’s “significant advantage” might be thought to undermine not only its legitimacy as a variety of particularism but also its overall plausibility regardless of whether it deserves to be called a form of particularism. To see the problem here, I need to elaborate on two points regarding what it means (for criterial particularism) to accommodate richly described scenarios as criteria. First, as my discussion of the abusive boyfriend example from the previous chapter suggests, criterial particularism reflects the thought that we need not (and should not) hedge our bets when it comes to such cases. What the boyfriend is described as doing in this case simply is cruel and a reason against doing it, full stop. There is no cause to inject any qualifiers here (e.g. ‘normally’, ‘everything else being equal’, etc.). Thus, criterial particularism not only recognizes the existence of moral principles (in contrast to unreconstructed particularism) but also recognizes the existence of unhedged moral principles (in contrast to defeasibility particularism).

While this first point might seem to amount to nothing more than a turf battle between different sorts of particularists, the second point cannot be discounted in this way. For the second point relates to what it means for rich descriptions to be criteria in the relevant sense. As we just saw, it means that criterial particularism is committed to unhedged moral principles. However, these unhedged moral principles are also supposed to be criteria in Wittgenstein’s sense and criteria understood in this sense help give the meaning of the term(s) at issue. Therefore, the unhedged moral principles to which criterial particularism is committed need to be understood as analytic and, correspondingly, as something we must accept if we are to be (conceptually) competent
with the term(s) at issue. This appears to be an example of what Austin called “taking it back” with a vengeance. Indeed, criterial particularism appears to be a form of “constitutive generalism,” a view that has been thought to be the polar opposite of particularism.\footnote{McKeever and Ridge, \textit{Principled Ethics: Generalism as a Regulative Ideal}. Unless otherwise noted, parenthetical page numbers in this chapter refer to this book.} (12) More importantly, apart from whether it still counts as a form of particularism, criterial particularism’s commitment to analytic moral principles raises serious doubts about its overall plausibility in light of Moore’s famous open question argument.

2. To help sort out some of the issues here, consider particularism as glossed by Sean McKeever and Michael Ridge, the authors of an important, recent monograph on the subject.\footnote{See previous footnote.} Consistent with my own classificatory scheme laid out at the beginning of Chapter Three, McKeever and Ridge suggest that “Different forms of particularism are defined by the different negative claims they make about moral principles.” (14) They then proceed to list five different negative claims about moral principles and, correspondingly, the five different versions of particularism that flow from each of these negative claims, starting with the strongest (or most ambitious) version. They are:

1. There are no true moral principles.  
   (Principle Eliminativism Particularism)

2. There is no good reason to think there are any true moral principles.  
   (Principle Skepticism Particularism)

3. No finite set of moral principles is sufficient to capture all moral truths.  
   (Principled Particularism)

4. We ought not rely upon moral principles.  
   (Principle Abstinence Particularism)
5. The possibility of moral thought and judgment do not depend on the provision of a suitable supply of moral principles. (Anti-Transcendental Particularism) (pp. 15-20)

Of course, McKeever and Ridge realize that we need to know what it is for something to be a moral principle if we are to understand claims about them. Towards this end, they helpfully discuss six different conceptions of moral principles (given that “[o]n any plausible account, not just any generalization, employing a moral predicate counts as a moral principle” (5)). According to these six different conceptions, moral principles serve the following purposes, respectively:

(i) to state sufficient conditions for the application of some moral concept (Principles *qua* Standards),

(ii) to guide the actions of conscientious moral agents (Principles *qua* Action Guides),

(iii) to do both (i) and (ii) (Principles *qua* Action-Guiding Standards),

(iv) to provide an algorithm for moral decision-making (Principles *qua* Algorithmic Decision-Procedure),

(v) to reveal preconditions of conceptual competence with moral terms (Principles *qua* Preconditions of Conceptual Competence),

and

(vi) to serve as truth-makers for particular moral truths (Principles *qua* Truth-Makers). (pp. 6-14)

For McKeever and Ridge, then, one counts as a particularist to the extent that, according to one or more conceptions of moral principles (see i-vi), one accepts some negative claim(s) (see 1-5) about the role of moral principles in moral thought and practice. By contrast, one counts as a generalist to the extent that, according to one or another of these conceptions of moral principles, one grants such principles some role in moral thought and practice (and thereby rejects one of particularism’s negative claims).
So, we are all particularists now. Or, at least most of us are if we accept McKeever and Ridge’s general framework. For if moral principles are understood as algorithmic decision-procedures, then most moral philosophers, including by their own admission McKeever and Ridge, count as particularists even in the strongest, Principle Eliminativism sense. That is, most moral philosophers recognize the need to employ judgment when applying moral principles to the facts of particular cases and the corresponding impossibility of a mechanical decision-procedure in this area. Likewise, however, we are all generalists too. For again, if moral principles are taken to be mere action guides – useful rules of thumb for doing the right thing - it is hard to imagine anyone denying their existence. But this only shows that the main debate between particularism and generalism hinges on moral principles otherwise conceived, that is, not moral principles *qua* algorithms and not moral principles *qua* rules of thumb.

Specifically, McKeever and Ridge hold that particularism is best understood as making negative claims about moral principles *qua* standards, or, in the alternative, moral principles *qua* action-guiding standards (to the extent that moral principles *qua* mere standards do exist, etc.).

But where does this leave criterial particularism, the view being explored and defended here? If we accept McKeever and Ridge’s gloss of particularism – that is, as

---

3 For instance: Scheffler, *Human Morality* at 43.
4 Still, given their framework, McKeever and Ridge would understand the difference between particularism and generalism as a matter of degree. Thus, perhaps understanding particularism as I propose in the Introduction – as being fundamentally about a commitment to the uncodifiability of moral terms – has the further advantage of allowing the line between particularism and generalism to be (more) sharply drawn. I regard this as an advantage because many have thought particularism to be a very distinctive if not radical view, and my framework for understanding particularism better captures this common thought.
one (or more) of five negative claims about moral principles *qua* standards – then what I have been calling criterial particularism should not count as a genuine form of particularism. This was the first problem I mentioned earlier, and the McKeever-Ridge framework brings it out clearly. For so-called criterial particularism – however it should be classified – is committed to both the existence of principles *qua* standards and a fundamental role for these principles in moral thought. Moreover, the latter half of this commitment distances criterial particularism even further from genuine particularism on McKeever and Ridge’s view. For, as I alluded to earlier, criterial particularism’s commitment to a fundamental role for moral principles *qua* standards amounts to a commitment to these same moral principles as preconditions for conceptual competence with moral terms. And the idea not only that true moral principles exist but that our understanding and accepting them constitutes our competence to deploy moral concepts is what McKeever and Ridge dub “constitutive generalism,” a view that McKeever and Ridge consider to lie on the “outer bound of generalism” (12) and that – for reasons discussed below – they reject.

3. Given its endorsement of unhedged moral principles constitutive of competence with moral concepts, criterial particularism faces another, arguably more significant problem than simply defending its claim to the mantle of particularism. For, as mentioned earlier, this second problem threatens to undermine criterial particularism’s plausibility regardless of whether we accept its claim to particularism’s mantle. The

---

5 As it turns out, my so-called ‘criterial particularism’ has a good claim to the name of particularism for reasons I discuss in the Introduction. Also see the previous note.

6 Interestingly, as we see below, McKeever and Ridge acknowledge that at least one other “card-carrying particularist” (i.e. Margaret Little) appears committed to constitutive generalism. (107-8)
source of the problem here is not simply that criterial particularism endorses unhedged moral principles but that these same principles are supposed to reflect what (conceptual) competence with moral language requires. But – and this is the problem - Moore’s famous “open question argument” arguably shows that no moral principles are analytically true (i.e. true simply by virtue of the concepts involved).

Consider my recasting of Moore’s argument. It concludes that questions of the form ‘x exhibits non-moral features N, but does it exhibit moral feature M?’ are conceptually open in that such questions need not signify any failure on the questioner’s part to understand the moral term at issue. Of course, Moore does not deny that some questions can signify a lack of understanding on the questioner’s part, e.g., “I know that the umpire called that pitch strike three, but why is the batter out?” However, on Moore’s view, moral concepts are different. We are not supposed to be able to provide standards (i.e. sufficient conditions for applying moral concepts) in non-moral terms.

McKeever and Ridge are impressed by Moore’s argument, and take it to pose a serious obstacle for any view such as criterial particularism which holds that some moral principles are analytically true and, thereby, constitutive of competence with moral terms. To avoid the charge that Moore’s argument begs the question against naturalism – the view that moral terms can indeed by defined in natural (or non-moral) terms – McKeever and Ridge read Moore’s argument charitably, treating the conceptual openness of moral terms vis-à-vis non-moral descriptions as a live issue and as therefore, as moving from the apparent conceptual openness of moral terms vis-à-vis non-moral descriptions to the

---

desired conclusion via an inference to the best explanation, in other words, the actual conceptual openness of the moral vis-à-vis non-moral best explains its appearing so.

Defenders of analytic moral principles would seem to have only two options for avoiding the force of Moore’s open question argument. About the first option, McKeever and Ridge write:

The constitutive generalist could try to deny the putative data, and hold that ordinary speakers do not find the relevant questions to be open. But this seems desperate. The relevant questions simply do seem open to competent speakers across a very wide range of proffered analyses of moral terms. (100)

Moreover, while McKeever and Ridge grant that “[I]t remains in principle possible that ... a constitutive generalist will offer an analysis that does not leave competent speakers thinking the relevant questions are open,” they greatly discount this possibility given that “disagreement on moral question is so pervasive and deep.” (100)

The only other apparent way to avoid Moore’s argument: give a better explanation of the fact that moral terms appear conceptually open vis-à-vis nonmoral descriptions than the fact that they actually are open in this sense. To illustrate the kind of story that defenders of analytic moral principles need to tell, McKeever and Ridge draw an analogy with linguistics. In trying to arrive at the correct standards of application for the predicate ‘is grammatical’ in some natural language, McKeever and Ridge tell us that:

The linguist’s job is to gather data as to how competent speakers classify sentences as grammatical or ungrammatical, controlling for performance-based errors [due to, e.g., various possible cognitive impairments], and then articulate a set of . . . rules that would best explain those classifications. If the job is done correctly then there is a useful sense in which the linguist’s rules provide a
 Crucially, however, the results of the linguist’s investigations – rules that in some sense define ‘is grammatical’ – are likely to be “so unintuitive and complex that competent speakers will not recognize their correctness when presented with them.” (101) And this fact about the linguistic concept of the grammatical – that it admits of a correct, reductive definition but competent users of the concept may not recognize this definition’s correctness – should help make the point of the linguistics analogy clear. For the appearance that ‘is grammatical’ is conceptually open is best explained not by its actually being conceptually open – *ex hypothesi*, there is a correct, reductive definition and so it is not, in fact, conceptually open – *but* by the unintuitiveness and complexity of the correct definition. Thus, defenders of analytic moral principles could take the same general approach. They could argue that various moral terms are not conceptually open vis-à-vis non-moral properties since appearances to the contrary are best explained by the unintuitiveness and complexity of the actual definitions of such terms. Not surprisingly, this is the path chosen by the primary defenders of analytic moral principles that McKeever and Ridge consider. However, it is not the path that criterial particularism takes.

From a particularist perspective including, especially, the criterial view being defended here, this defense of analytic moral principles is deeply unappealing. For, as their linguistics analogy strongly suggests, McKeever and Ridge apparently consider any

---

8 One wonders why McKeever and Ridge felt the need to qualify the upshot of the linguist’s investigations with the phrase ‘there is a useful sense in which’. That is, one wonders about the implied sense in which the linguist’s investigations do not provide a definition of ‘is grammatical . . . ‘ (101)
endorsement of analytic moral principles – at least if such principles are unhedged – as tantamount to a claim that morality can be codified, that the meaning of moral terms can be (reductively) defined in non-moral terms. What is so unappealing is not simply the tacit assumption that definition must be the concomitant of endorsing analytic moral principles but that, in assuming this, the alternative response to Moore’s argument that criterial particularism pursues gets overlooked or, worse, obscured from view altogether. However, as criterial particularism crucially maintains, endorsing a whole variety of different analytic moral principles that state many different non-moral conditions sufficient for the application of some moral term such as ‘cruel’ does not commit us to the view that cruelty can be defined in non-moral terms, not even in principle. If this is correct, then some new logical space would indeed open up: not only for allowing an alternative response to Moore but also (and relatedly) for making sense of a new form of naturalism about ethics. And this is just what criterial particularism amounts to: a new form of (non-reductive) naturalism in ethics.

Before trying to define the novel claim that we have just been discussing, consider first how it enables criterial particularism to respond differently to Moore’s open question argument. By changing the basic Moorean line of questioning about the conceptual openness of moral terms, criterial particularism’s novel claim makes it quite plausible – and far from “desperate” – to challenge the idea that the intuitive responses of ordinary speakers support this openness. For if criterial particularism is correct, then we should ask ordinary, competent speakers only whether they regard questions such as:

‘Bothered by the incessant crying of his girlfriend’s toddler, Kenny grabs the little boy and throws him against
the wall with all his might, causing massive internal injuries and killing the child. But is this cruel?\(^9\)

as conceptually open. That is, if criterial particularism is correct, we need to ask different questions to gauge conceptual openness in the relevant sense. Our questions should not aim to test the adequacy of some analytic definition of, say, ‘cruelty’ but instead to merely uncover whether ‘cruel’ is conceptually open vis-à-vis this or that non-moral description (viz. my description of Kenny’s action).

