

THE UNIVERSITY OF CHICAGO

VALUE REALISM AND THE FIRST-PERSON PERSPECTIVE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE HUMANITIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHILOSOPHY

BY

CLAIRE ELIZABETH KIRWIN

CHICAGO, ILLINOIS

JUNE 2019

For my mother.

TABLE OF CONTENTS

<i>Acknowledgements</i>	vii
CHAPTER 1: ARGUING FOR VALUE REALISM	1
1. Introduction.....	1
2. ‘Value’ and ‘real’	3
2.1 ‘Value’	3
2.2 ‘Real’	4
3. Characterizing realism and anti-realism.....	7
3.1 My proposal.....	8
3.2 Realism	12
3.3 Error theory.....	12
3.4 Expressivism.....	13
3.5 Neo-Humeanism and neo-Kantianism	16
3.6 Conclusion.....	20
4. How I shall argue for realism.....	21
4.1 Coming to see something.....	21
4.2 The argument	23
4.3 Realism, and nothing less	26
4.4 Autobiography	28
CHAPTER 2: DECIDING WHAT TO DO	29
1. Introduction.....	29
2. What do you do when you’re deciding what to do?.....	31
3. The role of (real) value in deciding what to do	38
3.1 ‘Getting something right’.....	38

3.2 An anti-realist rejoinder	43
3.3 Making (the right kind of) space for desire in deliberation	49
4. Motivating an investigation of ‘the first-person perspective’	57
5. Conclusion and recap	62
CHAPTER 3: ALIENATION AND TRANSPARENCY	65
1. Introduction.....	65
2. Two accounts of the first-person	66
3. Making sense of alienation	72
3.1 First-person, third-person, alienation	72
3.2 Two kinds of alienation	77
4. Further defense of transparency.....	81
4.1 Transparency itself as a form of alienation?	81
4.2 Interpreting transparency	85
4.3 A case study: alienated desires	92
4.4 The possibilities of vision / Loving the Alien	95
5. Re-unifying the first-person perspective	97
6. Conclusion	101
CHAPTER 4: THE TRANSPARENT FIRST PERSON AND THE VALUE-LADEN WORLD	104
1. Introduction.....	104
2. Tackling tasks from Chapter 2.....	105
3. Desire and deliberation	109
3.1 Deciding whether to go to the party.....	109
3.2 The alienated case.....	112
3.3 One more complication.....	115

4. On what comes first.....	120
4.1 Further generalizing the argument.....	120
4.2 ‘Directions of fit’.....	121
4.3 ‘Pulling oneself up by the bootstraps’	123
4.4 ‘The authority of reflection’	128
4.5 Conclusion.....	133
5. A puzzle remains.....	136
CHAPTER 5: THE IDIOSYNCRASY OF THE PRACTICAL	138
1. Introduction.....	138
2. My proposal.....	142
2.1 Pluralism and expertise	142
2.2 Expertise and what a person sees: a story.....	143
2.3 The case of value	147
2.4 The basis of expertise.....	149
2.5 A case study: loving one’s children	151
3. The bullet I am biting.....	153
3.1 Taste and preference as value-expertise	153
3.2 Clarification of the current state of play	155
3.3 Defusing explanations for the seemingly implausibility of the view	156
3.4 The phenomenology of coming to ‘get’ something.....	163
4. Disagreement and values in tension	165
4.1 A case study: opera and socialism	165
4.2 Must one of us be wrong?.....	169
4.3 A new approach to the problem of moral disagreement.....	171
5. Conclusion	174

CONCLUSION.....	179
1. What do you do when you're deciding what to do? (Redux).....	179
2. On what I have shown, or not shown.....	182
3. Afterword: both seeing and doing	184
BIBLIOGRAPHY	185

ACKNOWLEDGEMENTS

It turns out that writing a dissertation is really, really hard. I could not have done it without the help, love, support, and solidarity that I have been lucky enough to receive from so many people.

First, my wonderful committee—Agnes Callard, David Finkelstein, Brian Leiter, and Candace Vogler. I think it is fair to say that in terms of your philosophical temperaments and commitments, you constitute quite a motley crew. For me, this has been a wonderful experience. It is within the space of this mutually respectful difference that I have been able to develop something that feels to me like the beginnings of a philosophical voice and vision that is genuinely my own. Thank you to all of you for believing in me and in this project.

Throughout the process, I have depended in countless ways upon my fellow graduate students and other members of the UChicago community. Within my own cohort, so much love and gratitude is owed to Amichai Amit, Nethanel Lipshitz, Rory O’Connell, and Francey Russell, who have been my greatest philosophical companions and also just dear friends since we first met eight years ago. Within the philosophy department more broadly, I would like to thank in particular the community of women graduate students. Parts of my time here have been difficult and painful, and I could not have made it through those times without the support of these strong and smart and inspiring women. Beyond the philosophy department, I have been energized and uplifted by the solidarity and support of my colleagues and comrades who work with me at the Master of Arts Program in the Humanities (MAPH), as well as my fellow graduate workers in Graduate Students United (GSU). As I write this, GSU is in the process of a strike authorization vote; it is my sincere hope that by the time anyone is in a position to read this, the University will have agreed to recognize and bargain with our democratically-endorsed union.

To those who have looked after my soul: First, my mother. I know that my moving across an ocean to spend years doing this strange incomprehensible work is not exactly the life that you originally envisaged for me, but you have always supported and nurtured me in becoming the person that I am, and I could not have done this without you. Although we do not ever really talk ‘philosophy’ as such, still your influence on this project has been immense. You are the person who taught me to notice the beauty of plants and wildlife, and the importance of good friends and not taking life too seriously, and these forms of initiation into a rich world of value are what provided the template for the form of argument that I have attempted to give in this dissertation. Next, my sister, Angela. To say that you are an inspiration would be an wild understatement. You are a terrifying force of nature, and I love you more than I can say. In addition to my mother and sister, I have been lucky enough to have an honorary mum, Cath, and two honorary sisters, Joanne and Lynsey. You guys have been there through everything, and you are my home. Likewise, those friends who are closer to me than my own heart: Rey Conquer, Alec Hinshelwood, and Nakul Krishna. Last but not least, I am immensely grateful to Robyn Rihanna Fenty, whose music provided the soundtrack that kept me sane through the final push.

Finally, my partner, Pascal Brixel. It goes without saying that I have you to thank as the person who helped me to think through all of my philosophical ideas, and who helped to restore my confidence in my own abilities during the numerous times that it got very wobbly. But I would also like to say thank you for all the work that you did that has traditionally been ‘unspoken’ work: for all

the times that you cooked for me, and all the times that you washed the dishes and did the laundry. The history of philosophy does not feature very many women, and at least one major reason for this is that most people find it impossible to do the kind of work that this sort of endeavor requires while being at the same time responsible without respite from the demands of domesticity. So thank you for every mundane piece of work that you did so that I could calm down enough and concentrate enough to actually write something.

CHAPTER 1: ARGUING FOR VALUE REALISM

1. INTRODUCTION

My central claim in this dissertation is that value is *real*, in the metaethicist's sense: it is, as it is variously put, 'out there', 'discovered rather than created', 'part of the furniture of the world' or of 'the fabric of the universe'. It has often been remarked, by realists and their opponents alike, that one does not often see much in the way of a positive argument offered for such a position:

As many critics are quick to note, much of the work by advocates of [...] strongly realist views consists either in criticizing alternative views or in responding to traditional objections to realism. There is very little by way of positive argument for Robust Realism (or related views) in the current literature. (Enoch 2011: 10)

[M]oral realists spend a good deal more time defending moral realism against objections than they do arguing more directly in its favor. (Loeb 2003: 35, fn. 6)

Realists try to establish the normativity of ethics by arguing that values or obligations or reasons really exist, or, more commonly, by arguing against the various forms of skepticism about them. (Korsgaard 1996: 19)

It is very difficult to argue for [the possibility of realism], except by refuting arguments against it. (Nagel 1986: 143)

In defense of her position, then, the realist may attempt to respond to the traditional list of objections to realism: value is too 'queer' to be acceptable within a naturalistically plausible ontology;¹ value is (causally) explanatorily redundant in our best accounts of the world;² there is widespread disagreement on questions of value to an extent that undermines the claim that our

¹ (Mackie 1990).

² (Harman 1977), (Harman and Jarvis Thomson 1996), (Wright 1992: ch. 5).

opinions are responses to something real, and these disagreements seem in addition to be intractable in a distinctive way;³ it is unclear how we are to understand the so-called ‘supervenience’ of evaluative facts on non-evaluative facts;⁴ and even if there *were* such a thing as value ‘out there’, the realist cannot explain how such a thing could have the relevant motivational power,⁵ or the relevant normative authority⁶ over us. In addition, the realist may seek to defend her position by joining the general fray in arguing against individual anti-realist alternatives, attempting in this way to pick off her opponents one-by-one.

In this project I am going to attempt to give a genuinely positive argument for realism. At the same time I am going to try to show what I think is wrong with anti-realism, not so much in terms of any of its particular iterations (although I will have things to say about these too), but rather anti-realism as a general metaethical approach—anti-realism as such. This strategy means that I face the following two problems. First, it may be that—as Nagel seems to be suggesting in the above quotation—there is something about realism itself as a position that makes it especially difficult to argue for positively (and lest one is inclined to reply, snarkily, that this ‘something’ may be the fact that it is false: bear in mind that a glance at the history of philosophy provides us with no shortage of positive arguments for false positions). Second, the sheer variety of anti-realist positions makes it difficult to see how one might even begin in arguing, as I shall be claiming to do, against ‘anti-realism’ as a whole. My main goal in this opening chapter is to show how I plan to deal with these two problems. First, however, I shall spend some time making some clarificatory remarks about the nature of the position I am defending here.

³ (Mackie 1990), (Ayer 1957), (Stevenson 1963).

⁴ (Blackburn 1993b), (Blackburn 1993c), (Gibbard 2003: ch. 5).

⁵ The original source for this worry is of course Hume; for a more recent reformulation, see (Smith 1994).

⁶ (Korsgaard 1996), who takes herself to be following Kant.

2. 'VALUE' AND 'REAL'

2.1 'VALUE'

In defending the reality of what I am calling 'value', I am entering a metaethical debate that has, both in its historical and more contemporary forms, made use of a wide range of terms to name and talk about the thing under dispute. As Sharon Street explains while introducing her own (anti-realist) position:

I employ the term *reason* (always in the sense of a *normative* reason) as a catch-all normative term [...]. Nothing substantive hinges on this choice of language; my points could also be couched in the language of *value*, *should*, *ought*, *goodness*, what *makes sense*, what's *rational*, *worthwhile*, and so on. (Street 2008: 209, fn. 5)

Compare (Enoch 2011: 93, fn. 5); similar remarks are common enough throughout the metaethical literature. It is not that such authors deny that there mightn't, in various contexts, be important differences between these terms. Rather, the thought is that there is nonetheless some vaguely unitary phenomenon that these various terms are intended to pick out or gesture towards, and which is the object of our shared philosophical conversations and arguments, at least to the extent that these remain within a strictly metaethical register. For my part, I shall use the term 'value' for these purposes, although in my Conclusion I shall offer some considerations in favor of keeping hold of certain distinctions between these terms, even within the context of purely metaethical concerns.

My own account aims to remain strictly metaethical in the sense that I do not mean here to be committing myself to any claims about what sorts of things are actually valuable. In the course of making and illustrating my case, I shall often have certain instances or examples of value in mind, and I shall draw on these at various points, but my argument does not require that the reader accept

any particular evaluative judgment implied by my uses of these examples. And in particular, in defending the reality of value in general, I do not take myself to have established that *moral* value—which we may wish to delineate as a set of claims with certain particular content or formal features—is real or binding. I shall not, that is, be claiming that there is one universal set of rules or demands or principles that applies to everyone regardless of who they are or what their situation is, nor shall I be defending any particular moral principles. Indeed, my concern here is not primarily with ‘morality’ in this sense at all. Instead, in speaking of value, I have in mind a wide-ranging and multifarious phenomenon, potentially encompassing such things as: aesthetic and artistic value, the value of friendship and of certain traits of character, the value of pleasure, of intellectual pursuits and knowledge, of a noble action or a good joke.

2.2 ‘REAL’

In my initial statement of this dissertation’s central claim—the reality of value—I drew on the various tropes of the metaethics literature to indicate what was meant by this notion of reality: value is ‘part of the fabric of the universe’, ‘out there’ (as opposed, presumably, to ‘in here’, a phrase which is to be accompanied by a gesture towards one’s own head), and so on. In section 3 of this chapter I shall argue for a particular way of understanding what is really at issue in the debate between realist and anti-realist. But before doing so, I must clarify a more basic point.

My goal in this project is to address an interlocutor who is concerned about the reality of *value* in particular. Philosophers are, of course, capable of raising doubts about the metaphysical status of all sorts of things. But here, for the purposes of this debate, the thought is that there is or might be something *especially* or distinctively worrying about value. The question about value’s reality is thus a question about the reality of value *as compared to various other things*—things that are accepted, by both my interlocutors and myself, as counting as ‘part of the furniture of the world’ in the

relevant sense. These include such things as the findings and objects of the natural sciences, but also, in a more mundane way, such things as actual furniture (Don Loeb counts “chairs and houses” among the objects whose reality does not worry us in the way that the reality of value, as he thinks, ought to worry us (2003: 36)), as well as rocks, volcanoes, houseplants, and indeed us ourselves, as well as the qualities and features that make all these things the sorts of things they are (a blue-painted wooden chair; a happy Chicagoan). In this way, my argument about the reality of value is presented within the context of what we might think of as a sort of naïve realism—or at least something that it’s okay to talk about in ways that suggest such a picture—about a whole range of other familiar parts of our lives.⁷ Once again, such an approach is common enough within the literature with which I am in conversation. There will of course be anti-realists about value whose anti-realism in this arena really stems from a more fundamental set of doubts that stretches well beyond the case of value, but my argument here is not primarily pitched towards such people.

For some of my opponents, to reject the reality of value is to place it in roughly the same class as things such as witches, ghosts, Santa Claus, unicorns, and so on. So: chairs and tables are real; phlogiston and unicorns are not—and value, according to such authors, is to be placed on the ‘phlogiston and unicorns’ side of the divide. David Gauthier, for example, having been persuaded by Harman’s arguments against realism, concludes that “[o]bjective value, like phlogiston, is an unnecessary part of our explanatory apparatus, and as such is to be shaved from the face of the

⁷ There is one particular case of something agreed by both sides to be ‘real’ that will turn out to be crucial to my argument, and that is the case of psychological attitudes (beliefs, including evaluative beliefs, desires, hopes, intentions, and so on). As I shall show in section 3 of this chapter, all of my opponents here take for granted that the reality of such psychological attitudes is *not* being called into question in the way that we are calling into question value’s reality.

universe by Ockham's razor" (Gauthier 1986: 56).⁸ Loeb, bristling at the implied suggestion that his rejection of realism might mean that he is committed to a form of 'relativism'—according to which "the ultimate *validity* of moral judgments" is determined by "human tendencies, cultural values, contextual factors, historical forces, hereditary dispositions, and/or sentiment,"⁹—is at pains to point out that, in fact, on his view "moral judgments *have no* ultimate validity, as there are no moral facts or correct answers to our moral questions" (Loeb 2003: 31, fn.1). Richard Joyce chooses to title the book in which he expounds his own anti-realist position *The Myth of Morality* (2001). And Mackie, as part of his 'argument from queerness', asserts that objective values would have to look something like Plato's Forms—a claim about which he appears to feel mildly apologetic, since it seems to him to involve accusing his opponents of uncritically accepting the existence of something (i.e., value) which is akin to "what are admittedly among the wilder products of philosophical fancy" (i.e., Plato's Forms) (Mackie 1990: 41). And Mackie seems to be worried that this is perhaps not a very polite thing to do, even to one's philosophical opponents.¹⁰

While these authors' opposition to value's reality is intended to relegate it to roughly the same category as such things as phlogiston and unicorns, to present my opponents as a whole to be taking such an attitude towards value would, however, be grossly misleading. For example, for Korsgaard, the whole point in rejecting realism about value is precisely in order to *save* the idea that morality is binding on us. In a vaguely similar sort of vein, various authors who claim to trace their

⁸ And a little later: "To be sure, inference to the best explanation will not show that there is no objective value, any more than it will show that there are no fairies at the bottom of the garden. We are content to put objective value on a par with the fairies" (Gauthier 1986: 58–59).

⁹ (Robinson 2002: ix–xi), quoted in (Loeb 2003: 30, fn.1).