Now it seems fair to say that most ordinary speakers would not regard the question ‘But is it cruel?’ as conceptually open when asked of an action that satisfied the non-moral description I gave of Kenny’s action. Nor is this non-moral description unique in partially fixing the meaning of ‘cruel’. Unfortunately, other paradigmatic examples (or criteria) of cruelty abound and could be used to generate an indenumerable variety of non-moral descriptions that each help to partially fix what we mean by ‘cruel’. (By calling such other examples ‘paradigmatic’ or ‘criterial’, I just mean that competent speakers would not regard our use of ‘cruel’ as conceptually open vis-à-vis the non-moral descriptions of such examples.) The real question here about such examples and their descriptions is not what to call them but whether they exist. That is, whether cases exist that, non-morally described, conceptually foreclose any question about a moral term’s applicability. Since criterial particularism rejects as inapposite efforts to capture (or codify) the meaning of moral terms in a complete reductive analysis, it does not try to redescribe cases more and more abstractly to fit some preconceived definition. Rather, owing to its forbearance from trying to (reductively) define moral concepts, criterial

\(^9\) We might also ask ‘But does all this count against Kenny’s doing it?’
particularism allows us to take cases as we find them in all their rich specificity.\(^{10}\) And, given this rich specificity, ordinary speakers will indeed regard questions about the applicability of certain moral terms as conceptually foreclosed. Or, returning to my earlier example, ordinary speakers will consider anyone who questions the cruelty of Kenny’s action as lacking (sufficient) understanding of what it means for something to be cruel. That is, if you do not think that what Kenny does is cruel, then you do not understand cruelty.

4. Even if – for the reasons just discussed – Moore’s open question argument poses no problem for criterial particularism and the sort of analytic moral principles it endorses, McKeever and Ridge offer a closely related objection that must also be taken seriously. (107) Specifically, based on a set of alleged counterexamples, they object to any view such as criterial particularism that takes competence with moral concepts to require the acceptance of certain moral principles. Here are McKeever and Ridge addressing an interpretation of defeasibility particularism that commits it to analytic moral principles, albeit hedged ones:\(^{11}\)

Someone can be competent with moral concepts without embracing any of the defeasible generalizations [Little] discusses. A nihilist can deny that pain or lying (or anything else) provides defeasible reasons or indeed any kind of reason at all without conceptual confusion. To take another example, a hedonist would deny

\(^{10}\) Of course, given what we learned from Murdoch in Chapter Three, “… tak[ing] cases as we find them …” can be problematic. On occasion, we may realize that we have no descriptions, whether richly specific or not, that fit well the situation in which we find ourselves. Other times we should at least be open to the possibility that this is the case. For Murdoch holds that the crucial step in moral thought and deliberation is arriving at an appropriate description of our situation.

\(^{11}\) The suggestion that at least some of defeasibility particularism’s hedged moral principles might be analytic appears in a relatively early essay by Margaret Little alone (and not in her later articles co-authored with Mark Lance). Little, "On Knowing the 'Why': Particularism and Moral Theory," *op. cit.* Perhaps Little and Lance would not endorse the “constitutive generalist” reading.
that lying in itself provides a defeasible reason (or any kind of reason at all) but is
not thereby conceptually confused. Certain forms of Kantianism and virtue ethics
deny that pain in itself provides a defeasible reason (or any kind of reason at all).
However, Kantians and virtue ethicists are not thereby guilty of conceptual
confusion. (108)

In other words, defeasibility particularists characteristically endorse moral principles such
as (a) ‘Typically, if an action is an instance of lying, then this fact alone is a reason
against doing it’ and (b) ‘Everything else being equal, if an action involves the infliction
of pain, then this fact alone counts against doing it’. However, even with the qualifiers,
hedonists will reject (a) while certain Kantians and virtue theorists will reject (b). Yet no
one who subscribed to any of these moral views and who, by implication, rejected the
principle impugned by their respective view would thereby count as incompetent with
moral concepts. Thus, at least with respect to the kind of moral principles that
defeasibility particularism endorses, McKeever and Ridge conclude that such principles
cannot be analytic.

The worry here, of course, is that this sort of objection – or really, these sorts of
counterexamples – can be modified to fit the kind of moral principles that criterial
particularism endorses as analytic. And no doubt, with a little ingenuity, these
counterexamples can be amended to threaten criterial particularism. However, the threat
that these counterexamples pose is more apparent than real. Granted, principles that
might seem to be the kind that criterial particularism endorses as analytic (i.e. principles
with the same rich specificity as my earlier examples) can be formulated so that hedonists
et al. will reject them. Consider, the following variations on a theme:

(c) ‘If an action involves grabbing an incessantly coughing
yet fitfully sleeping five-year-old and slamming him
against a wall with all one’s might, thereby causing
massive internal injuries and killing him instantly, then this fact alone is a reason against doing it.’

(d) ‘If an action involves grabbing an incessantly crying infant and throwing him against a wall with all one’s might, thereby causing massive internal injuries and killing him, then this fact alone is a reason against doing it.’

(e) ‘If an action involves grabbing an incessantly coughing yet fitfully sleeping five-year-old and slamming him against a wall with all one’s might, thereby causing massive internal injuries and killing him instantly, then it is cruel.’

(f) ‘If an action involves grabbing an incessantly crying infant and throwing him against a wall with all one’s might, thereby causing massive internal injuries and killing him, then it is cruel.’

Notice that (c) and (d) have been formulated so that hedonists will reject the former and Kantians the latter, yet no one is suggesting – not even criterial particularists – that hedonists and Kantians must therefore be conceptually confused about the applicability of the moral term ‘is a reason against’. But for the non-analyticity of moral principles (c) and (d) to tell against criterial particularism, they must be the sort of principles that criterial particularism takes as analytic. While (c) and (d) possess the requisite specificity, they appeal to this specificity to conceptually ground thin moral terms such as ‘is a reason against’, not thick moral terms such as ‘cruel’ or ‘barbarous’. And even though many ordinary, competent speakers might regard hedonists and Kantians as conceptually confused to the extent that they reject (c) and (d), respectively, criterial particularism is more charitable. For criterial particularism only insists on the analyticity of moral principles involving thick moral terms. In other words, unhedged principles with their richly detailed descriptions only serve as direct conceptual constraints on the applicability of thick moral terms. By contrast, thin moral terms – both the contributory
(e.g. ‘is a reason against’) and the verdictive variety (e.g. ‘is right’) – are not constrained conceptually in the same direct way, thereby freeing hedonists, Kantians, and other moral theorists to pursue unity and simplicity without risk of conceptual confusion (at least on these grounds). Still, criterial particularism holds that our use of thin moral terms can run up against conceptual constraints. However, while the conceptual constraints on thin moral terms also ultimately appeal to the richly detailed, non-moral descriptions that constrain thick moral terms, these descriptions operate to constrain thin moral terms in a complex and indirect way that cannot be captured in a strict (i.e. unhedged) principle. Therefore, criterial particularism is not vulnerable to counterexamples from this quarter either.

So what about the nihilist, the source of McKeever and Ridge’s other counterexample to views that endorse analytic moral principles? At least two kinds of nihilism can be distinguished: we need to know which of these two McKeever and Ridge intend. Nietzsche, for instance, arguably accepts merely rhetorical nihilism. Despite all his protestations to the contrary – to the effect that he is abstaining from any and all evaluative commitments – he nevertheless stands committed to all sorts of values, non-traditional and non-Christian perhaps but evaluative commitments all the same. However, even if he only pretends to genuine nihilism, Nietzsche does give us a picture of this alternative. Genuine nihilism, that is, involves doing what Nietzsche only pretends to do: “standing outside” (or abstaining from) absolutely all evaluative commitments.

Presumably McKeever and Ridge intend their nihilist to be the global sort. For local nihilists have to endorse criterial particularism’s analytic moral principles about
thick moral terms just as hedonists and Kantians would. While local nihilism involves objecting to specific thick terms and the evaluative commitments they involve (e.g. ‘chaste’ or ‘virginal’), she could also abstain from using them and, thereby, stand outside the cognitive-evaluative “package deal” that this or that specific thick concept involves. But even so, she could not deny the various analytic moral principles that, according to criterial particularism, serve to conceptually limit the concept unless she truly did not understand the thick moral concept at issue.

A genuine nihilist, by contrast, might seem better placed to reject criterial particularism’s analytic moral principles without conceptual confusion. And perhaps this is correct. But genuine nihilism presents several other problems. First, it is not even clear what it would mean for someone to abstain from absolutely all evaluative commitments. Second, even if we can make sense of this idea, it is unclear whether so-called genuine nihilism is even possible. Perhaps Mr. Spock of Star Trek fame – the pure intellect – would be an example. But then again, from how his character is portrayed, it is clear that he does not live up to this “ideal”. Finally, even if we can make sense of someone’s being able to abstain from all evaluative commitments and even if this were possible for someone to do, it is far from clear that success in achieving this would have no impact on one’s cognitive status/conceptual competence. To assume that it would not is to assume that the cognitive and the cognitive aspects of our natures can be halved off from each other without altering the other. And to assume this is to assume quite a lot. Thus, because genuine nihilism seems sufficiently problematic and/or controversial, McKeever and Ridge’s appeal to it cannot do much work for their argument.
CHAPTER FIVE

On the Very Idea of Nondefinitive Analytic Naturalism: Rule-Following and an Alternative Account of Moral Reasoning

1. McKeever and Ridge’s important, recent discussion of particularism assumes that the more-or-less wholesale rejection of moral principles is what makes particularism worthy of serious consideration. For the rejection of a substantial role for principles in moral thought supposedly poses the deepest and, therefore, most interesting challenge to orthodox moral philosophy.¹ According to this way of thinking, the view that I have been defending here – so-called criterial particularism – holds relatively little interest since it accepts moral principles that are not only unhedged but analytic as well.

McKeever and Ridge apparently overlook the possibility of such a view because they assume that being committed to unhedged moral principles – let alone unhedged moral principles that are analytic – is tantamount to being committed to the codification of morality. But understanding how criterial particularism can plausibly urge the acceptance of moral principles that are both unhedged and analytic will show not only how, among the many moral particularisms, McKeever and Ridge missed this view, but also, more importantly, why it is criterial particularism that is the most challenging of orthodox moral philosophy and, hence, the most interesting – all despite the fact that it embraces unhedged moral principles.

To see why a commitment to even unhedged, analytic moral principles does not imply a commitment to defining (or codifying) morality, we need to return to Moore’s open question argument and a key assumption that it involves. Moore assumes that if the linguistically competent can sensibly question the analysis of the term given by some proposed definition, then the proposed definition fails. In other words, to succeed, definitions must seem obvious to the competent upon reflection.\(^2\) No doubt, since Moore first put forward his open question argument over a century ago, much has transpired in the philosophy of language to throw doubt upon this key assumption. Developments in the semantics of natural kind terms, for example, have made us comfortable with the idea that at least some terms (e.g. water) are amenable to successful, uncontroversial yet non-obvious definition. Indeed, the idea that moral terms might be understood (and ultimately defined) along broadly similar lines has become the overwhelmingly preferred way of responding to Moore’s famous argument.\(^3\)

Criterial particularism, however, does not take this route. Rather, criterial particularism urges a myriad of unhedged analytic moral principles, yet – precisely because it still accepts some version of Moore’s obviousness constraint – it also rejects any idea that moral terms can be defined. For all moral terms – especially the thin but also the thick – admit genuine hard cases in which even competent users can be of different minds as to whether the term at issue should apply, not simply borderline cases owing to generic kinds of vagueness. Thus, given the hard cases that arise for all moral terms, any proposed definition of a moral term will have to take a stand in such cases,

\(^2\) ‘Obvious’ here certainly does not mean being obvious, or even easy, to formulate. 
\(^3\) At least for those of a broadly realist persuasion. Apart from Moore’s own non-naturalism, one could explain the conceptual openness and the corresponding failures of analytic definition by the allegedly non-cognitive (or expressive) function of moral terms.
leaving it open to sensible questioning from competent users and a failure as a
definition.⁴

To be clear, criterial particularism freely grants that some terms (e.g. ‘water’) are
amenable to non-obvious definition and, thus, concedes that Moore’s key assumption
requires qualification given more recent developments. But criterial particularism does
not concede that moral terms should be understood along the lines of natural kind terms
and, thereby, freed of Moore’s obviousness constraint on adequate definitions. Now I
have no intent to simply argue against the natural kind understanding of moral terms and
in favor of the alternative that Moore and criterial particularism shares (call it the
common sense understanding). For ultimately I think the proof is in the pudding, so to
speak. That is, to best evaluate opposing ways of understanding moral terms, we should
set them out and see what each can do (i.e. in terms of saving the appearances, explaining
anomalies, etc.) With this in mind, consider that the natural kind understanding of moral
terms has been received enthusiastically – at least by those inclined towards realism (or
cognitivism) – because it provides a way around Moore’s open question argument that

⁴ Interestingly, if Moore’s requirement that definitions be obvious to the competent is
retained for moral terms (as criterial particularism would do), a challenge to traditional
moral theory does loom after all. For traditional moral theory attempts to define key
moral terms with the goal of reducing/eliminating moral disagreement and, at the same
time, making moral progress. See, for example, McKeever and Ridge, *Principled Ethics:
Generalism as a Regulative Ideal* at 181-189 and, more modestly, T.M. Scanlon, "The
1-23. However, if genuine hard cases exist and we retain Moore’s stricture on definitions
– as criterial particularism would do for moral terms – then this central goal of traditional
moral theory is a dead-end, doomed to failure. It is far from clear, however, that this
project was ever anything except Quixotic. So, arguably, no great loss. (Of course, I do
not mean to impugn the nobility of this project. Sure, it would be nice if moral
philosophy had the transformative power that McKeever and Ridge imagine it has. My
complaint instead concerns how deeply unrealistic it is. Moreover, criterial
particularism’s gappy view of moral semantics helps explain why it is unrealistic.)
doesn’t involve Moore’s non-naturalism. But criterial particularism seems equally well-positioned in this regard. For criterial particularism also avoids the force of Moore’s open question argument, albeit by refraining from any attempt to define moral terms and endorsing as analytic only those unhedged moral principles that anyone competent with the moral terms at issue would also endorse. So the relative dialectical parity to this point between moral realism of the natural kind variety, and criterial particularism ought to make us wonder why the position in the conceptual “problem space” created by Moore’s famous argument and occupied by criterial particularism has gone almost completely unnoticed. Perhaps it is simply the vogue and influence that the new developments in the semantics of natural kind terms have enjoyed. Surely this is part of the reason. But criterial particularism owes much of its inspiration to Wittgenstein, and his ideas have not been without “vogue and influence” themselves. So, something else must be at work. Uncovering this “something else” will help to not only locate criterial particularism more precisely on the conceptual map but also set the agenda for the rest of this chapter.