¹⁰ Luckily for me, I share David Enoch's sentiment here: "I will not be offended if you call me a Platonist" (Enoch 2011: 8).

philosophical lineage back to Hume¹¹ are keen to establish that people do, in fact, have such things as *reasons* to act or to refrain from acting in all sorts of ways; the point of these kinds of positions is largely to spell out a positive account of where such reasons come from and how they work. And finally, expressivism (perhaps most especially in its modern ‘quasi-realist’ incarnation) aims to give an account of the semantics of value-talk that will allow us to salvage the validity of value-talk generally.

Now we may wonder, regarding these latter more positive or optimistic-sounding versions of anti-realism, whether it might nonetheless be correct to say that still they hold (or must, ultimately, hold) that, at the very least, value *understood as the realist understands it* really would have to be counted along with (and relegated along with) unicorns and phlogiston. (I’ll discuss a version of this thought in relation to expressivism and quasi-realism below.) But in any case, what is clear is that my opponents’ shared commitment to the claim that value is ‘not real’ gives rise to a variety of different attitudes towards the concept of value, and a variety of different positions on the sort of role that it might yet properly play within our philosophical thought and within our lives. In the following section, then, I will discuss some of the differences that hold between my various opponents, and also do some work to try to reveal a central underlying unity.

3. CHARACTERIZING REALISM AND ANTI-REALISM

I have said that one of the things I hope to do here is to argue against anti-realism as a *general* position. My strategy against the various forms of anti-realism is not to be one of divide-and-conquer, but rather by way of one (more or less) unified argument. But given the sheer variety of

¹¹ I am thinking primarily of Mark Schroeder (2007) and Sharon Street (e.g. (2008), (2010)).

positions than can be gathered under the broad heading of anti-realism¹² it is difficult to see how this might be done. In this section, I lay the groundwork for such a unified argument by making a case for a certain way of understanding the central distinction between realist and anti-realist, and then by detailing the application of this account to various different metaethical positions, and attempting to deal with some complications that arise along the way.

3.1 MY PROPOSAL

My proposal is that we can usefully understand the distinction between realist and anti-realist by thinking about the relationship that a given metaethical position takes to hold between what I shall refer to as a person's 'psychological attitudes' (or 'attitudes'), on the one hand, and what we are to say about value, on the other. In particular, we can characterize the distinction between realist and anti-realist in terms of the order of priority that a given position takes to hold between our attitudes and what I'll call 'the arena of value'. This terminology of an order of priority is left, at this stage, deliberately vague, for different forms of anti-realism will wish to fill out the idea in different ways (as I shall explore below). And likewise, my phrase 'the arena of value' is intended to be understood (at least for now) in an entirely metaphysically non-committal sense, since I am using it to cover not only the more metaphysically robust metaethical accounts but also, for example, the expressivist's attempt to give legitimacy to evaluative language without having to commit to anything problematically metaphysical.

Let me begin with the concept of psychological attitudes. In section 2.2 above, I noted that the disagreement between myself and my opponents concerns the reality of value *in particular*—of

¹² The terminology used to describe opponents of value realism varies, with some authors wanting to speak in terms of (or to distinguish between) 'anti-realism', 'non-realism', 'irrealism', and so on. Here I shall simply refer to all opponents of realism as 'anti-realists', and in this section I shall spell out how I shall be characterizing this opposition.

value, that is, as opposed to a wide range of things (atoms, chairs, rocks, people, the Supreme Court) the reality of which is *not*, at least for the purposes of this particular metaethical debate, being called into question. And one especially interesting part of the world that realist and anti-realist alike tend to be happy thinking of as ‘real’ for these purposes are people’s psychological attitudes. Both realist and anti-realist typically do not question the idea that people, as a matter of psychological fact, do go about believing, doubting, hoping for, planning, desiring, intending, fearing, and indeed even valuing things. It is widely accepted that—considered merely as a matter of descriptive fact about the psychology of persons—such attitudes, including those attitudes which are explicitly evaluatively-inflected, are *not* to be treated as a cause for concern in the way that value *is* a cause for concern.

Plenty of authors are explicit in their endorsement of this idea. T. M. Scanlon, for instance, in the course of explaining why his notion of ‘taking something to be a reason’ ought to be entirely acceptable from the point of view of his anti-realist interlocutors, asserts that:

The state of taking something to be a reason, as I interpret it, is a purely psychological state, just as ‘naturalistic’ as the state of adopting a plan, accepting a norm, or feeling approval [here Scanlon is referring to attitudes that form the basis of various anti-realist proposals]. Even a belief in witches is a naturalistic psychological state in this sense. (Scanlon 2014: 62)

Likewise, Loeb, in the course of exploring what it would be to be a realist about gastronomic value,¹³ explains that:

To get [gastronomic value properties] sharply into focus, however, it is necessary to distinguish them from other, less controversial properties in the neighborhood. [...] It is a perfectly ordinary psychological fact that I don’t like the taste of green peppers, for example. Likewise, there are perfectly ordinary *social* facts about these matters of taste—facts about broad patterns of intersubjective response to certain types of foods. Presumably, for

¹³ His purported aim in undertaking this exploration is to reveal the analogies between gastronomic realism and moral realism in order to undermine moral realism. However, by the end of the paper, he appears to have gone some way towards persuading himself of the plausibility of gastronomic realism (Loeb 2003: 47). See my Chapter 5 for more on this topic.

example, “everyone likes ice cream,” is a fairly accurate generalization. The question is not whether facts like these exist, for surely they do. (Loeb 2003: 32)

And similarly, Street explains that her own metaethical position rests on the idea that we can get a grip on what is involved in “the attitude of taking or judging something to be a reason” even if we don’t yet know what reasons or values themselves are, and concludes that we can derive “a naturalistically acceptable understanding of reasons as ‘constructed out of’ or ‘legislated by’ such judgments” (Street 2008: 241).

Now from a certain philosophical perspective, such seemingly blasé embrace of the ‘naturalistic acceptability’ of these sorts of attitudes will strike us as jarring. For there is a familiar tradition of deep concern about the possibility of fitting concepts of mindedness and intentionality into a genuinely naturalistic ontology.¹⁴ And yet, when we look at the detail of anti-realist positions, what we find is a commitment—whether implicit or explicit—to the idea that at least people’s evaluative and practical attitudes are *not* to be called into question in the way that value *is* to be called into question. Indeed, as far as I can tell, every major example of an anti-realist position in metaethics relies on such an assumption in the very spelling out of their position: the error theorist assumes that we do in fact have beliefs about value, for it is just these beliefs that she holds to be false; the expressivist holds that we have a range of non-cognitive attitudes that we express in our moral talk; the neo-Humean holds, depending on the variant, that we desire things or make normative judgments, for these attitudes are what explain the reasons that we have; and the neo-Kantian accepts all the intentional psychological goings on that the neo-Humean does, and perhaps also some extras.

¹⁴ See for instance (Quine 1960: ch. 6) and (Churchland 1989).

If we suppose, as seems reasonable, that at least some anti-realists have at least some sympathy with the concern about the naturalistic acceptability of mindedness and intentionality, how are we to understand the lack of anxiety these authors show about making use of such attitudes within their metaethical theories? It seems to me that my remark in section 2.2 is relevant here: at the heart of the metaethical debate between realists and anti-realists is the worry that, amongst all the entities about which philosophers can properly raise doubts, there might be something *especially* or *distinctively* worrisome about value in particular. Thus, although we may in other contexts want to raise concerns about psychological attitudes, in the context of the metaethical debate we will agree to leave these doubts aside. For however philosophically worrisome these concepts might be in themselves, they are perhaps less worrisome than value. Roughly speaking, the idea is that psychological attitudes, though they might turn out not to be *really* real, are nonetheless *more* real—according to the metaethical anti-realist—than value. Or such is my surmise, based on the use that anti-realists wish to make of such attitudes.¹⁵

My account of the distinction between realist and anti-realist, then, is based on the order of priority (where what ‘priority’ means will be filled out in more detail in the following sections) that each takes to hold between, on the one hand, psychological attitudes of various kinds, and on the other, ‘the arena of value’, whatever that shall amount to for the position in question. In broadest outline: what distinguishes realist from anti-realist is that the former takes the arena of value (which, for the realist, is simply evaluative reality) to ‘come first’, while the latter reverses this ordering and takes attitudes to come first. Anti-realists, then, are united in an implicit commitment to the idea that we are able to get a philosophical grip on human attitudes *prior* to our understanding of value, and

¹⁵ See (Wedgwood 2007) for an interesting defense of normative realism on the basis of the normativity of the intentional. My approach here is rather different from Wedgwood’s, but there are nonetheless some points of commonality.

that whatever account we are to give of value will be given in some sense in terms of these attitudes. I'll now give a little more detail about how this basic idea is cashed out in various different instances of anti-realism, and discuss some particular points of interest that arise in relation to some of these positions.

3.2 REALISM

My claim, as a realist, is that evaluative reality comes first, and that our attitudes are answerable to it in various ways: our evaluative beliefs aim to represent that reality accurately, and thus may be straightforwardly true or false, while other less overtly cognitive attitudes (desires, fears, intentions, attitudes of approval or disapproval) may be rendered appropriate or inappropriate relative to the demands of this realm. In this sense, then, realism gives a clear priority to the arena of value over human attitudes. For the realist, value comes first, and we are answerable to it.

As I am conceiving the conversation with which I am concerned, the value realist need not deny that there remains a set of pressing *general* problems about how we are to understand the relationship between the mind and the world, in particular in the wake of Cartesian and Humean skepticism (and the special insight into these problems—if not, perhaps, the solution to them—offered by Kant in the First Critique). My point as a realist about value is to deny that there is some special, or especially pressing, form of this problem that arises in the case of value in particular. It is in *this* sense that the realist claims that value has ‘priority’ over our attitudes: namely, it has *whatever* sort of priority reality in general will turn out to have over our attitudes.

3.3 ERROR THEORY

The error theorist’s account of the relationship between attitudes and ‘the arena of value’ is similarly straightforward. Indeed, it echoes nicely Scanlon’s example of someone’s believing in witches. It’s

certainly true, according to the error theorist, that people do come equipped with a wide range of attitudes, including indeed beliefs about things being valuable. The problem is, in short, that there is no such thing as value, just as there is no such thing as witches, and so those beliefs are without exception false. The example of the error theorist is thus a sort of limit-case of giving priority to attitudes over value in her account, for according to her, the attitudes exist but the value doesn't. The error theorist begins from the arena of attitudes, and that is also where she ends; we may believe all kinds of things are valuable, but we're wrong. All we're left with are the beliefs.

3.4 EXPRESSIVISM

Expressivism occupies an interesting position within the taxonomy of metaethical positions that I am developing here. It is first and foremost a semantic theory, not a metaphysical one, and this complicates things somewhat. Still, I shall try to explain why I believe my analysis of the distinction between realist and anti-realist nonetheless applies in this case, and thus how my arguments in the dissertation as a whole can be seen to be relevant to expressivists (including in the more recent 'quasi-realist' iterations of the position).

The central expressivist move is to avoid both realism and error theory by offering a revised semantics for value-talk. While the surface grammar of a statement such as 'x is bad' seems to suggest that the speaker is attributing a property to a thing (badness to x), the expressivist suggests that we need not accept this surface grammar. Early forms of expressivism¹⁶ aimed to replace this picture with the claim that value-talk involves instead the expression of simple non-cognitivist attitudes (of disapproval, say, or disapproval combined with a demand made of the listener that they, too, disapprove). So according to such an approach, the statement 'x is bad' may be understood as akin to something along the lines of 'x? boo, hiss!' Since such a statement no longer seems to be

¹⁶ (Ayer 1957), (Stevenson 1963), (Hare 1952).

representing anything about how things stand in the world, there is no longer any need to posit some value-laden reality that is being represented, and likewise since the statement is no longer truth-apt the error-theoretic threat has been avoided.

Since its ‘boo, hiss!’ beginnings, expressivism has become ever more complex and sophisticated. Two developments are particularly worthy of note. First, there have been attempts to construe the relevant non-cognitive attitudes in ways that permit of the sort of logical connections and inferences that will allow the expressivist to solve the so-called ‘Frege-Geach problem’. And second, there has been a move towards embracing deflationary or minimalist accounts of truth that are supposed to allow the expressivist to attribute truth to value-statements while still avoiding the metaphysical extravagancies of realism. In this vein, so-called ‘quasi-realist’ accounts¹⁷ claim to be able to deliver all the goods of realism while maintaining fully naturalistic credentials. These claims have been attacked from a number of different directions—and including authors on both the realist and anti-realist sides of the aisle. I do not wish to re-invent any of these multitudes of wheels, and so I will simply try to explain briefly what I think matters about expressivism for my purposes.

Although expressivism is pitched as first and foremost a semantic account, nonetheless it does come with certain metaphysical commitments. These commitments are, however, primarily negative ones. The initial motivation for the view is precisely to avoid both realism and error theory; the thought is that a realist metaphysics is untenable, and so a cognitivist semantics will inevitably lead to error theory. It is a central component of the expressivist’s position—even the quasi-realist position—that she rejects the realist’s metaphysical account of value. Further, this means that she must be willing to allow for a certain mode of value-talk which is *not* given a strictly expressivist reading—namely, the metaethical mode of discussion in which the fundamental difference between

¹⁷ (Blackburn 1984: ch. 6), (Blackburn 1993a), (Blackburn 1998); (Gibbard 1990), (Gibbard 2003).

‘realism’ and ‘quasi-realism’ can be stated.¹⁸ It is within this mode of conversation that our disagreements about the metaphysics of value can be laid out, and here we will see that expressivism is still fundamentally an account which gives priority to attitudes—the attitudes which are expressed in value-talk—over the arena of value. In this case, the priority amounts to the fact that in the metaethical story the expressivist tells, it is a person’s attitudes which determine what evaluative claims she will express—and this, then, is the *end* of the story about value. Humans possess certain attitudes, which are expressed, and this expression is what gives us the entire realm of (talk about) value.

To what extent can this characterization of the priority of attitudes be avoided by a flight into truth-minimalism on the part of the expressivist? For instance, perhaps she can claim that quasi-realism is really what we might call a *no-priority* view: since quasi-realists (and other truth-minimalist expressivists) can claim truth for their normative statements, the case might be made that really attitudes and value (the *truth* about value, she can add) come on to the scene at the same time.

As I see it, the question here comes down once again to whether or not what we are really worried about is *value* in particular. I take it that a defining characteristic of *metaethical* expressivism (including quasi-realism) is that it is value that is singled out for the expressivist treatment. Part of the point of expressivism, to the extent that it is really a metaethical thesis, must be to allow that *some* talk is *not* (merely) expression. Once again, the metaphysical commitment reveals itself, this time in the form of a metaphysical contrast: some forms of language are accepted as fully representational, because there *is* a world that can be represented in language. It is just that value is not a part of that world. To the extent that the expressivist wishes to hold onto this sort of contrast, then, she also

¹⁸ For some discussion of the complaint that quasi-realism cannot adequately distinguish itself from realism, and thus does not count as a properly separate metaethical position, see for instance (Dworkin 1996), (Drier 2004).

holds an attitudes-first view when it comes to value. In the realm of value-talk, attitudes have priority in a sense that they do not in the realm of representational talk. And this is enough for expressivism to be adequately characterized as anti-realist under the account that I have proposed.

3.5 NEO-HUMEANISM AND NEO-KANTIANISM

I shall have quite a bit more to say about some of the major proponents of neo-Humean and neo-Kantian views in the following chapters, and so for now I shall try to be brief.¹⁹ For the neo-Humean, the source of value is typically understood to reside in desires or desire-like attitudes: some things are (let's say) good-to-me or to-be-done-for-me because I happen to possess certain desires that render them such. For the neo-Kantian, value's dependence on the mind takes on a more rational, and thus (or so the neo-Kantian hopes) more universal structure: regardless of the particular idiosyncratic desires I may have, I am nonetheless rationally bound to make certain normative judgments, and my own commitment to these judgments is what renders the objects of these judgments valuable to me, and likewise for each rational agent. Thus we may say of the neo-Humean and neo-Kantian that although they agree with the realist that some things are, as a matter of fact, 'valuable', there is an important sense in which they do not conceive of this value as residing, in the final instance, 'out there' in the world in the way that the realist does; rather, its source is 'in here', in something in me. As Korsgaard puts it, "obligation derives from the dictate of the agent's own mind" (Korsgaard 1996: 31), and then later "I believe that all values and reasons are human creations, and that the materials from which they are created are things like our own desires"

¹⁹ I shall have almost nothing to say here about the historical Hume or Kant. I believe it may be that within Kant's work we can find a more successful version of what I above called a 'no-priority' view. Occasionally, Korsgaard (my central example of a neo-Kantian) says things that suggest that she too might wish to be considered 'no-priority' in this way. However, I think that ultimately her commitments are too strongly anti-realist for this to be plausible as an overall reading of her position.