2. To bring out the precariousness of the logical space that criterial particularism wants to occupy and, thus, to further explain why this position may have been neglected, consider the following three claims about moral semantics:

---

5 Criterial particularism arguably counts as a form of naturalism owing to the fact that the moral principles it endorses, while not defining the moral terms involved, do constitute substantial, non-moral constraints on our use of these terms.
I. Our use of moral terms is guided by objective features of the world.  
   \textbf{[Cognitivism]}

II. Our use of some moral term $m$ is guided by objective features of the world
    \textit{if and only if}
    a definition (i.e. necessary and sufficient conditions for the application) of $m$
    can be given in terms of objective, non-moral features of the world.  
   \textbf{[Naturalism]}

III. Upon inspection, the correct definition of a moral term must seem correct to
     anyone competent with the term.  
    \textbf{[Analyticity]}

These three claims form an inconsistent triad that helps elucidate the logical space available for responding to Moore’s open question argument. Moore himself, by insisting on Analyticity, maintains a grip on Cognitivism only by denying Naturalism. However, few find Moore’s way of holding onto Cognitivism while rejecting Naturalism – that is, his claim that ‘good’ stands for a simple, non-natural property – very satisfying, let alone convincing. Thus, most respond to Moore’s argument in the other two, more-or-less straightforward ways suggested by this triad. On one hand, we have seen that many would loosen the requirements for defining moral terms a la natural kind terms – that is, reject Moore’s insistence on Analyticity – making it possible to combine Cognitivism and Naturalism without inconsistency. On the other hand, others continue to feel the pull of Analyticity yet regard Moore’s eschewal of Naturalism as a desperate, unmotivated move to save Cognitivism. These philosophers accept the latter two theses as legitimate constraints on any semantics of moral terms, but realize that, taken together, they rule out Cognitivism. Accordingly, these philosophers take moral terms to play noncognitive (or expressive) roles instead.

Where does this leave criterial particularism? At first glance, we may appear to have canvassed all the possibilities without mentioning it. Still, logical space remains.
To clarify, criterial particularism fits into the same broad class as Moore’s own view since they both accept Cognitivism and Analyticity at the price of what I have been calling Naturalism. But, unlike Moore, criterial particularism does not reject Naturalism – or, at least, the motivation behind it – unconditionally. Instead, criterial particularism would distinguish between Defining Naturalism (DN) and Constraining Naturalism (CN).

(DN):
Our use of some moral term $m$ is guided by objective features of the world if and only if a definition (i.e. necessary and sufficient conditions for the application) of $m$ can be given in terms of objective, non-moral features of the world;

(CN):
Our use of some moral term is guided by objective features of the world if and only if (a variety of) sufficient conditions for the application of $m$ can be given in terms of objective, non-moral features of the world;

It then rejects DN but not CN. Of course, DN is the “Naturalism” in my inconsistent triad and so, by rejecting it, criterial particularism does share some ground with Moore. On the other hand, by urging CN instead, criterial particularism distinguishes itself from Moore’s non-naturalism and arguably qualifies as a form of naturalism. Moreover, if we can accept CN (as opposed to DN), we no longer have an inconsistent triad. We can consistently maintain Cognitivism, Naturalism (CN), and Analyticity: precisely what criterial particularism would have us do.

But does a (merely) constraining naturalism even make sense? Perhaps the difficulty in seeing how criterial particularists can remain analytic naturalists about ethics without being committed to the definition of moral terms in natural (or non-moral) language explains why such a position has been underappreciated if not ignored. However, making sense of an analytic naturalism that is constraining but not defining does not appear so hard. For a roughly analogous position has been recognized by
philosophers of law, seemingly without difficulty, as one way of understanding “gaps” in the law and the judicial discretion that such gaps create. As Brian Bix puts the question to which the legal analogue of constraining naturalism is one possible answer:

Are there ‘gaps’ in the law because language, rules, or the law have ‘run out’ entirely, or because the meanings of the relevant materials . . . no longer determine a particular answer or interpretation, even though they do establish a range of permissible (and, by inference, impermissible) answer and interpretations?6

The second of these explanations reflects the legal analogue of constraining naturalism. About the corresponding discretion created by these gaps and how this might be understood, Bix helpfully adds:

[It is] the negative analogue of the ‘one right answer’ approach. That is, judges have discretion if it is not the case that, as a matter of law, there is only one right answer to the question before them. Thus . . . judicial discretion is consistent with there being for a particular case either a range of possible right answers (along with a range of possible wrong answers) or no justified characterization of any answer as ‘right’ or ‘wrong’ before the decision is handed down.7

So, for questions that come before a judge, the legal analogue of constraining naturalism holds that the law does not determine one right answer but still yields a range of acceptable (and unacceptable) answers. Put another way, just like certain objective, non-moral features of the world arguably do in ethics, the materials that constitute “the law” – however broadly (or narrowly) defined – play a substantial, guiding role even when they do not determine a uniquely correct result.

This certainly seems like a picture of rational guidance that makes sense, not only for the law but also for ethics. Nevertheless, in ethics, such a picture – to wit, an analytic naturalism that is constraining but not defining – does not even seem to be an option for

---

7 Ibid at 26.
most philosophers. Indeed, non-naturalism appears to have more takers! Assuming that constraining naturalism makes sense given what we have just seen of its legal cousin, one suspects that something deeper – some fundamental conception (or misconception) – must help obscure the possibility of a constraining naturalism in ethics.

To get clearer about this, recall that the legal analogue of constraining naturalism arises as a response to the problem posed by gaps in ‘what the law requires’. But why would gaps in the coverage of some predicate pose a problem? Of course, as we shall see, the specific problem posed by gaps depends on the kind of predicate at issue. Generally, however, it has been assumed that when a case falls into the gaps in some predicate’s coverage, it is arbitrary – or, at least, arbitrary from the perspective of the rule governing our use of the predicate – whether the predicate applies in such a case. Thus, many have worried about the prospects of gaps in the predicate that judges need apply and the assumed implication of such gaps, namely, the inevitable arbitrariness – at least insofar as what the law requires – of certain judicial decisions (often the more controversial ones). Various ways of addressing this concern about arbitrariness in the law have been mooted, including the one canvassed by Bix that gaps in the law need not have this implication.8 About gaps in moral predicates, by contrast, the worry has a more metaphysical cast. If moral predicates admit gaps, it is thought that our use of them cannot reflect a grasp of the ways things are objectively (i.e. independently of us). In

---

other words, if moral predicates admit gaps, then supposedly they do not express genuine concepts.\(^9\)

But while the prospect of gaps gives rise to different concerns in the case of moral as opposed to legal predicates, the impetus for these concerns is ultimately the same: the threat of arbitrariness. That is, if the rule that governs our use of some moral predicate were such that, in a whole range of cases, we could apply the predicate willy-nilly, few philosophers – not even this one! – would count our use of such a predicate as a response to the way things are objectively. In other words, the deep-seated assumption that helps obscure the possibility of a constraining naturalism lies not in the move from:

(ii) Over a range of cases, the rule governing our use of some predicate \(m\) allows that \(m\) can be applied willy-nilly.


rather, the fundamental misconception comes earlier in the move from the mere fact that:

(i) A predicate \(m\) admits gaps.


to (ii), precisely the move that constraining naturalism and its legal cousin rejects. The question remains as to why the necessity of the move from (i) to (ii) – at least in the domain of moral predicates – is so often (or commonly) taken for granted.

Before taking up this pivotal issue, we should briefly consider an objection to the foregoing analysis. The objection at issue maintains that the very idea that a predicate

\(^9\) Presumably this more metaphysical worry does not arise for ‘what the law requires’ because, even if it admitted no gaps, it would still fail to express a genuine concept given (i) the standard for genuine concepthood in play and (ii) the fact that, by all accounts, the law involves much that is not independent of us in the relevant sense (the past decisions of individual judges, for example).
has gaps entails its arbitrariness, that is, (i) entails (ii). According to this objection, we should accept the following account of what it is for a predicate to admit gaps:

\[(\text{GAPS1}): \text{A predicate } m \text{ admits gaps if and only if for some possible uses of } m, \text{ the rule governing our use of } m \text{ fails to give any guidance whatsoever regarding whether } m \text{ applies.}\]

But such an account begs the question against constraining naturalism and its legal cousin, assuming what they ask us to reject: the move from gappiness to arbitrariness, from (i) to (ii) above. We could instead account for gappiness like this:

\[(\text{GAPS2}): \text{A predicate } m \text{ admits gaps if and only if for all possible uses of } m \text{ (except those involving ordinary vagueness), the rule governing our use of } m \text{ fails to unequivocally determine that } m \text{ either applies or does not apply.}\]

Like Bix’s idea of judicial discretion as the negative analogue of Dworkin’s one right answer theses, (GAPS2) is similarly broad-minded since the crucial notion of a rule’s failing to determine unequivocally whether a predicate applies is consistent with both (a) the rule’s providing no guidance whatsoever and (b) the rule’s providing guidance without determining one right answer. Because (GAPS2), unlike (GAPS1), does not exclude the possibility of constraining naturalism, it clearly should be preferred in this context.

3. We can now turn to the pivotal question of why it is commonly assumed that if a moral predicate admits gaps, it must be arbitrary in its application. We should notice, however, that this pivotal question brings us back to the beginning of our attempts to make sense of constraining naturalism. For if being able – at least in principle – to give necessary and sufficient conditions for a predicate’s application means that the predicate admits no gaps, our pivotal question ultimately amounts to questioning naturalism – or, more precisely, defining naturalism – from the opposite direction. In other words,
exploring why gappiness is commonly taken to entail arbitrariness also means exploring
our initial question of why objective guidance (the opposite of arbitrariness) is commonly
thought to require the provision – at least in principle – of necessary and sufficient
conditions (the opposite of gappiness). Moreover, approaching our pivotal question from
this initial angle leads straight into the most sustained attempt to answer the question.

The most sustained attempt to answer our pivotal question can be found in John
McDowell’s exploration and analysis of resistance to the idea that what virtue requires
cannot be codified, not even in principle. McDowell claims that resistance to virtue’s
uncodifiability flows from “[a] deep-rooted prejudice about rationality” according to
which “acting in the light of a specific conception of rationality must be explicable in
terms of being guided by a formulable universal principle.”

From McDowell’s perspective, then, the crucial question is this: why is it thought that, to be rational and
consistent in whatever one is doing, one’s behavior must be guided by something
expressible as a universal principle? Furthermore, if what one is doing is trying to apply
a moral predicate such as ‘cruel’, then McDowell’s question is just a more general
version of our own.

To answer his (and our) question, McDowell takes up Wittgenstein’s famous,
much debated discussion of rule-following in *Philosophical Investigations*. First and
foremost, McDowell urges that – pace Kripke and others – Wittgenstein’s argument
should not be read “as making a sceptical point.”

Thus, about Wittgenstein’s imaginary
pupil in Section 185 who applies the instructions “Add 2” as we would expect to 1000
but then continues 1004, 1008, 1012 . . ., McDowell elaborates:

---

10 McDowell, “Virtue and Reason” at 58.
11 Ibid at 59.
That is, the point of Wittgenstein’s excursus on rule-following is not skeptical but deflationary. His examples and considerations are supposed to show that our confidence in certain ways of projecting our predicates as correct has its basis in nothing but our shared forms of life. Indeed, Wittgenstein’s point is even stronger: only our shared forms of life could play this grounding role. Crucially, moreover, this modest conception of what grounds our confidence need not – nor, arguably, did Wittgenstein intend it to – raise any doubts about the possibility of our projections being responsive to the way things are objectively. Thus, consistent with a grounding in our shared forms of life, our projections of, say, moral predicates can still be guided by the sort of objective, non-moral features of the world that naturalism, broadly speaking, requires. The only caveat: these objective, non-moral features of the world possess normative force and guide our projections not in some brutally causal way but only through their being taken up and recognized as having such force in our shared forms of life.13

So what is meant by ‘our shared forms of life’? On this key notion, I can do no better than McDowell does and likewise defer to Stanley Cavell. As quoted by McDowell, Cavell writes:

---

12 Ibid at 59-60.
13 To help clarify the role of our shared forms of life in mediating the guiding role of certain objective, non-moral features of the world, we might draw an analogy to Gareth Evans’s work on demonstrative reference, thinking of our shared forms of life as the modes of presentation for the various objective, non-moral features that guide our moral projections. Accordingly, an emphasis on our shared forms of life – no more than Evans’s emphasis on modes of presentation – need not set up a barrier between us and the objective world.
We learn and teach words in certain contexts, and then we are expected, and expect others, to be able to project them into further contexts. Nothing insures that this projection will take place (in particular, not the grasping of universals nor the grasping of books of rules), just as nothing insures that we will make, and understand, the same projections. That on the whole we do is a matter of our sharing routes of interest and feeling, modes of response, senses of humor and of significance and of fulfillment, of what is outrageous, of what is similar to what else, what a rebuke, what forgiveness, of when an utterance is an assertion, when an appeal, when an explanation – all the whirl of organism Wittgenstein calls “forms of life.”