(Korsgaard 2009: 209); for Street, “[t]hings are valuable ultimately because we take them to be” (Street 2008: 207). Both of these positions, then, accord priority to attitudes over the arena of value.

There are interesting similarities in the ways that neo-Humeanism and neo-Kantianism have been developed as metaethical positions in recent years. According to one influential interpretation of the relationship between the two positions,²⁰ both begin from the idea that a person comes equipped with a particular set of values, desires, normative judgments, and so on, and both agree that a person’s reasons are given by what follows as a formal or logical matter from this collection of practical attitudes. The difference between the two views, according to this reading, concerns just how much of substance we can say about what *does* follow. In particular, the neo-Kantian claims that *any* such bundle of attitudes entails certain particular moral conclusions (the categorical imperative, the value of humanity as such). The neo-Humean is not so optimistic: she allows, to use Street’s favored example, that there could be an “ideally coherent Caligula” (eg. (Street 2010: 371)), and that such a character would not be subject to any moral requirements.

In both its neo-Humean and neo-Kantian forms, then, this sort of view gives priority to attitudes over the arena of value. The same priority commitment is true of Mark Schroeder’s neo-Humeanism, which I shall discuss in detail in Chapters 2 and 4. However, it is worth noting that while Street (who actually considers her view to be a Kantian-Humean hybrid) classifies herself as unapologetically anti-realist, Schroeder considers his view to be a form of reductive naturalist realism. I shall not go into his reasons for holding so (see chapter 4 of his (2007)), but it is clear that according to the taxonomy that I am laying out here, he counts as an anti-realist, and I shall speak of him as such throughout.

²⁰ (Street 2012: esp. 40–42).

Finally, let me make a brief remark on Bernard Williams. Both Street and Schroeder mention Williams's "Internal and External Reasons" as a precursor to their own views. And this paper is widely treated within today's metaethics literature as *the* seminal expression of a neo-Humean position. I think that actually it is rather unclear what Williams' metaethical position is. Certainly, he says a number of things that incline one to align him with Street and Schroeder, as they themselves would like to place him. But there are reasons to doubt this alignment, and some of those reasons have independent metaethical interest. In particular, it seems to me that there is a very important difference between the claim, on the one hand, that desires give rise to reasons²¹ *without* our needing to say anything more about, for instance, any source of goodness in the thing itself, and on the other, the claim that it is a *necessary* condition on a person's having a reason that she possess some relevant piece of psychological material that would allow this reason to motivate her. The latter claim allows (though it does not require) that it may be that our attitudes must work in concert with some source of goodness in the world in order to give rise to reasons. And I believe that there is *some* reason to attribute something more like this latter view to Williams. Consider, for instance, Williams's example of the man who is unkind to his wife (Williams 1995: 39). Williams has us suppose that the man does not care about the fact of his unkindness, that "there is nothing in his motivational set that gives him a reason to be nicer to his wife" (Williams 1995: 39). On Williams's account, then, the man *has no reason* to be nicer to his wife. So far, so neo-Humean. But notice what Williams says next:

There are many things I can say about or to this man: that he is ungrateful, inconsiderate, hard, sexist, nasty, selfish, brutal, and many other disadvantageous things. I shall presumably say, whatever else I say, that it would be better if he were nicer to her. (Williams 1995: 39);

²¹ More precisely: that they do so under certain formal constraints, such as constraints of coherence, or by passing a categorical-imperative-type test of the kind Korsgaard has in mind.

What Williams seems to be insisting upon is a point about *reasons* in particular—where this term is *not*, for Williams, standing in for the entire gamut of normative concepts in the way that it does for (say) Street and Schroeder. Williams wishes to say that this man does not *have a reason* to act otherwise, despite the fact that we may truly say a wide range of evaluatively-loaded things about him. Thus, when Talbot Brewer criticizes Williams here for failing to notice that all of these *other* claims (about the man’s sexism, or selfishness, etc) entail that the man *does* have the reasons Williams denies he has, I suspect he may be missing Williams’ point. Brewer says:

These are terms of condemnation, not dispassionate descriptions. To condemn a person in any of these terms is to imply that they ought to act differently, and this in turn in to imply that they have a reason to act differently. (Brewer 2002: 448, fn. 14)²²

I do not think that Williams means these terms (sexist, selfish, and so on) to be “dispassionate descriptions”. The issue, I think, is that more recent metaethicists have read Williams’s work through the lens, specifically, of our own metaethical concerns about the reality or otherwise of value. And while I do think that Williams shares some of these concerns, it is not clear to me that these are his main concerns at these parts of the text.

The point of this discussion is not really to defend Williams from the charge of neo-Humeanism. There are certainly plenty of things that Williams says throughout his work that suggest an anti-realist reading of his metaethics, such as it is. (And going forward, I’ll follow the general trend in the current metaethics literature of treating him as more or less a neo-Humean.) Instead, my point is to indicate that one may have a position on, for instance, the nature of practical rationality, that may be described as ‘Humean’ (or perhaps ‘instrumentalist’) without this entailing that one is an

²² To be clear, although I suspect Brewer’s charge against Williams in particular may miss its mark somewhat, I am actually sympathetic to his argument in this paper more generally.

anti-realist about value.²³ Since this point seems sometimes to be elided within the realist/anti-realist debate (I discuss a version of this issue in relation to Schroeder in Chapter 4, 3.1), it is worth clarifying here.

3.6 CONCLUSION

In these ways, then, these various examples of types of anti-realist positions can be understood as different instances of this same very broad-strokes picture of the relationship between human attitudes on the one hand and value (such as it is) on the other—a picture, that is, that sees our attitudes as coming first and (at least in those cases where value or value-talk is taken to have any legitimacy at all) as forming a basis of some sort for value. As I understand it, then, this is what most fundamentally marks the distinction between realist and anti-realist. And this point of commonality between the otherwise highly diverse anti-realist positions—their attitudes-first metaethical picture—is what will allow me to address my opponents with one unified line of argument in what follows. I believe that my characterization of the realist/anti-realist distinction does single out a point of dispute that is metaethically very important. But, to be clear, I am not claiming that this is the only acceptable way of marking the distinction between realist and anti-realist. And ultimately, this account is offered in a stipulative spirit: this is what *I* shall mean in the context of my arguments for realism and against anti-realism.

²³ Cf. eg. (Vogler 2002), (Svavarsdóttir 2008).

4. HOW I SHALL ARGUE FOR REALISM

4.1 COMING TO SEE SOMETHING

I opened this chapter by noting the widespread observation that not much has been offered in the way of positive arguments for metaethical realism, and raised the possibility that there might be something about realism itself that somehow makes it especially difficult to offer substantive arguments for the position. Why might this be? And what does it mean for the realist who, like me, would like to offer such an argument?

My route into thinking about the question came via the sustained critique of realism that Korsgaard offers at the beginning of *The Sources of Normativity*. A major part of her objection to realism is what she sees as the realist's inability, or perhaps refusal, to say anything useful to someone who is in a state of doubt about whether or not she actually has to comply with some (purported) moral demand. Korsgaard cites a number of realists²⁴ who appear to struggle to justify their own position because the reality of value seems to them self-evident. That may be, thinks Korsgaard, all well and good for those realists, but it is not much help to anyone else:

If someone finds that the bare fact that something is his duty does not move him to action, and asks what possible motive he has for doing it, it does not help to tell him that the fact that it is his duty just is the motive. That fact isn't motivating him just now, and therein lies his problem. In a similar way, if someone falls into doubt about whether obligations really exist, it doesn't help to say 'ah, but indeed they do. They are *real* things'. Just now he doesn't see it, and herein lies his problem.

I find Korsgaard's worry here extremely interesting, and I think there is something important to it. It is related to the vague worry that one sometimes hears regarding the (im)possibility of 'arguing with the skeptic'. And what Korsgaard hopes to do with her own account is to offer an argument that

²⁴ For instance, Nagel, whom Korsgaard quotes (a little out of context, as it turns out) as saying that "[i]n arguing for this claim, I am somewhat handicapped by the fact that I find it self-evident" ((Nagel 1986: 141), cited at (Korsgaard 1996: 41)).

ought, at least (whether in reality it does or not) to persuade someone who is skeptical or doubtful about the legitimacy of the demands that morality makes on us. In very rough terms, she aims to do so by showing that the person is ultimately being inconsistent or irrational if she does not assent to the legitimacy of these demands. And the fact that this is Korsgaard's strategy in *arguing* for her own metaethical picture is intimately linked to the neo-Kantian structure *of* that position. She thinks that it is only by appealing to things that a person is in some sense *already* committed to that one could possibly achieve the sort of rational persuasion that she is after. (I return to this aspect of Korsgaard's position in Chapter 4, section 4.4.)