Talk of our shared forms of life, then, refers to “... our sharing routes of interest and feeling, modes of response, senses of humor and of significance and of fulfillment, of what is outrageous, of what is similar to what else ...” *inter alia*. Moreover, this sharing of routes of interest, feeling, and significance, etc. lies so deep that it represents a kind of agreement we typically take for granted and hardly, if ever, notice. Nevertheless, this fundamental agreement in forms of life underlies both our confidence that the one who continues “Add 2” with 1004, 1008, ... is the strange one and our confidence when it comes to thick ethical terms that, for instance, what the abusive boyfriend does is cruel. Or, at least, so holds Wittgenstein on the reading offered by McDowell (and Cavell).  

While the grounding claim that McDowell (and Cavell) attribute to Wittgenstein is controversial, I introduce it now not to defend it – I do this in sections 4 and 5 – but because it is central to McDowell’s account of the philosophical resistance to the idea of virtue’s uncodifiability. Moreover, as I suggested earlier, McDowell’s account also explains resistance to criterial particularism’s animating idea, namely, that objective

14 McDowell, "Virtue and Reason" at 60.
15 On these issues, I have also been helped by reading D. Z. Phillips, *Introducing Philosophy: The Challenge of Scepticism* (1996) at 64-80 which pages refer to a chapter entitled “Criteria and Forms of Life.”
guidance for our projections of moral predicates can be had without necessary and sufficient conditions of applicability for such predicates.

On McDowell’s view, then, it would be fair to say that resistance to these two related ideas – that is, the uncodifiability of virtue and the availability of objective guidance without necessary and sufficient conditions – flows directly from an inability to accept Wittgenstein’s grounding claim. For, as Cavell goes on to say in the passage quoted above, the idea that “Human speech and activity . . . rest upon nothing more . . . than [our shared forms of life] . . . is a vision as simple as it is difficult, and as difficult as it is (and because it is) terrifying.” McDowell describes this terror and its upshot as:

[A] sort of vertigo, induced by the thought that there is nothing but shared forms of life to keep us, as it were, on the rails. We are inclined to think that [this] is an insufficient foundation . . . that [our shared forms of life] cannot be a shared conceptual framework within which something is, given the circumstance, objectively the correct move . . .

Therefore, thinking that we need a more substantial foundation to ground the correctness of our going on one way as opposed to another with a predicate, we “recoil” from the supposed abyss to an inflated conception of rule-following with two components. McDowell describes this inflated conception of rule-following and its two components in these words:

[F]irst . . . grasp of the rules is a psychological mechanism that (apart from mechanical failure, which is how we picture mistakes and so forth) guarantees that we stay in the straight and narrow; and, second . . . the rails – what we engage our mental wheels with when we come to grasp the rules – are objectively there, in a way that transcends the “mere” sharing of forms of life (hence, for instance, Platonism about numbers).

---

16 McDowell, "Virtue and Reason" at 61.
17 Ibid.
We shall consider the shortcomings of both components of this inflated conception of rule-following when we take up Wittgenstein’s grounding claim in the next section. For now, it is the second of these two components – the idea of rules as rails stretching to infinity independently of the reactions and responses that constitute our shared forms of life – that causes trouble for criterial particularism and its animating idea of a naturalism that constrains without purporting to avail itself of such rules.

Specifically, the idea that objective guidance requires rules *qua* infinite rails threatens criterial particularism’s idea of a merely constraining naturalism because rules *qua* infinite rails seem to imply the in-principle availability of necessary and sufficient conditions even if such conditions are so long and involved as to be beyond our ability to explicitly formulate them. Thus, if rules *qua* infinite rails are both required for objective guidance and imply the availability of necessary and sufficient conditions, then constraining naturalism – the idea of objective guidance without necessary and sufficient conditions – is impossible. This is McDowell’s answer to our pivotal question. The very idea of criterial particularism with its merely constraining naturalism may strike many as out of bounds, not to mention obscure, because of a tendency to think objective guidance – and thus genuine concepthood – requires rules understood as rails stretching to infinity. Of course, if Wittgenstein (a la McDowell and Cavell) is correct, then not only the psychological component but also the rules as infinite rails component of our inflated conception of rule-following must go.

---

18 Alternatively, one might take the idea of rules as rails stretching to infinity to be satisfied by the existence of a simple universal corresponding to the predicate at issue. Indeed, Moore takes just this approach with his non-naturalism. But surely this is a non-starter. Excluding this interpretation, then, the idea of rules as infinite rails does imply the in-principle availability of necessary and sufficient conditions.
4. How, then, does objective guidance depend upon the fact that we agree in our reactions and responses to criterial cases? Consider first what it means for our use of a term to be objective, that is, for our use of a term to be guided by something objective. First of all, the putatively objective feature of reality by which we guide our use of a term – call this a pattern of application – must “elude analysis in subjective terms.”

Alternatively, the pattern of application by which we guide our use of a term must extend “independently of the actual outcome of any investigation, to the relevant case,” providing an investigation-independent standard of correctness for our use of the term. In short, for our use of a term to be objective, we must be following a rule.

So now, let us consider what I call the grounding claim. That we agree in our reactions and responses to criterial cases is supposed to make possible (or ground) objective guidance. As we just saw, the (objective) meaningfulness of our use of moral terms depends on the existence of an investigation-independent patterns of application that constitute rules (or standards of correctness) for our use of each such term. However, for our use of, say, ‘cruel’ to be meaningful, it is not enough that a merely causal relation obtain between our utterances of ‘cruel’ and some investigation-independent pattern of application. Rather, if one of us is to really mean something when she labels an action ‘cruel’, the pattern of application that regulates our use of ‘cruel’ must not only be investigation-independent (for our use to be meaningful) but also graspable by us (for the meaning to be ours, that is, something that we intend). This is the first of two related points behind the grounding claim. That the investigation-independent pattern of application (or rule) regulating our use of ‘cruel’ must be

---

graspable by us also constitutes a dramatic constraint on the investigation-independent pattern of application that we could possibly mean by ‘cruel’. Specifically, in order to be graspable by us, the investigation-independent pattern of application that we mean by ‘cruel’ must be bound up with – i.e. it cannot be specified independently of – our shared reactions and responses to criterial cases. This is the second point. Together with the first, it yields the grounding claim, namely, the claim that agreement in reactions and responses to criterial cases is necessary for us to mean something by the term ‘cruel’.

Here is an argument for the grounding claim based upon these points. Call the first point (1), the second point (2), and the grounding claim (GC).

(1): If our talk of cruelty is to have objective significance, then our talk of cruelty must be governed by a rule that is graspable by us.
(2): If our talk of cruelty is to be governed by a rule that is graspable by us, then the rule governing our talk of cruelty cannot be specified independently of our reactions and responses to criterial cases of cruelty.
(GC): If our talk of cruelty is to have objective significance, then we must agree in our reactions and responses to criterial cases of cruelty.

However, strictly speaking, the grounding claim does not follow from these two points alone – as this argument makes clear. The problem is that the second point only presupposes something that needs to be made explicit. Call this presupposition (+).

(+): If our talk of cruelty is to be governed by a rule that is graspable by us, then we must agree in our reactions and responses to criterial cases of cruelty.

Call the more fully expressed version of the second point (2+).

(2+): If our talk of cruelty is to be governed by a rule that is graspable by us, then we must agree in our reactions and responses to criterial cases of cruelty and the rule governing our talk of cruelty cannot be specified independently of these shared reactions and responses.

Now, given (1), we arrive at (GC) whether we proceed via (2+) or, more simply, via (+).
But what is it for a rule (or investigation-independent pattern of application) to be ‘graspable by us’? For a person to grasp a rule in the relevant sense, she must have in her mind – that is, within her subjective mental purview – something that “obliges [her] subsequently … to judge and speak in certain determinate ways …”\textsuperscript{20} Thus, for a rule (or investigation-independent pattern of application) to be ‘graspable by us’, it must be something that each of us can bring within our own individual minds and that, once there, provides determinate guidance for our own application of the relevant concept. Call this sense of grasping (or understanding) a rule ‘subject-oriented’.

Admittedly, the subject-oriented view has been famously criticized by Saul Kripke. Describing the subject-oriented view, he writes, “We all suppose that our language expresses concepts – ‘pain’, ‘plus’, ‘red’ – in such a way that, once I ‘grasp’ the concept, all future applications of it are determined (in the sense of being uniquely \textit{justified} by the concept grasped).”\textsuperscript{21} In other words, the subject-oriented view of understanding is something “we all suppose.” It is a piece of common sense. However, as Kripke goes on to say, “[I]t seems that no matter what is in my mind at a given time, I am free in the future to interpret it in different ways – for example, I could follow the skeptic and interpret ‘plus’ as ‘quus’.”\textsuperscript{22} That is, given Kripke’s arguments, it seems that no possible contents of a person’s mind can provide the sort of guidance that, according to the subject-oriented view, grasping a rule is supposed to provide. Thus, Kripke skeptically concludes that, “We must give up the attempt to find any fact about me [i.e. about the contents of my mind] in virtue of which I mean ‘plus’ rather than ‘quus’, and

\begin{itemize}
\item \textsuperscript{20} John McDowell, "Wittgenstein on Following a Rule" in \textit{Mind, Value, and Reality} (1998), 221-262 at 221.
\item \textsuperscript{21} Saul Kripke, \textit{Wittgenstein on Rules and Private Language} (1982) at 107. Actually, as some readers may have noticed, I think Kripke overinflates our commonsense view of the nature of concepts – or at least our view of certain concepts – when he declares, “We all suppose that … once I ‘grasp’ [a] concept, all future applications of it are determined.”\
\item \textsuperscript{22} Ibid.
\end{itemize}
must then go on in a certain way."²³ For Kripke, then, the subject-oriented view of grasping a rule cannot be maintained.

Despite his criticisms, Kripke grants the subject-oriented view the status of common sense. This status arguably constitutes some presumption in its favor. With respect to the first point behind the grounding claim – that the subject-oriented view of grasping a rule is correct – we can grant this presumption and turn to the second point behind the grounding claim. For understanding the second point – how the subject-oriented view of grasping a rule constrains what rules can be if we are to grasp them (and that a conception of rules within these constraints is coherent) – involves seeing how the subject-oriented view of grasping a rule can be saved. That is, understanding the second point involves seeing that neither Kripke’s “sceptical conclusion” nor his “sceptical solution” are mandatory.

I shall approach the second point behind the grounding claim by way of a blind spot in Kripke’s case for his “sceptical solution.” The key assumption that goes unquestioned relates to the nature of rules. Kripke simply assumes that a rule – to be a rule at all – must conform to a certain conception, namely, rules as rails already there, running to infinity. According to this conception of rules:

On each occasion when the application of ‘cruel’ to some action is an issue, the cruelty rule itself determines when ‘cruel’ applies, independently of both the reactions and responses that we initially draw upon and the reactions and responses that we come to have in learning the concept of cruelty; the point of the reactions and responses that we initially draw upon and those that we come to have is to get our practice of judging and speaking in line with the rule’s impersonal dictates.²⁴

²³ Ibid at 108.
²⁴ I paraphrase here a passage from McDowell, "Wittgenstein on Following a Rule" at 231. Here is the passage in McDowell’s own words:
At each stage, say in the extending of a series, the rule itself determines what comes next, independently of the techniques that we learn in learning to extend it; the point of the learning is to get our practice of judging and speaking in line with the rule’s impersonal dictates. Ibid.
This can be put another way. Compare what this conception of rules implies about an omniscient god. Such a being would need neither to share in our initial reactions and responses nor to learn any further reactions and responses in order to know whether some novel action is right. For if this conception of rules is correct, an omniscient god could grasp the entire rule “laid [out] to infinity” and, via direct inspection of the rule so grasped, ascertain whether ‘cruel’ applies to the action in question. Of course, since our minds are much more limited, we can grasp (in the subject-oriented sense) only a finite fragment of any rule so understood. And since any finite fragment can be interpreted in different ways, grasping such a fragment cannot do what grasping a rule must, namely, provide determinative guidance for the application of a concept/term. Consequently, if this is what a rule is – i.e. something already there, running to infinity, that covers all possible uses of a term – then the subject-oriented view of grasping a rule cannot be maintained.

But why assume that this is what rules must be? Rules can be thought of differently. Barry Stroud writes:

Logical necessity [or the necessity that forces us to judge an action ‘cruel’ once we grasp the meaning of ‘cruel’] . . . is not like rails that stretch to infinity and compel us always to go on in one and only one way; but neither is it the case that we are not compelled at all. Rather, there are the rails we have already traveled, and we can extend them beyond the present point only by depending on those that already exist. In order for rails to be navigable they must be extended in smooth and natural ways; how they are to be continued is to that extent determined by the route of those rails which are already there.25

In other words, on this alternative, rules are traced out by our shared reactions and responses – “the rails we have already traveled.” Because they are not already there, running to infinity, independently of what we might say or think, rules so conceived do not provide univocal guidance for every possible use of a term. Nevertheless, while they

do not dictate one single way that we must “go on,” rules so conceived provide guidance as to future usage. Rules so conceived can do this because (1) they are bound up with our shared reactions and responses, and (2) given our shared reactions and responses, only certain ways of “going on” strike us as “smooth and natural.”