Now, as a realist, I think that the source of the normative authority that value has over us is (and here I gesture vaguely beyond myself) 'out there'. It is this sort of commitment of mine that leads Korsgaard to make the following rather memorable assertion:

What is really wrong with substantive realism is its view about the source of normativity. Why do we use normative concepts like good, right, reason, obligation? According to the substantive realist it is because we grasp that there are things that have normative properties. [...according to the substantive realist, w]e have normative concepts because we've spotted some normative entities, as it were wafting by. (Korsgaard 1996: 44)

In the face of this accusation, it would be natural enough to accuse Korsgaard of straw-manning the realist. Indeed, plenty of self-described realists have been at pains to insist that their realism absolutely does *not* commit them to anything that might be described in Korsgaard's rather disparaging terms.²⁵

²⁵ See for instance (Scanlon 2014). Nagel also goes to great lengths to insist that whatever may be said of his realism, he is *not* any kind of Platonist ("What we aim to discover by this method is not a new aspect of the external world, called value [...]. It is important not to associate this form of realism with an inappropriate metaphysical picture: it is not a form of Platonism" (Nagel 1986: 139)). As a result, Korsgaard is a little kinder to him than to other realists, allowing that there is a reading of his position on which he is not really a *substantive* realist (the bad kind) at all (e.g. (Korsgaard 1996: 41, fn. 68)).

Suppose, however, that we try to take this sort of characterization of realism seriously. And since I am a realist, this will mean that I am supposing that there *are* some normative entities around the place, as it were wafting by.²⁶ But my skeptical reader, of course, does not agree. How, then, might I go about trying to change her mind? It seems to me that part of the task will be to get her to *see* something that she does not currently see, or perhaps more accurately to see that she *can* see something that she does not realize that she can see. And in fact something like this is going to be part of what I am going to try to do in this project. Alongside a set of other arguments, I will also be trying to help the reader to spot some normative entities, as it were, as they waft by.

4.2 THE ARGUMENT

The narrative structure of my argument is built around an investigation of the activity of trying to decide, on some occasion, what one is going to do. I begin from this topic in Chapter 2, and then make a set of arguments through chapters 3, 4, and 5 which allow me to return to it at the end of the project and offer my conclusion: when you are trying to decide what to do, you look outwards towards a value-laden world that you hope will guide you in making your decision. (And if all goes well, you will indeed be so guided.)

However, although the notion of deciding what to do shapes the narrative arc of the project as a whole, this is not because I take the activity of deciding what to do to play some especially metaphysically privileged role in our lives or our nature as agents. I have chosen to structure my arguments around this part of our lives simply because I think it is an effective tool, in two

²⁶ And in fact I think there probably are. I don't know what sort of thing 'wafts', exactly—butterflies maybe? Or elegant ladies, whose wafting is followed by a trail of expensive perfume? Well, such things are, I think, or can be, 'valuable things', and perhaps that is close enough to saying that they are 'normative entities'. But to say this at this stage of the dialectic between realist and skeptic is presumably not going to be very helpful or persuasive. I hope, however, that by the end of this project it will be clearer why this seems to me like both a natural and a true response to Korsgaard's remark here.

important respects. First, as we'll see in Chapter 2, it covers the right sort of ground to allow me to establish the points I want to make about the relationship between our attitudes and the arena of value (i.e., the topic that I have set up here as the defining question separating realist from anti-realist). And second, I believe that paying sustained philosophical attention to this part of one's experience can usefully orient one towards the role that realist value is already playing within one's actual life. It is within the context of this sort of reflection that a person might be brought to notice some of those wafting normative entities. But there are surely many different argumentative routes through which these points might be established; reflection on what we do when we're deciding what to do is not the only possibility. It is simply the one that I have chosen.

As the title of this project suggests, my investigation of the question of what you do when you're deciding what to do will involve a detour of sorts into the philosophy of mind, and in particular into the topic of what I will speak of as 'the first-person perspective'. This notion plays a number of different roles within my argument for realism. On the one hand, the first person perspective is part of my topic here: it is in part via some arguments situated within the philosophy of mind and concerning the distinctively first-personal way that we relate to (some of) our own mental states that I shall be arguing for the reality of value. On the other hand, the idea that you as my reader are yourself a person, occupying the very perspective that is the topic of some of these chapters, means that the notion of the first-person perspective takes on an important second kind of significance in relation to the methodology I employ in making my argument. At the most general level, my approach takes the form of an invitation to my reader to reflect on aspects of her experience and life and to say how she finds things to be therein. It is in part by appeal to your own experience of the first-person perspective that I make my case for the various conclusions I draw along the way.

It is also the case, however, that in conceiving of my project in this way, I hope to be able to lessen the pressure that arises from a certain sort of objection to my argument. The objection is as follows. What I claim to be offering in this project is an argument for value realism. Yet the form this argument takes is such that an opponent may be inclined to object that this is not really what I have shown at all. Instead, she may insist that—to the extent that my arguments work at all—all they really show is that we’re merely ‘first-personally committed to realism’, or that we ‘can’t help but think’ in terms of value as the realist understands it. And this (the opponent continues) by no means settles the question of whether or not realism is true—which is, surely, what we were supposed to be interested in determining.

For certain anti-realists, even my having achieved as much as this opponent allows will already be fatal to their position (Street, for instance, holds that a necessary component of her anti-realist account is the claim that we are first-personally committed to the rejection of realism). But for others of my opponents, merely establishing that (as they might want to put it) ‘things seem a certain way from within a certain perspective’ won’t establish anything interesting regarding the metaethical question about the realness of value that we started with. Within the context of this project, I cannot hope to address this concern in any detail. However, let me offer some considerations that will, I hope, serve to reduce the force of the worry somewhat.

My opponent here attempts to characterize my conclusion in terms of a claim about how things look from within a particular perspective. But my interest is not really in establishing, in the abstract, the nature of some particular entity or phenomenon called ‘the first person perspective’, nor in proving (merely) that ‘from this perspective’ such-and-such follows or must be endorsed. Rather, my interest is in you and what you think; how you find, after certain kinds of reflection, things to stand concerning you and your relationship to the world, in particular evaluatively

speaking. My goal is not to get you, after reading my arguments, to be willing to say ‘ah yes, I suppose maybe it could be that we’re first-personally committed to realism’ (after which you may well be inclined to add ‘...but so what? Is realism true?’). Rather, my goal is to get you to be willing (or just very slightly less unwilling) to say ‘huh, yes; I guess maybe I’m kind of a realist’ (after which there is not so much space to add the ‘...but so what?’). Even better: perhaps I might get you to be willing to assert the value of some particular thing, and to mean by this assertion that you take the source of that value to be in the object itself, rather than in some psychological state of yours. If I were able to reach, or even just begin to approach, such an outcome, then the gap that my opponent had seen opened up between what I have said I am trying to show and what I have in fact shown is at least markedly shrunk down.

4.3 REALISM, AND NOTHING LESS

Let me now clarify a potential misunderstanding of my view. In the previous section, I emphasized the way in which my argument both concerns the nature of the first-person perspective and at the same time involves a sort of first-personal appeal to you, my reader. This way of speaking about my project may invite a certain sort of misreading. For while the metaethics literature has in recent years seen steadily increasing interest in the idea of the first-person perspective, the most frequent occurrence of the notion has been within the context of constitutivist approaches to metaethics. The constitutivist argues that the agential perspective itself involves or entails certain normative commitments, and takes this entailment to establish the legitimacy of these commitments.²⁷ Given my own focus on the notion of the first-person perspective (and indeed on the activity of ‘deciding what to do’, which must at least bear some connection to the concept of the ‘agential perspective’),

²⁷ See especially Korsgaard (1996), (2009) and Velleman (2000).

it may appear that what I am defending is not quite realism in the fully red-blooded sense, but rather a variety of constitutivism. This would, however, be a misunderstanding of my position.