5. About our alternative conception of rules, we might now worry: is it coherent? That is, can a “rule” traced out by our reactions and responses to some finite number of cases do what a rule must, namely, provide guidance for an indefinite number of future applications of a concept? We appear to be confronted by a dilemma. On one hand, if a “rule” is defined by our reactions and responses to some finite number of cases, then we seem left with just a finite fragment. And, as we have seen, finite fragments suffer all the problems made vivid by Kripke. On the other hand, the only alternative able to avoid these problems seems to be the “infinite rails” conception of rules to which Kripke subscribes. And, of course, the “infinite rails” conception of rules is the one we wanted to avoid.

We seem to have backed ourselves into a corner. The way out begins with recognizing that, for our alternative conception of rules to be coherent, we cannot speak sensibly of rules independently of what is involved in someone’s grasping them. That is, how we conceive of the former and how we conceive of the latter mutually constrain each other. Consider first how our alternative conception of rules is constrained by our conception of grasping a rule. On the latter, grasping a rule just means coming to share the reactions and responses to criterial cases that characterize competent users of the relevant concept. Put another way, someone’s sharing these reactions and responses manifests her understanding of the relevant rule. A rule, then, so conceived cannot possess any greater determinacy than the reactions and responses the sharing of which constitutes understanding the rule. Significantly, this goes some way towards dissolving the dilemma with which we appeared to be confronted in the previous paragraph. For it is not the case, according to our alternative conception of rules, that there exists a
complete, fully determinate set of constraints of which our reactions and responses to criterial cases give us but a partial glimpse. Rather, the pattern of application suggested to us by our reactions and responses to criterial cases is all that there is with which to specify the content of the rule we are following. There is nothing, even in principle, beyond this. When someone arrives at the point where she shares the reactions and responses to criterial cases characteristic of competent users, she has brought within her subjective mental purview not just a finite fragment of a rule but the rule itself. Thus, by scaling back what a rule is, we put a comprehensive grasp of a rule within our reach, and thereby avoid the problems (made vivid by Kripke) with a less than comprehensive grasp of a rule.

However, despite our success in avoiding the first horn of the dilemma we faced, we might still be pushed onto the second horn and forced to accept the “infinite rails” conception if we cannot explain how a rule, on our alternative conception, provides guidance for future usage. To see how rules so conceived provide guidance, we need to consider how our conception of grasping a rule is constrained by our conception of rules. While our alternative conception of rules rejects the picture of rules as independently existing, infinite rails, it does not deny that a rule – however conceived – must be normative to count as a rule at all. The essential normativity of rules affects how we can think of grasping a rule. Specifically, grasping a rule – i.e. sharing the reactions and responses to criterial cases characteristic of competent users of the relevant concept – must be understood not merely as being disposed to react and respond in certain ways but also as sharing a “sense of how to go on” with other competent users of the concept. It is this “sense of how to go on” – constituting, in part, our grasp of the relevant rule itself –

26That this is it can leave us with the feeling of vertigo identified by Stanley Cavell. Moreover, that we feel vertigo helps lead us to think that only “rules as infinite rails” can provide the kind of guidance we need. Of course, Wittgenstein’s later philosophy is geared towards helping us avoid these feelings of vertigo while doing without the mythology of “rules as infinite rails”.
that guides future usage beyond the cases, criterial and otherwise, to which we have been exposed.

We should now be able to see how our alternative conception of rules coheres. Properly understood, our alternative does three things. First, it reveals the hole in Kripke’s case for his “sceptical conclusion.” Second, it makes tenable our intuitive, subject-oriented view of understanding. Finally, it supports what I have called the grounding claim (GC). To elucidate matters further, consider the following example. Confronted with a situation in which I can either lie or tell someone the truth, I confidently believe that telling the truth would be cruel based on my understanding of both the situation and ‘cruel’, the latter understanding being one I share with other competent users of the term. Thus, I judge that this action would be cruel according to my shared understanding of cruelty. As we have seen, however, Kripke regards the idea of my use being guided by my (subject-oriented) understanding of a term – even one shared with other competent users – as commonsensical but nevertheless illusory.

Let us briefly review his complaint to see how our alternative enables us to dismiss it. Allegedly, on Kripke’s view, my understanding of ‘cruel’ is radically indeterminate. It is not even supposed capable of distinguishing whether by ‘cruel’ I understand the concept cruelty as opposed to the concept gruelty (where an action is ‘gruel’ if and only if it is cruel before the year 2010 and not cruel thereafter). Nothing that might be brought before my mind as part of my “learning” the concept expressed by ‘cruel’ (and that, correspondingly, might constitute my understanding of ‘cruel’) can distinguish whether I understand cruelty as opposed to gruelty. For instance, even if cruelty was much less complicated and capable of simple definition, my understanding of the definition would not make my understanding of ‘cruel’ any more determinate. For my understanding of any definition is itself indeterminate. Generally, we make no progress accounting for my understanding of one concept by appealing to my understanding of another when, as here, the very idea of understanding a concept is at
issue. Thus, Kripke would challenge the commonsense claim implicit in my example, namely, that I understand cruelty and that, according to this understanding, I judge that telling the truth would be cruel in the situation envisioned. He would counter that, according to my understanding of cruelty, I should—with as much justification—judge that lying would be cruel in this situation.

Taking my understanding of ‘cruel’ to be constituted by my understanding of a definition leads nowhere since it involves a vicious regress. With this much, adherents of the alternative conception of rules urged here can agree. But Kripke’s “sceptical conclusion” that I understand nothing determinate by the term ‘cruel’—not even telling the truth as opposed to lying in a situation that clearly calls for the former—need not be accepted. On the alternative conception being urged here, something can be brought before my mind that enables me to achieve an understanding of ‘cruel’. Specifically, bringing clear cases of cruelty before my mind—provided these cases represent a suitable variety—can enable me to understand the concept expressed by ‘cruel’. Granted, by Kripke’s lights, even this ostensive approach to learning a concept must fail to yield any understanding of ‘cruel’. For him, no matter how many clear cases of cruelty I am exposed to as part of my instruction, exposure to such cases does not enable me to understand that ‘cruel’ means cruelty as opposed to gruel. But this disagreement over the pedagogic efficacy of exposure to criterial cases reflects the fundamental difference between Kripke’s conception of rules and the alternative conception being urged here. As we have seen, Kripke assumes that rules exist independently of our responses and reactions and, running to infinity, provide univocal answers regarding a concept’s application for any context in which the question of its application could possibly arise. That is, on this picture, rules are supposed to represent a (maximally) complete set of constraints. And indeed, given this picture of rules, exposure to criterial cases yields only part of a rule’s reach and leaves it entirely open which rule (or concept) it is that my mind is engaging. Or, we might say, my mind fails to engage with any one concept on
Kripke’s view. The sample cases to which I am exposed represent part of the extension of not only cruelty but also gruelty (and an indefinite number of other “bent” concepts as well). By contrast, however, if rules are only as determinate as the responses and reactions the sharing of which manifests understanding them, then per commonsense we have a firm grasp of many concepts, including many moral concepts. Or, at least, we have as firm a grasp as any other competent user of such concepts. We need not demean our achievement simply because we fail to measure up to some fantastical, inhuman conception of what objective understanding and rule-following is supposed to be.

6. We can now take up the issue of how criterial particularism conceives of moral reasoning. In earlier chapters we saw how criterial particularism seems to promise, if not require, something new and different along these lines. First, like Murdoch, criterial particularism sees no need to be uncomfortable with talk of moral vision. But to make such claims of assurance fully convincing, criterial particularism needs an account of moral reasoning that puts to rest concerns about dogmatism. Second, for criterial particularism to make good on its compatibility with anti-Humeanism – i.e. the claim that there are genuine thoughts that necessarily motivate – criterial particularism needs to be able to show how such (necessarily motivating) thoughts can arise via the use of reason or reasoning in some sense. Moreover, as we saw in Chapter One, not just any account of moral reasoning will do. Specifically, no sort of deductive account will work. Most importantly, however, criterial particularism requires an alternative account of moral reasoning to make good on its promise of a non-definitive analytic naturalism. For as we saw earlier in this chapter, a merely constraining analytic naturalism distinctively holds not only that the application of moral concepts admit ‘hard cases’ (i.e. cases in which the concept neither clearly applies nor clearly does not) but also, despite this lack of
determinative guidance, that it makes sense to ask whether moral concepts should apply in hard cases (i.e. whether moral concepts apply in hard cases is not simply arbitrary but subject to rational, though not determinative, constraints). Put another way, by urging constraining naturalism as an alternative to defining naturalism, criterial particularism maintains that even in cases in which the rule governing a moral concept fails to provide definitive guidance for the concept’s application, how we go on to apply the concept in such cases remains subject to rational guidance. Thus, for criterial particularism, any alternative account of moral reasoning must, first and foremost, elucidate the kind of rational guidance at issue.

Fortunately, since the primary motivation for an alternative account of moral reasoning arises from the need to understand the rational constraints that fail to specify one right answer but nevertheless bind us in hard cases, we know where to begin. For in the wake of our discussion of the grounding claim, we know both that our ability to grasp thick moral concepts depends upon our sharing various responses and reactions and that our sense of how to “go on” with thick moral concepts – the sense that rationally guides future applications – reflects these reactions and responses. So, in developing an account of moral reasoning that meets criterial particularism’s needs, we should begin with our reactions and responses.

In taking on this task, however, we should be aware that our responses and reactions will need to yield something different from what they have yielded already in defense of the grounding claim. Recall that the realistic conception of rules implicated by the grounding claim appeared to create a dilemma. To avoid the second horn of this dilemma, we need to articulate a more nuanced account of how our reactions and responses guide future applications.

---

27 See section 5 above.
dilemma, we needed to realize that the responses and reactions we share – the responses and reactions that both make possible and reflect our grasp of thick moral concepts – constitute no inert, brutally causal phenomenon but a more-or-less lively sense of how to apply such concepts to future cases. And, as we saw above, it is this sense of how to “go on” reflected in our shared responses and reactions yet possessed by each of us individually that rationally guides future usage, thereby securing the normativity that it was feared could not be had if rules were anchored in our responses and reactions. What we need to realize now is that, when it comes to hard cases, this kind of rational guidance is no help precisely because it flows from our shared sense of how to “go on”. For this shared sense gives out in hard cases. Indeed, the very fact that it gives out is what makes hard cases hard.

Looking at the same point from a different angle, notice that our grasp of a rule – that is, our coming to share in certain responses and reactions – can be said to (rationally) guide future applications in two different ways corresponding to the two broad kinds of cases – easy and hard – that we confront. In brief, easy cases of a concept’s applicability are clear and straightforward while hard cases range from the merely difficult and controversial to those that may not appear – at least not initially – to even raise the question of the concept’s applicability. Now the rational guidance enlisted to defend the grounding claim and derived from the sense we each share of how to go on with a concept – this is the kind that governs easy cases. Therefore, in developing an account of

---

28 Under hard cases, I include ‘cases that may not appear to even raise the question of a moral concept’s applicability’ to reflect Murdoch’s emphasis on description as the crucial moment in moral thought. In other words, as I understand Murdoch, part of the reason why coming up with appropriate descriptions is the crucial moment in moral thought is that this task is nowhere as easy it sounds.
moral reasoning that satisfies criterial particularism’s needs, we must look beyond our shared sense of how to go on that our responses and reactions reveal to the underlying responses and reactions themselves. Moreover, we must consider these underlying responses and reactions in their myriad, cultural manifestations even to the extent that they are not shared widely, let alone by all competent speakers. For when our responses and reactions do not naturally chime in unison, they are not simply mute. Our responses and reactions still have something to say about how we should go on with a concept even when they fail to jibe. As we shall see, it is to these intimations that we appeal in reasoning about hard cases of a concept and in trying to choose between the different ways we could project the concept beyond what our shared sense dictates.

To clarify the role of our responses and reactions and further frame the alternative account of moral reasoning that follows, I return to Stroud’s helpful gloss on Wittgenstein. Though quoted earlier, it is worth quoting again:

Logical necessity . . . is not like rails that stretch to infinity and compel us always to go in one and only one way; but neither is it the case that we are not compelled at all. Rather, there are the rails we have already traveled, and we can extend them beyond the present point only by depending on those that already exist. In order for the rails to be navigable they must be extended in smooth and natural ways; how they are continued is to that extent determined by the route of those rails which are already there.29

In other words, as maintained earlier, we can be rationally compelled without conceiving of rules as infinite rails. Going on rationally with a concept “can . . . only [be done] by depending on those [rails] that already exist,” “the rails we have already traveled.” These rails – or, less picturesquely, our responses and reactions to the various cases we have thus far confronted in applying a concept – constrain how we go on with a concept in two

29 Stroud, "Wittgenstein and Logical Necessity".
ways. As we know, one sort of rational constraint derives from the responses and reactions we have to easy (or criterial) cases and the shared sense of how to “extend [the rails] beyond the present point” that these responses and reactions reflect. More importantly, given our current focus, this shared sense of how to go on does not “compel us always to go [on] in one and only one way” and, therefore, another sort of rational constraint comes into play. So, Stroud agrees that our responses and reactions constrain how we go on with a concept even when they fail to speak with one voice. More interestingly, Stroud suggests how our responses and reactions constrain us in such cases. To wit, our projections in hard cases must be “navigable” (or make sense) and, to be navigable, they must extend our responses and reactions in “smooth and natural ways.”