There are a number of differences between myself and the constitutivist. Perhaps the most obvious (although in some ways the least significant) is that my account, unlike the typical constitutivist's, remains at the strictly metaethical level. As I stated at the beginning of this chapter, I have nothing here to say about what sorts of things are valuable, and I do not suggest that reflection on the first person perspective is likely to offer any special insight into these first-order evaluative truths. Next, although the constitutivist and I are both talking about something that it makes sense to call 'the first-person perspective', my account of what that perspective looks like stands in stark opposition to at least some central cases of constitutivists. For Street, for example, speaking of this perspective is meant to signal the particular and idiosyncratic set of beliefs and evaluative commitments that an individual happens to possess; as I shall be arguing in Chapter 3, this approach to the first-person perspective leaves us with a picture not of a genuinely first-personal form of self-relation but rather of a case of alienation.

The most important difference between myself and the constitutivist, however, lies in the ultimately anti-realist stance that the constitutivist takes towards value. Unlike the constitutivist, I am not saying that value is real because you are already committed to it. Rather, I am saying that value is real quite apart from what you may think about the matter, and I am going to try to persuade you that in some sense you already know this, by drawing your attention to those aspects of your life in which you yourself are already seeking out and making decisions on the basis of what is actually valuable. Compared to the constitutivist, then, the position I am advancing is a genuine robust realism about value, not an ersatz or semi-real or mind-constructed realism. The fact that I am making my case by drawing on the idea of the first-person perspective, and by appealing to your

own first-personal experience as a thinking, acting being, does not alter or undermine this point. It is still realism, and nothing less, of which I am seeking to persuade you here.

4.4 AUTOBIOGRAPHY

I don't know how many pieces of philosophical writing have ever gotten any readers actually to change their minds about their particular topic. But it is perhaps worth mentioning that the reason that I have developed my defense of realism in the form that I have is that my arguments here trace the lines of thought that were what initially changed my mind about the metaethical status of value. A number of years ago, I was very comfortably on the skeptical side of the equation. It seemed obvious to me that the basic insight behind Mackie's 'argument from queerness' was right, and belief in this strange stuff called value was utterly unwarranted. Science, I reasoned, looks into how things are and will be in the universe, and the scientific method, while not without its difficulties, both in theory and in practice, seems to be a pretty decent way of figuring out how things are and will be. The notion that the universe might also come overlaid with a huge set of facts of a quite different form—facts about how things *should* be, whether in reality they are or not, and for which we have nothing even remotely comparable to the scientific method for investigating—struck me as just the sort of excess for which (to echo Gauthier's remark) Occam's razor was sharpened.

I can still feel the force of this sort of worry. I can also see the appeal of trying to make some sort of room for value talk along the various lines that the more optimistic anti-realists have suggested, in order to hold on in some way or another to this pretty fundamental part of our lives without having to invoke realist value in any robust sense. Eventually, however, I came to find my own anti-realism (in both its full-blown error-theoretic and its less bleak anti-realist forms) untenable. In the rest of this dissertation, then, I shall lay out the considerations that made me change my mind and endorse realism.

CHAPTER 2: DECIDING WHAT TO DO

1. INTRODUCTION

Suppose you have to decide, on some occasion, what to do: which dish to order from the menu, or which job to accept from a set of offers; what to do with a free afternoon, or with the rest of your life. What is happening—what are you up to—when you try to make such a decision?

This is the question that structures the narrative arc of this project as a whole. It is the topic of this chapter, in which I shall clarify how I am understanding the activity that I am calling (for now) ‘deciding what to do’, and draw out the problems that I need to resolve in order to make my argument work. It is then the topic to which I return at the end of this project in order to put together the pieces of the intervening arguments to establish my conclusion—my answer to the question of what you do when you are deciding what to do.

This answer (the short version of it, anyway) will be that what you are doing is looking outwards, towards a value-laden world which you hope will guide you in making your decision. And if all goes well, you will indeed observe some relevantly value-laden part of the world, and it *will* so guide you. If I can persuade you of such a picture, its metaethical significance is twofold. First, in so far as you yourself are already seeking value understood as the realist understands it, you are already committed to (at the very least) the possibility of such value. You are in some sense, that is, already a realist. Second, through identifying the activity of deciding what to do as a process of seeking value, my hope is that you might on some actual occasion of deciding what to do notice an instance of real value itself, and be in a position to see it as such. Along the way towards this conclusion, I establish additional results that either weaken the case for anti-realism or strengthen the case for realism. In

its broadest terms, however, my argument for realism is based around the asking and answering of the question of what we do when we're deciding what to do.

Given that the notion of deciding what to do plays such a central role in the development of my argument, it is crucial that I clarify the sense in which its playing this role is, perhaps surprisingly, a largely contingent feature of the nature of the case I am making. Although the idea of deciding what to do provides a guiding thread through the argument as a whole, this is not because of some especially privileged role that I take this notion to play either in our lives in general or in relation to the concepts of action or agency as such. I do, as it happens, think that our ability to decide what to do, in the sense detailed in the following section, is an important part of our being the sorts of creatures that we are—and I think that this is so despite its being, as I explain below, a fairly infrequent occurrence in our lives. However, this conviction of mine does not actually do any significant load-bearing work in the argument that follows. In particular, I want to stress that my argument does *not* take the form of a constitutivist argument founded on any claim about what is constitutive of agency, action, etc, as such. Whether or not the activity of deciding what to do plays any sort of constitutive role in relation to any of these concepts, I shall neither be assuming nor arguing anything of the sort here.

Instead, as I noted in Chapter 1, the role that the notion of deciding what to do plays in my argument is that of a useful tool. I have chosen to structure my argument around this part of our lives simply because it seems to me to be a helpful way in to thinking about the question of the relationship between (on the one hand) the nature of value and (on the other) our psychological attitudes—that is, the relationship that I established in Chapter 1 as the defining issue in the realist/anti-realist debate. It is through reflection on the activity of deciding what to do that I hope to be able to draw your attention to the ways in which you already seek (and, if you are lucky, not in

a state of depression, and so on, *find*) value in the world, prior to and not dependent on your own attitudes—value, that is, as the realist understands it. Reflection on what we do when we’re deciding what to do has no special claim to being the only possible way of doing this. The notion of deciding what to do, then, though central to the way I have chosen to develop my argument here, is at the same time in a sense inessential.

2. WHAT DO YOU DO WHEN YOU’RE DECIDING WHAT TO DO?

The activity of deciding, on some occasion, what you’re going to do, is I hope a familiar enough one. I mentioned some examples above, including cases where you are faced with a determinate list of options (the menu; the fortuitous multitude of job offers) as well as cases in which your reflections must be more open-ended in their possibilities (the free afternoon; the exhilarating or panic-stricken moment of wondering what one is going to do with the rest of one’s life). Either way, the phenomenon I have in mind involves the undertaking of an explicit and at least somewhat conscious process of trying to settle on a particular course of action, and it is worth noting that this is not actually something we do all that frequently. At least in this sense, then, I share Nomy Arpaly’s conviction that this notion can sometimes tend to be overemphasized in philosophical thought. “Characters in Hollywood movies” notes Arpaly, “encounter a lot of car chases. Characters in novels rarely wash their hands or do their laundry. And in the work of moral psychologists, people deliberate and reflect *a lot*” (2003: 20). And yet in reality, most—indeed, the vast majority—of our activities and actions are done without any such conscious prelude. Almost all of the time, we just *do* things—habitually, thoughtlessly, instinctively, improvisationally, in response to social cues, as an unreflective continuation of what we were already doing (maintaining a fascinating conversation; washing the dishes; pursuing a life in academia), or because this is ‘what’s done’ in this situation, or

whatever. ‘Deciding what to do’ in the explicit and conscious sense that I have in mind, then, often takes the form of an interruption, or at least a pause, in the ordinary flow of things. And this may seem to stand in some tension with the centrality it is accorded by many philosophers writing about agency or ethics.¹

In addition to this issue—and here I am departing from Arpaly’s complaint—it seems to me that philosophers have often been working with rather a narrow conception of what this activity must or ought to look like. One of the reasons that I’ve introduced my topic under the title ‘deciding what to do’ rather than that of ‘deliberation’ or ‘reflection’ is because I would like to allow a broader variety of forms of thought to come under our consideration than these latter terms have tended to suggest to philosophers.