About what this means more precisely – this talk of smooth and natural ways of going on – it is enough for now to observe two things. First, this talk implies some kind of continuity requirement. Second, whether some way of going on in a hard case meets this requirement can be supported by appealing not just to the responses and reactions we share but also – and more especially – to those we do not share. For showing that some way of going on in a hard case meets the continuity requirement will involve showing that it maintains continuity with what has gotten us this far, namely, our shared sense of how to go on. However, because this shared sense reflects only a subset of our responses and reactions, it can be interpreted differently – and different ways of going on in hard cases can legitimately be said to maintain continuity with it – according to which strands of our broader responses and reactions are thought to figure more centrally (or importantly) in it. Thus, we must consider our broader responses and reactions in their myriad, cultural manifestations because only by juxtaposing our shared sense of how to
go on with some of these broader, underlying responses and reactions can we highlight what aspect(s) of our shared sense we regard as important and, correspondingly, how the different projections being proposed in a hard case maintains continuity with this (or these) important aspect(s).\(^{30}\)

6. We can now begin setting out the alternative account of moral reasoning that criterial particularism requires. As with criterial particularism itself, the account that follows will be informed by Wittgensteinian ideas about not only criteria (via McDowell) and rule-following (via Stroud and McDowell) but also some that have not been mentioned yet. Again, my goal is not exegesis that tries to stitch these ideas together in a seamless interpretation of Wittgenstein’s views but “data-mining” that extracts from Wittgenstein helpful nuggets about certain issues.\(^{31}\) Two notions from Wittgenstein have yet to be explicitly introduced but will loom large: (i) aspect perception and (ii) the related idea of perspicuous representation. About aspect perception (or seeing as), we need say little right now since this idea is already sufficiently familiar from Wittgenstein’s duck-rabbit illustration.\(^{32}\) But about Wittgenstein’s less familiar notion of perspicuous representation, it will be helpful to say more.

\(^{30}\) Interestingly, the way I suggest we can reason about projecting moral concepts into hard cases has an affinity with the approach to moral-political argument favored by Michael Walzer. Michael Walzer, *Interpretation and Social Criticism* (1987). Walzer also recognizes a continuity requirement of sorts since, on his view, justifying a certain distribution of some good (e.g. medical care) requires showing the proposed distribution accords with what he calls the social meaning of the good in question.

\(^{31}\) As it happens, the different ideas that I take from Wittgenstein do seem to mesh quite well, even tightly. Perhaps they might provide the basis for an interesting new reading of Wittgenstein. But, as I say, that is not my concern here.

\(^{32}\) Ludwig Wittgenstein, *Philosophical Investigations* (1953) at 166. Section or page numbers preceded by *PI* refer to the *Investigations*. 
In the course of an important series of remarks that illuminate his method in

*Philosophical Investigations* and other later writings, Wittgenstein tells us that,

“[Philosophical] problems are solved, not by giving new information, but by arranging what we have always known.” (*PI* § 109) Shortly after this remark, he adds:

The concept of a perspicuous representation is of fundamental significance for us. It earmarks the form of account we give, the way we look at things ...

A perspicuous representation produces just that understanding which consists in ‘seeing connections’. Hence the importance of finding and inventing intermediate cases. (*PI* §122)\(^{33}\)

Now it may not be clear why I included the earlier remark from Section 109, but a perspicuous representation involves just what Wittgenstein talks about there: a certain arrangement of things that, in some sense, we already know. That is, if done well, how we arrange (or present) certain facts can get us to see these same-old facts in a new way, making us aware of relations implicit in these facts that we had previously overlooked.

Unfortunately, despite its “fundamental significance,” Wittgenstein barely mentions what goes into a perspicuous representation, or what makes for a convincing arrangement of familiar facts. In the *Investigations*, we have only the one, brief sentence

\(^{33}\) In quoting from *PI* §122, I have slightly altered Wittgenstein’s original ordering of his remarks to fit my emphasis. Specifically, the two short paragraphs I quote have been reversed from the order in which they appear in the *Investigations*. However, Wittgenstein uses my ordering in an almost identical passage from his remarks on Frazer’s *Golden Bough*, where he writes:

The concept of perspicuous representation is of fundamental importance for us. It denotes the form of our representation, the way we see things . . .

This perspicuous representation brings about the understanding which consists precisely in the fact that we “see the connections.” Hence the importance of finding connecting links.

from §109. Nevertheless, in some brief remarks on aesthetics from roughly the same period in which he wrote the relevant sections of the *Investigations*, we get a somewhat clearer view of the very idea of a clear view (or perspicuous representation), including what is involved in generating one.

The brief remarks that illuminate the idea of a perspicuous representation come from Wittgenstein’s Cambridge lectures in the 1930’s and were transcribed and later published by various attendees including, most famously, G.E. Moore. In the relevant part of his notes, Moore records the following:

*Reasons*, he said, in Aesthetics, are “of the nature of further descriptions”: e.g. you can make a person see what Brahms was driving at by showing him lots of different pieces by Brahms, or by comparing him with a contemporary author; and all that Aesthetics does is “to draw your attention to a thing”, to “place things side by side”.

Alice Ambrose, then a student of Moore’s and who apparently attended the same lectures, recorded something similar:

What is the justification for a feature in a work of art? . . . what reasons can one give for being satisfied? The reasons are further descriptions. Aesthetics is descriptive. What it does is to draw one’s attention to certain features, to place things side by side so as to exhibit these features. To tell a person “This is the climax” is like saying “this is the man in the puzzle picture”. Our attention is drawn to a certain feature, and from that point forward we see that feature. (pp. 38-39)

---

34 Though perhaps the entire “argument” of the *Investigations* is best understood as Wittgenstein’s attempt to provide perspicuous representations of our life with language in its various manifestations, especially those that have given rise to philosophical problems. Thus, in the *Investigations* at least, he appears content to show – rather than say – what perspicuous representation involves.


Finally, in other lectures specifically on aesthetics, Wittgenstein reportedly said: “[W]hat we really want, to solve aesthetic puzzlements, is certain comparisons – grouping together of certain cases.”

Taken together, these remarks bring out several key points about perspicuous representations. First, their relevance is not limited to problems in the philosophy of language and philosophical method. For, as these remarks suggest, Wittgenstein would resolve so-called “aesthetic puzzlements” in the same way, that is, by using perspicuous representation. Moreover, Moore also records the thought that ethics involves the same kind of “reasons”. Second, as the talk of ‘reasons’ and ‘justification’ implies, trying to assemble a perspicuous representation of, say, “what Brahms was driving at” in a particular piece, is a kind of reasoning. In other words, it involves providing your audience with the kind of reasons that Wittgenstein is talking about in these remarks. And thus, lastly, these remarks tell us something about the kind of reasons (or reasoning) that trying to assemble a perspicuous representation involves. For one thing, the reasons involved are supposed to be descriptions as in ‘This is the climax’ (said of Brahms’ piece) or ‘Here are the ears, here is the mouth’ (said of the duck-rabbit figure to someone who isn’t seeing it as, say, a rabbit). Moreover, especially in complex cases, these reasons-cum-descriptions often need to do more than simply point out some feature of the case at hand. They may need to describe and thereby draw our attention to one or more related examples, setting the case at hand and the other example(s) “side by side” for our consideration. Of course, this kind of reasoning by persuasive description may need to

---

38 Moore, "Wittgenstein's Lectures in 1930-33" at 106.
do yet other things to successfully achieve a perspicuous representation. So, while these brief remarks provide some grist for our alternative account of moral reasoning, we still need more.

7. To continue building our alternative account of moral reasoning, we next consider how John Wisdom\(^{39}\) appropriates these notions of perspicuous representation and aspect perception. A chief concern that Wisdom raises about perspicuous representations is this: how can they do any rational-cognitive work if they merely rearrange what we already know and thus, at least in a sense, do not tell us anything new? Or, as he characteristically puts it: “How does anyone ever say to another anything worth saying when he doesn’t know anything the other doesn’t know?”\(^{40}\) (L 2) For his answer, Wisdom asks us to:

Imagine someone . . . trying on a hat. She is studying the reflection in a mirror like a judge considering a case. There’s a pause and then a friend says in tones too clear, ‘My dear, it’s the Taj Mahal.’ Instantly the look of indecision leaves the face in the mirror. All along she has felt there was about the hat something that wouldn’t quite do. Now she sees what it is. (L 2)

In this context, uttering the description ‘It’s the Taj Mahal’ is enough (of a reason) to induce a perspicuous representation. For the woman trying on the hat now sees it under some very different aspects (e.g. ‘monumental’, ‘too magnificent by half’) whereas previously she may not have thought of it under any definite aspect. Accordingly, this

\(^{39}\) Mark Rowe, from whom we shall hear more below, recently called Wisdom “the one analytic philosopher who made Wittgenstein’s remarks on [aesthetic] explanation absolutely central to his conception of philosophy.” Mark W. Rowe, "Criticism Without Theory" in *Wittgenstein, Aesthetics and Philosophy*, ed. Lewis (2004), 73-93 at 83. However, I first learned of Wisdom’s relevance to particularism from Peter Levine, *Living Without Philosophy: On Narrative, Rhetoric, and Morality* (1998).

woman now has a clear view of the hat: she knows *that* it won’t do, and *why*.\(^{41}\)

Furthermore:

[T]he hat could be seen perfectly clearly and completely before the words ‘The Taj Mahal’ were uttered . . . [T]o call a hat the Taj Mahal is not to inform someone that it has mice in it . . . it is more like saying to someone ‘Snakes’ of snakes in the grass but *not* concealed by the grass but still so well camouflaged that one can’t see what’s before one’s eyes. (L 2-3)

In other words, the friend’s saying ‘It’s the Taj Mahal’ fails to give the woman trying on the hat any new information in the sense of facts impersonally available to any competent observer. There are no such facts – no information - about the hat that the friend and the woman trying it on are not both fully aware of – no mice, no hidden surprises. They both have eyes, and can both literally see the hat equally well. However, the friend also *sees* the hat as excessive, over the top, or just simply *too much*, and – significantly – he hits upon the words ‘It’s the Taj Mahal’ to convey this (i.e. to convey how he sees the hat in this richer, aspect–laden sense). So despite there being a sense in which the woman’s friend does not tell her (or rely upon) anything about the hat, current trends in fashion, or the Taj Mahal that she doesn’t already know, her friend does *show* her – make her aware of – something about the hat that had thus far escaped her. Consistent with this, Wisdom emphasizes that, “It is possible to have before one’s eyes all the items of a pattern and still to miss the pattern.” (G 153) Moreover, where we are talking about a “recognition of patterns . . . easily missed,” we are also talking about “what is so and therefore . . . the

\(^{41}\)This recalls the idea of self-justifying perceptual experiences from foundationalist epistemologies such as Chisholm’s and recently rehabilitated by Bill Brewer, *Perception and Reason* (1999), and, assuming that we can make good on our right to such talk, it will help us avoid the charge of dogmatism broached at the close of Chapter Three.
facts, though not in the simple ways which first occurred to us.” (G 154) Therefore, perspicuous representations do make a rational-cognitive difference.\textsuperscript{42}

But Wisdom is also concerned about \textit{how} perspicuous representations make a rational-cognitive difference and thus, more specifically, about the reasoning that goes into them. Consider again the woman trying on the hat and her friend’s remark. The reason-cum-description he gives – ‘It’s the Taj Majal’ – does not directly point to a feature of the hat but rather affects “certain comparisons – [a] grouping together of certain cases.” By setting the hat and the Taj Mahal “side by side” in this way, the friend makes it much clearer – much easier for the woman herself to see – why the hat won’t do. But Wisdom does not end with this example. He anticipates someone objecting:

If you want to bring out the fact that we sometimes use words neither to give information as when we say ‘That will be fifteen guineas’ nor to express and evoke feeling as when we exclaim ‘Fifteen guineas!’ but to give a greater apprehension of what is already before us then why don’t you choose a better example? . . . Why choose as an example a statement so preposterous and loosely worded that the question ‘Is it true?’ is hardly a question at all . . . (L 6)

To which he responds:

That is why I chose it. We all know . . . that there are questions which though they don’t call for further investigation but only for reflection are yet perfectly respectable because the reflection they call for may be carried out in a definite demonstrative procedure which gives results Yes or No. [But] . . . this isn’t the only sort of reflection and . . . the other sorts are not poor relations…. For they too take us towards a better apprehension of reality . . . We do not say that drama, novels, poetry, never show us anything of the truth. (L 6-7)

Still, Wisdom acknowledges that, “[W]e are apt to half-feel that what is said in poetry is always more a matter of fancy than of fact, that it is not within the scope of reason.” (L 7)

\textsuperscript{42} Or, anyways, they can make such a difference for those who find them minimally persuasive (i.e. persuasive enough to bring the relevant object under the new aspect at least temporarily).
To rebut this nagging suspicion, Wisdom adopts three strategies. First, he argues from paradigm cases, relying especially on a legal analogy. In the more familiar, less idiosyncratic contexts that these paradigm cases represent, we also use descriptive language to help clarify what we think even when all the facts are agreed, and doing this in these contexts surely counts as reasoning. For instance:

In courts of law it sometimes happens that opposing counsel are agreed as to the facts and are not trying to settle a question of further fact . . . but are concerned with whether Mr. A who admittedly handed his long-trusted clerk signed blank cheques did or did not exercise reasonable care, whether a ledger is or is not a document, whether a certain body was or was not a public authority. (G 157)

And, as Wisdom notes elsewhere, “[W]hen the facts are agreed upon we must still hear argument before we give judgment.” (L 7) Moreover:

The facts agreed upon, still a question is before the bar of reason as when, the facts agreed upon, still a question comes before a court. ‘Was there negligence or was there not?’ To such a question maybe the answer is ‘Yes,’” maybe the answer is ‘No,’ maybe the answer is neither ‘Yes’ nor ‘No.’ But the question is not beyond the scope of reason. On the contrary it calls for very careful consideration and not the less when what’s relevant is not enumerable because there’s not a separate name for every relevant feature of the case and an index to measure its weight. (L 21)

In other words, the question often arises in the law: how should certain facts that all parties accept be characterized? As this or as that? Accordingly, a great deal of legal argument involves doing just the kind of thing that we have been considering: describing some agreed-upon set of facts in such a way as to rearrange our understanding of these facts and, correspondingly, prompt us to see them in a different way (e.g. as exhibiting reasonable care or not). And surely what we call opening and closing arguments in a trial – despite what else they may unfortunately incorporate – involve genuine reasoning.

Likewise, then, the more general kind of movement in thought of which legal argument is
but one instance – “a move in thought which from a mass of data extracts and assembles … something which . . . doesn’t go beyond the data [but] gives an apprehension of reality which before we lacked” – also constitutes a genuine kind of reasoning.

This kind of argument by analogy comprises Wisdom’s first strategy. His second strategy for defending all this – that all these varied attempts to see things differently by (re-) describing them amounts to a genuine form of rational reflection\(^{43}\) – involves looking under the hood, so to speak, of this putative kind of reasoning and analyzing its general structure. For if it amounts to a genuine kind of reasoning, we ought to be able to discern some such structure. And, while Wisdom readily concedes that this alternative form of reflection is not “carried out in a definite demonstrative procedure which gives results Yes or No,” he clearly wants to resist the view expressed in Hume’s famous remark\(^{44}\) and show how reason is nevertheless at work when we follow either our own or another’s (re-) descriptive prompts and reflect in this alternative mode. Accordingly, he observes:

In such cases . . . the process of argument is not a chain of demonstrative reasoning. It is a presenting and re-presenting of those features of the case which severally co-operate in favour of the conclusion . . . The reasons are like the legs of a chair, not the links of a chain. (G 157)

\(^{43}\) Earlier I alluded to other paradigm cases that Wisdom used to support his first strategy, his argument by analogy. However, I only discussed one: the legal paradigm. The other main paradigm to which he appeals – critical (or aesthetic) reasoning – I discuss below in connection with seeing aspects.

Though I do not discuss it here, I believe a third paradigm would be the process of artistic creation itself. It is almost a cliché that serious artists try to get their audience to see something familiar in a new way, and the process of fashioning the work so that it has (or can have) this effect would therefore be another instance of trying to arrive at a perspicuous representation.

\(^{44}\) I mean the famous advice with which Hume closes *An Enquiry Concerning Human Understanding* declaring that everything except abstract or experimental reasoning is sophistry and illusion.
Furthermore:

[A]lthough the discussion is *a priori* and the steps are not a matter of experience, the procedure resembles scientific argument in that the reasoning is not vertically extensive but horizontally extensive – it is a matter of the cumulative effect of several independent premises, not of the repeated transformation of one or two. (G 157)

Finally, he adds:

[B]ecause the premises are severally inconclusive the process of deciding the issue becomes a matter of weighing the cumulative effect of one group of severally inconclusive items against the cumulative effect of another group of severally inconclusive items … (G 157)

This glance under the hood, then, no doubt reveals some important points about the putative reasoning involved in assembling perspicuous representations. For now, however, we need only notice that the general structure uncovered by Wisdom further supports removing the putative label from this kind of reasoning.

Lastly, Wisdom suggests yet another strategy for regarding the assembly of perspicuous representations as falling within the bounds of (i.e. as constrained by) reason. Ironically, while less obvious and direct, this last defense is arguably the most probative.

Now a perspicuous representation is a representation *of* something. Thus, as we have seen, assembling a perspicuous representation of some object involves describing it. However, when we describe an object with the aim of getting another to see it in a new light – as a nonobvious extension of some familiar concept – we often describe it in ways that may be controversial, poetic, obscure, and even preposterous. Consider, for example, the following objects that someone sought to perspicuously represent and the descriptions that were used to do so:

(i) ‘It’s the Taj Mahal’ said of a hat.

(ii) ‘He was God incarnate’ said of Jesus.
(iii) ‘It’s an apartheid system’ said of U.S. immigration policy.

(iv) ‘If they be two, they are two so/As still twin compasses are two’ said of two lovers.

And precisely because constructing a perspicuous representation often involves using descriptions such as these, this activity can be difficult to see as constrained by reason. But Wisdom shows that the descriptions involved – even the controversial, poetic, obscure and preposterous – are subject to rational constraints. That is, we have at least some idea of “what to look for” – what to attend to – for descriptions like this to be borne out by the facts. Take one of Wisdom’s chief examples, (ii) above. If we call someone God incarnate, we know that (within the Christian tradition at least) how he conducts and comports himself must evince extraordinary goodness. Conversely, since Nero was a cruel tyrant, we also know that saying ‘In Nero God was incarnate’ is absurd. Anyone with the slightest idea of who Nero was and what it could mean to say of someone that he is God incarnate would understand that such a thing could not be said of Nero. Or, as Wisdom puts it, such a description is “against all reason and therefore not beyond the scope of reason.” (L 20)

In fact, we can go even further. For the descriptions that perspicuous representations typically involve are subject to rational constraints in a sense – or, from a direction – that Wisdom neglects. Recall Wittgenstein’s comment that, in trying to perspiciously represent something, we would need to find or invent “intermediate cases” and notice that the descriptions we are talking about often highlight just such cases for our consideration. In (iii) above, for instance, when a proponent of immigration reform says of the status quo ‘It’s an apartheid system,” she raises apartheid South Africa as an intermediate case, a case that can – and the reformer hopes will – serve as a bridge from
our existing view of the immigration status quo to some new, clearer view of the matter. The point of saying ‘It’s an apartheid system’, then, is to get us to see current U.S. immigration policy not only as an apartheid system (at least in some respects) but also—and ultimately—as cruel, unjust and urgently needing reform. In other words, seeing the U.S. immigration status quo as apartheid is only meant to be a stepping stone to seeing it as cruel, unjust, etc.

Besides elucidating further the analogical nature of the reasoning involved in perspicuous representations, this discussion of descriptions as intermediate cases brings out how such descriptions face rational constraints from another direction, not just the one that Wisdom discusses. On one hand, as we have learned from Wisdom, what apartheid means rationally constrains what kinds of things we can describe as apartheid systems. That is, our rough idea of an apartheid system must fit some significant aspects of current U.S. immigration policy if the description ‘It’s an apartheid system’ is to play its intended role in bringing about a perspicuous representation of such policy. On the other hand, since apartheid is intended here as an intermediate case, it must also fit with the ultimate view of current U.S. immigration policy being urged: that such policy is cruel, unjust, and urgently needing reform. Put more generally, the ultimate view of things that one wants to represent rationally constrains the intermediate cases to which one can appeal, and, correspondingly, the descriptions that one can use. And, given this constraint, the description ‘It’s an apartheid system’ makes sense while, e.g., ‘It provides the lifeblood driving much of our economy’ does not. For, of these two, only the former links in any straightforward way with the ultimate view being urged, namely, U.S. immigration policy as cruel, unjust, and urgently needing reform. That so many of the
descriptions used to construct perspicuous representations are rationally constrained in an additional way beyond the one identified by Wisdom only reinforces his claim that assembling perspicuous representations through creative description lies within reason’s purview.

Having discussed how we can bring someone to see an object – a hat, a person, a situation – in a new light and defended this process of perspicuous representation as a genuine kind of reasoning, the relation between perspicuous representation and aspect perception should now be clear itself. For the whole point of a perspicuous representation and the reasoning-cum-describing that goes into it is just to get someone to see a familiar object under the aspect of some familiar concept, albeit as a novel extension of said concept. With this relation made explicit, however, we need to confront some misconceptions about aspect perception that could be thought to threaten our hard-won understanding of perspicuous representation as a legitimate kind of reasoning. Indeed, we can further support this alternative account of moral reasoning by examining how aspect perception admits a great deal more depth and richness than many believe.

8. For a better understanding of aspect perception and its suitability as the end of the distinctive kind of reasoning we are considering, we can turn to the analysis offered by Mark Rowe, someone who has also worked the Wittgenstein/Wisdom vein being mined here.45 In the course of developing an alternative account of aesthetic (or critical) reasoning along broadly similar lines to the alternative account of moral reasoning being developed here, Rowe acknowledges that, “Some philosophers have felt uneasy about the analogy between criticism [i.e. critical reasoning] and the prompting of aspects . . . “ (75)

45 Parenthetical page numbers in this section refer to Rowe, "Criticism Without Theory".
But this unease, Rowe suggests, arguably stems from the excessive focus on visual figures such as “the ubiquitous duck-rabbit” and the misconceptions about aspect perception that arise from this distorted focus. He observes:

If the duck-rabbit were seen as the paradigmatic example of all seeing-as, then we would expect the following to be the case: [i] the ambiguity of the figure would always reside in what was represented; [ii] seeing-as would always involve an illusion of depth or three dimensions; [iii] in all cases the imagination would add something (the rabbit's body for instance) which was not shown in the picture; [iv] the ambiguity would always be binary – either this or that. (75)

However, as Rowe suggests:

[I]f we consider a line of six dots . . . which can be grouped in a variety of different ways (1-5, 2-4, 3-3, 4-2, and so on) we can see that these generalizations do not hold. When our imagination regroups the dots, there is no illusion of three dimensions; there is no change in what is represented (because it is not a representational figure); no extra features are added by the perceiver's imagination; and the figure can be grouped in at least five obvious ways. (75)

Eschewing the misconceptions highlighted by this simple line of six dots and the different ‘aspects of organization’ it suggests, we should realize that:

(i) the underlying ambiguity that aspect perception presupposes can sometimes be found not in what is represented by the “data” but in how we organize them,
(ii) aspect perception need not involve an illusion of depth or three dimensions,
(iii) aspect perception need not involve our imagination’s filling out what would otherwise be an incomplete visual picture, and
(iv) the underlying ambiguity that aspect perception presupposes need not be binary.

As we shall see, (i) and (iv) possess the greatest significance for us. Still, given that they reflect the limits of visual paradigms, (ii) and (iii) also point in the right direction, towards overcoming other misconceptions about aspect perception.

Beginning the second stage of his analysis of aspect perception, Rowe notes that “The row of dots and the duck-rabbit still share [at least] three features . . .”: (1) a finite
number of aspects, (2) an instantaneous and immediate change of aspect, and (3) a visual element. However, we would again misconceive aspect perception if we generalized these three features to “all cases of seeing-as.” (75-77) To see why the number of aspects under which a thing can be seen need not be limited, Rowe asks us to consider: “[A] group of five dots: four are positioned as if they stood at each corner of a square, the fifth stands in the middle.” About this figure, he then notes:

... I can see [it] as an upright pyramid seen from above or as an inverted pyramid seen from the same position.... I can [also] imagine the central dot as advancing (so that either the pyramid grows taller or the inverted pyramid becomes shallower) or as retreating (so that either the pyramid grows shallower or the inverted pyramid grows deeper). This is an example of . . . ‘aspects of degree’ because there are an infinite number of shapes which the figure can represent. (75)

It is a misconception, therefore, to think that we can see an object only under some limited number of aspects, let alone as only this or that. Moreover, as Rowe points out:

Most criticism seems more analogous to discussions of this kind of case [i.e. aspects of degree] than it does to the dramatic switching effect that can be caused by . . . the duck-rabbit . . . . [M]ost criticism induces not rapid changes of kind, but subtle shiftings of degree – more or less sentimental, for example. (75)

---

46 About the “subtle shiftings” in our experience of a work that criticism can induce, Rowe declares that, “Sometimes we will have difficulty in describing our change of experience in words at all.” (75) Ironically, Rowe’s talk here of “subtle shiftings” brings to mind other, more fundamental subtleties that are also related to criticism but which Rowe neglects. I refer to the subtle differences not in our experience but in works of art themselves which it is the task of good criticism to help make available, typically by doing just what Rowe dismisses: describing them in words.


[T]alk [about art] ... has a direct influence on taste. Careful studies have shown that what people physically see in a painting bears a close relation to what they have been told or taught ... which ... confirm[s] ... common observation. All teachers know that to educate is in large part to suggest ...
Thus, realizing that the different aspects under which we can see certain objects are essentially unlimited makes it easier to accept a central role for aspect perception in our reasoning about complex objects such as works of art, literature, and real-life situations.47

Related to the thought that aspect perception only involves a limited number of aspects is the idea that changes of aspect must be “instantaneous and immediate.” But this also misconceives aspect perception. For even some two-dimensional visual figures – not to mention more complex items – “strike us as initially chaotic” and with objects of both kinds “the mind finds the new aspect only much more gropingly and by degrees.” (75-76) The opening of e.e. cummings’ poem ‘ygUDuh’, for example, illustrates both how an aspect can gradually dawn on us and the reasoning-cum-description involved in bringing this about:

This is the whole point of Criticism … Nothing, therefore, is more important to the life of art … than tending and developing its critical vocabulary, just as nothing is more important to the connoisseur who wants to see rightly in order to increase his range of delights. I do not mean to say, of course, that merely by using the right words a person … will enlarge [her] stock of artistic pleasure and deepen [her] understanding of what [she] likes. The truth is rather that lacking those right words and true ideas, the fleeting perceptions will not [be able to] sustain themselves …

Furthermore, he adds:

It is an error to suppose that the right words rise naturally to the lips of any reasonable man when faced with the object of his interest. Critical terms are rare pearls … Ibid.

47 To the last kind of complex object – real-life situations – I intentionally omitted the qualifying phrase ‘that present moral or practical issues’ because it adds nothing, assuming that we take seriously the current point about aspect perception. For if real-life situations can be seen as an unlimited number of different things, then it would seem to be a corollary that any real-life situation can be seen as raising moral or practical issues.
YgUDuh

ydoan
yunnuhstan

ydoan o
yunnuhstan dem
yguduh.

As Rowe puts it, “When we are told that the opening line can be paraphrased as ‘You’ve got to’ … we gradually discover comprehension spreading throughout the entire poem.”

(78) In other words, the description ‘The opening line can be paraphrased as “You’ve got to”’ helps us see how to start making sense of not just the opening line but the rest of the poem (the second line as ‘You don’t’, the third line as ‘You don’t understand’, and so on). Again, provided we do not misconceive it, aspect perception fits well with our alternative account of reasoning. For the idea that coming to see some new aspect may be a gradual process – not necessarily sudden and dramatic – makes much more sense given the complexity of the hard cases that, according to criterial particularism, we reason about.

The last misconception about aspect perception that Rowe urges us to avoid has been close-at hand throughout our discussion: the idea that aspects must be visual, that seeing-as must be literally seeing. But this need not be so. Just recall our previous example. When considering e.e. cummings’ poem ‘ygUDuh’, the aspect that dawns upon us after a little help is not (primarily) visual but conceptual.48 Letters that first appear

---

48 I say ‘not (primarily) visual but conceptual’, although – and in part because – this poem:

... illustrates the close connection between the visual and the conceptual, and the naturalness of our use of all those terms that do service for both sight and knowledge. We can say “I see what you mean” or “I know what you mean,” and “know” is believed to derive from the same root as “ken.”
nonsensical we come to see as meaningful. Meaning, of course, is not something we literally see. It is something we understand: seeing as understanding.\footnote{49} Furthermore, while they may not leap to mind given the inclination to think of aspect perception in visual terms, conceptual aspects free of any element of visual depiction not only exist but are ubiquitous. For instance, the different ways in which sentences such as ‘Why buy the cow when you can have the milk for free’ or ‘Why buy the pig when all you want is a little sausage’ can be interpreted count as such aspects. Or consider when:

We open the newspaper and see something called a letter to the editor, and we read it as that. Then we are puzzled by its contents and try to decide how to take it, whether to read it as an ironic or as a straight-but-stupid letter to the editor.\footnote{51}

In other words, even a garden-variety activity such as reading a newspaper involves conceptual aspects given the different ways that a particular piece of writing can be understood: e.g., first, as a letter to the editor (or as an instance of some other genre); then, as ironic or as straight-but-stupid. More importantly, this last example points to how fruitful recognizing conceptual aspects can be. For not only individual words and sentences admit conceptual ambiguity – and, correspondingly, conceptual clarification – but so do the long expanses of words and sentences that comprise integral wholes such as

\begin{quote}
The eyes of the person who knows English will behave differently from the eyes of someone who knows only the alphabet when they scan astichintime. John Reichert, \textit{Making Sense of Literature} (1977) at 4. However, even to someone who knows English, scanning ygUDuh will not typically yield comprehension unless one has been given some clue. Making playful use of a certain East coast dialect – I suspect it’s Brooklyn-ese – cummings’ poem shows the importance of attending to how words sound rather than how they look, at least for the purpose of understanding them. For just as with dialects themselves, it is much easier to understand the words of cummings’ poem when spoken aloud than when read silently.\footnote{49} See previous footnote.\footnote{51} Reichert, \textit{Making Sense of Literature} at 5.
\end{quote}
novels, poems, and plays (and, yes, letters to the editor). And it is a short step from reasoned discussion of how – or under what aspects – we can best understand such complex works to reasoned discussion of how – or under the aspect of what moral concepts (if any) - we can best understand certain complex situations.

This last misconception dismissed, aspect perception – or “seeing” x under the aspect of F – often means not literally seeing but actively understanding x as a novel instance of F. Moreover, with space for conceptual aspects secured, aspect perception can readily comprehend the sophisticated distinctions that animate critical reasoning about art and literature. We should have no qualms, then, if our distinctive kind of reasoning via description has aspect perception as its goal. For if aspect perception thus conceived can comprehend the conceptual richness that aesthetic reasoning involves, it can also comprehend the distinctions involved in reasoning about how to apply moral concepts in hard cases. But we are getting ahead of ourselves. Rowe’s analysis of aspect perception focuses not on moral but on aesthetic reasoning, and we still need to discuss an important implication that aspect perception has for aesthetic reasoning. Ultimately, aspect perception also has a similarly important implication for moral reasoning as well. But again first things first.

9. Having shown that it admits sufficient depth and richness, Rowe maintains that aspect perception makes good sense of certain key features of aesthetic reasoning. Specifically, when understood in terms of aspect perception, criticism’s relationship to evaluation can be readily accommodated and explained. So what is the relationship

---

52 Notice that I do not say ‘complex practical situations’. For this would imply that hard cases of moral concepts can arise only in situations requiring a decision about how to act. But difficulties with moral concepts can arise in other contexts, say, in how we are to describe some past event.
between aesthetic criticism and evaluation? Consider Alfred Kazin’s remark: “Criticism exists, after all, because the critic has an intense and meaningful experience of a work.” Accordingly, the point of critical writing and reasoning is to convey this “intense and meaningful experience” to others, to help them on their way to sharing it. Or as Wisdom puts it:

[A] good critic by his art brings out features of the art he writes about, or better, brings home the character of what he writes about, in such a way that one can feel and see, see and feel that character much better than one did before. Such a critic tackles aesthetic problems, with his head and with his heart, with his heart and with his head and so tackles them seriously.53

Thus, aesthetic reasoning has as its primary purpose getting others to share in the critic’s evaluative responses to a work. This is the relationship between criticism and evaluation.54

To begin to see how analyzing critical reasoning in terms of aspect perception supports and explains the former’s intimate involvement with evaluation, consider first a related feature of critical reasoning. Rowe writes, “[C]ritical discussion does not issue in a proposition (‘This is good’ or whatever) but in an experience which may be described in a proposition.” (77) Or, as Stuart Hampshire puts essentially the same point:

53 John Wisdom, “Things and Persons” in Philosophy and Psycho-Analysis (1964), 217-228 at 223. Stuart Hampshire expresses much the same view when he writes:

One engages in aesthetic discussion for the sake of what one might see on the way, and not for the sake of arriving at a conclusion . . . if one has been brought to see what there is to be seen in the object, the purpose of discussion is achieved.


54 Quite rightly, then, Causabon in George Eliot’s Middlemarch would rank as a paradigmatically bad critic given his utter lack of any personal interest in or engagement with the paintings of Michelangelo and Raphael that he discusses with Dorothea on their Roman honeymoon.
Perhaps . . . there are particular features of the particular object which
make it ugly or beautiful, and these can be pointed out, isolated, and
placed in a frame of attention; and it is the greatest service of the critic to
direct attention in this analytical way. But when attention is directed to
the particular features of the particular object, the point is to bring people
to see these features, and not simply to lead them to say: ‘That’s good’.

But we have the same goal when talking about aspects. As Rowe emphasizes, “A
discussion about aspects has not had its intended effect if you simply know that or believe
(perhaps because reliable people keep telling you) that this figure can be seen as a rabbit
as well as a duck; you have to see it as a rabbit.” (77) Likewise with the Taj Mahal
remark: it has not had its intended effect if the woman trying on the hat fails to see it
differently herself. For instance, she might simply decide against buying the hat based on
her friend’s haughty, disapproving tone and a desire not to go against his judgment.
Accordingly, this kind of reasoning – the kind that prompts us to see aspects with helpful,
persuasive and/or creative descriptions – bears a striking resemblance to critical
reasoning. For criticism – at least good criticism – also fails to achieve its goal if it
produces merely second-hand knowledge or beliefs about a work. As we have seen, it
aims instead for the reader to appreciate the work, its features, and their combined
significance for herself, ultimately exercising – and, in the course of doing so, developing
– her own (aesthetic) judgment. Given this fundamental parallel, aesthetic (or critical)
reasoning makes a great deal of sense when understood in terms of aspect perception, that
is, as aiming at – and, when successful, inducing – a sophisticated kind of aspect
perception.

Most importantly, the basic structural similarity between discussions of aspects
and critical reasoning – each “deals with immediate experience . . . and if it changes that

55 Hampshire, "Logic and Appreciation" at 166.
experience in the appropriate way then it has been successful” – reveals how, by drawing on aspect perception, critical reasoning can make good on its aim of getting others to share certain evaluative responses to a work. The idea is straightforward. Just as when someone sees an object (say, U.S. immigration policy) as something else (say, an apartheid system) she has both her understanding of and her attitude towards the object simultaneously altered, so it is with aesthetic reasoning (at least when successful). Recall e.e. cummings’ poem ‘ygUDuh’. Given the right prompting (e.g. ‘The opening line can be paraphrased as “You’ve got to”’), we gradually come to see the poem under a new aspect. In doing so, we gain not only an understanding of what various lines mean but also new attitudes (e.g. interest, amusement ) towards the poem. Or consider this example from Wittgenstein who describes:

[A]n experience with the 18th century poet Klopstock. I had found that the way to read him was to stress his metre abnormally. Klopstock put u – u (etc.) in front of his work. When I read his poems in this new way, I said: ‘Ah-ha, now I know why he did this’. What had happened? I had read this kind of stuff and been moderately bored, but when I read it in this particular way, intensely, I smiled, said: ‘This is grand’, etc.\(^{56}\)

Wittgenstein tells, in other words, how reading these poems with a different emphasis leads to perceiving them differently, experiencing them as grand (instead of as boring). Moreover, in coming to perceive Klopstock’s poetry under this new aspect, Wittgenstein comes to posses not only a new understanding of Klopstock’s style “now I see why he did this” – but also new attitudes such as enthusiasm towards his work.

10. Let us call the distinctive kind of reasoning that we have been exploring ‘perspicuous description’ since it tries, by using such descriptions, to assemble

perspicuous representations and prompt the “seeing” of new aspects. We must now inquire more directly whether perspicuous description can serve as an account of moral reasoning. Fortunately, an affirmative response can be defended without much difficulty. Recall that, in the course of developing our account of perspicuous description, we discussed two paradigm cases, a certain sort of legal reasoning and critical reasoning about the arts. Now consider what these two recognized instances of perspicuous description have in common. First, with respect to the objects (or “data”) that each seeks to perspicuously represent, neither presupposes that there exists only one way of doing this. In other words, they each accept that it may be not only possible but perfectly reasonable to “see” their respective objects under more than one aspect. In the case of critical reasoning, this should be unexceptional. For it is a commonplace that great works such as Shakespeare’s Hamlet can be understood in multiple ways. (Indeed, this richness and depth of meaning is arguably a large part of what makes a work great.) Even with the legal paradigm, however, it is well-accepted that the facts in some law cases can be reasonably characterized in different ways – as exhibiting, for example, either negligence or due care – depending upon which facts are emphasized and which are explained away or otherwise minimized.

Both of these paradigms of perspicuous description share a second important feature. While each happily concedes that neither their respective objects nor the nature of perspicuous description requires that there must be only one correct way of perspicuously representing an object – this was the first thing they had in common – it would misconceive these paradigms to think that just any representation of a work of art or the facts in a legal case, respectively, could count as perspicuous. This is the second
thing they have in common: that what makes a representation perspicuous is not arbitrary but rather subject to certain rational constraints.

Take the legal paradigm first this time. The uncontroverted facts in a law case rationally constrain how it can be represented. Indeed, in some cases (colloquially referred to as ‘open and shut’), the facts may admit only one reasonable characterization. In other cases, the facts fail to determine one characterization as uniquely reasonable. Yet even in cases that admit competing reasonable characterizations – e.g. “whether Mr. A who admittedly handed his long-trusted clerk signed blank cheques did or did not exercise reasonable care” – the facts rationally constrain how such characterizations are put forward and defended. That is, to perspicuously represent Mr. A’s actions in either of these conflicting ways, not just any random characterizations of the facts will work. Rather, only descriptions that highlight certain key facts are apposite. For instance, to prompt a jury to see Mr. A’s actions as exercising due care, describing the color of Mr. A’s hair and eyes is a non sequitur. Instead, one needs to highlight such things as the many years that Mr. A’s clerk had worked for him, the special relationship that existed between Mr. A and his clerk, and the countless other times that Mr. A had entrusted his clerk with blank cheques without incident.

Roughly the same point holds of aesthetic criticism. That is, the object of critical attention rationally constrains the different ways in which it can be seen. Of course, given what a work of art or literature is – something specifically created to possess a certain richness and depth of meaning that can support a variety of interpretations – the rational constraints here are significantly looser. But they do exist. Thus, Wagner’s opera *Tristan and Isolde* can be understood in many different ways (or under many
different aspects). However, interpreting Wagner’s opera as certain things (as an automotive repair manual, for example) goes against all reason and is not, therefore, beyond all reason.


Lance, Mark and Margaret Olivia Little "Particularism and Anti-Theory (ms.)."