One idea in particular that I would like to avoid is a heavily ‘calculative’ picture, on which the activity of deciding what to do must involve identifying and somehow weighing specific considerations counting for and against various options.² Another picture (sometimes, but not always, presented in combination with the calculative model) which appears to be extremely widespread among ethicists and metaethicists, and which I would also like to resist, is what I shall call a ‘stepping back’ model. Perhaps the most familiar version is that given by Christine Korsgaard:

I desire and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn’t dominate me and now I have a problem. Shall I act? Is this desire really a *reason* to act? (Korsgaard 1996: 93)

And again:

¹ Arpaly mentions Michael Smith (1994: 157); she also has in mind Korsgaard (1996). I’ll consider some other examples in what follows.

² Chapter 7 of Schroeder (2007) (‘Weighting for Reasons’) gives a good sense of what this model looks like, and of the extensive literature geared towards attempting to make sense of such calculations.

When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on. (Korsgaard 1996: 100)

Thomas Nagel, similarly, states that when we “adopt the standpoint of decision,” it

introduces a subtle but profound gap between desire and action, into which the free exercise of reason enters. [...] Once I see myself as the subject of certain desires, as well as the occupant of an objective situation, I still have to decide what to do, and that will include deciding what justificatory weight to give to these desires. (Nagel 1997: 109)

Among philosophers who have recently begun to raise criticisms of the significance accorded to ‘deliberation’, ‘reflection’, and (which is often treated as interchangeable with these—see the final part of this chapter) ‘the first-person perspective’, it is typically something like the ‘stepping back’ picture which provides the archetypal target of their attacks. Arpaly, for instance, asserts that “[d]eliberation is a process that necessarily involves a first-person experience of distancing yourself from your emotions” (2003: 22); Hilary Kornblith characterizes “reflection” as “thinking about one’s own mental processes from a first-person point of view” (2012: 28) and imagines “the reflective agent, reflecting on his or her many and varied first-order desires” (2012: 79); John Doris characterizes the philosophical doctrine of “reflectivism” which is the object of his attack as holding that “[i]n an exercise of agency [...] a person correctly divines the beliefs, desires, and other psychological states relevant to her decision, makes her decision in light of these states [...], and acts accordingly” (2015: x).

Leaving aside what each of these writers thinks of the significance of deliberation for moral psychology or human action, what all of these accounts have in common is the idea that what deliberation amounts to is, first, the achieving of a kind of distance from certain of one’s attitudes (for Korsgaard, the relevant attitudes are desires, but as the above quotations suggest some philosophers have a broader range of mental phenomena in mind). This ‘distance’, it seems, is

supposed to allow one to view one's own attitudes more accurately and/or more objectively. Finally, with these attitudes properly in view, one then makes one's decision, which takes the form of deciding which of the attitudes surveyed to endorse.

This 'stepping back' picture is, I think, deeply problematic in a number of respects, some of which I will address in more detail in section 3.3 of this chapter, returning to the issue again in Chapter 4. For now, however, my main concern with this picture, and with the 'calculative' picture mentioned above, is simply that neither of them adequately capture the full range of different activities that I would like to consider under the topic of 'deciding what to do'.

It is true, I think, that *sometimes* the activity of deciding what to do involves something like attempting to get some distance from one's more powerful emotions or desires, as well as maybe doing some sort of check that one's relevant beliefs are reasonably accurate. Whether or not one in fact achieves any of this, it seems to me that at least sometimes this is a more or less accurate description of what takes oneself to be trying to do. Likewise, *sometimes* deciding what to do may take the form of explicitly identifying and somehow 'weighing' the various considerations that speak for or against some course of action (and if this way of proceeding strikes you as the way to go on some occasion, Microsoft Excel now comes with a spreadsheet template that allows you to list individual pros and cons and accord each a weight; the spreadsheet then helpfully does the math for you and settles the matter of which option is best overall). But other times, instead of or as well as such a process, deciding what to do can involve such things as: imaginative engagement with the options or possibilities, or some sort of creative play with ideas until things click into place and something presents itself as, yes, the thing to do here. Sometimes it means trying to quiet the mind's attempts to think and overthink its way to a solution, and to sit with that quietness until the answer appears, or sleeping on things in the hope that they will seem clearer in the morning, or resolving simply to

trust in one's gut instinct. One time-honored way of making a difficult decision about what to do is to flip a coin, and then to try to notice whether or not your heart sinks a little at the outcome; if it does, you do the other thing. (If the idea of doing the other thing also makes your heart sink a little, the method isn't so helpful, unfortunately.)

Consider, for instance, what you do when you're choosing a meal from a restaurant menu. Perhaps you *do* begin assailed by certain desires (for the mac & cheese, or the *foie gras*, or the lobster) from which you decide you need to 'distance yourself'. Perhaps you decide that for reasons of health or morality or your limited budget, none of *these* particular desires will have their day today. And maybe that leaves you with only the salad, and so your decision is made, and the case looks somewhat like the sort of thing Korsgaard, Nagel, et al have described. But what if this process of 'stepping back', such as it is, leaves you with a range of options, all of which are healthy enough, not egregiously cruel in their methods of production, and within budget? In this case, you are in the lucky position of getting to choose what to eat simply *on the basis* of what you desire, what you want, what you feel like eating today—which is not, it seems to me, obviously the same as choosing *which* desire to gratify from some position in which you stand over and above those desires (for more on this point, see section 3.3 below). It is not clear, either, that you must make your decision between these permitted options via a process of identifying features that count for and against the various dishes and then performing some sort of weighing and calculating move. Personally, I close my eyes and imagine chewing and tasting the foods on my shortlist one by one, and then at the end of this process, I usually just know what I want. And that's how I decide what to do.

As another example of the varied forms that deciding what to do can take, consider a person (call her Miranda) who is trying to decide whether or not to give her marriage another chance following an instance of infidelity on the part of her husband (call him Steve). Miranda has made an

extensive list of pros and cons that Steve possess as a husband, and remains just as undecided as ever. But of course she does; as her friend (call her Carrie) points out, Miranda is a lawyer, and can, as Carrie notes, “argue both sides of any case”. So she must, says Carrie, stop “thinking” and start “feeling”: “I’m sorry, Harvard, but I’m afraid you’re going to have to base this decision on your emotions”. She must, that is, *follow her heart*. Miranda, a cynic to her core, doesn’t even really know what such a suggestion means. But her dilemma remains, and time is running out. Miranda and Steve’s couples therapist has insisted that following a period of enforced separation, they will each decide, separately, whether they are able to let go of the past, forgive one another (for it turns out that neither has been a perfect spouse), and move forward. For some reason, the therapist has declared that they must specify a place and time to meet, such that turning up at this place and time will be the sign from each of them that they are willing to give their marriage another go. The place will be the center of the Brooklyn Bridge, and the time will be... any minute now. Miranda, sitting alone in a coffee shop near the Brooklyn Bridge, must decide whether or not she is going to turn up and meet Steve.³ Glancing up from her meticulously penned list of pros and cons, Miranda spots her reflection in a mirror on the cafe wall. The coffee she is drinking has given her a foamy milk-mustache, and she recalls a moment from a past dinner with Steve, in which he had defused the tension of an imminent fight between the two of them by goofily giving himself a milk-mustache to match hers. Just as she had back then, Miranda now suddenly half-smiles—involuntarily, and somewhat reluctantly, but certainly genuinely. At this moment we as audience know, trained as we are in following the cues of the romantic comedy, that the list of pros and cons has not merely

³ Assuming, that is, that he’ll turn up and give *her* another chance. It doesn’t occur to Miranda until rather late in the day that *Steve’s* not turning up is, in fact, a real possibility.

acquired a new and powerful pro, but rather is quite forgotten; that Miranda has made her decision; and that she will go to the Brooklyn Bridge.⁴

I want to allow that such processes of deciding what to do are perfectly acceptable and non-defective examples of the phenomenon, despite not conforming to the calculative or stepping-back pictures. This is not because I think one ought (in matters of love, or in general) to ‘follow one’s heart’ (whatever that amounts to) but rather simply because I think that there are a wide range of different ways of deciding what to do, even given that we are restricting ourselves to the—as noted above, rather limited—realm of explicit and conscious decision-making. If, however, you are suspicious of such cases and inclined to regard them as somehow defective or non-ideal instances of the activity of trying to decide what to do, that’s okay. Later in the dissertation, I’ll give you some reasons to take these sorts of cases more seriously,⁵ but for now, you can work with whatever kinds of examples of deciding what to do that strike you as most familiar or most characteristic of the activity.

In speaking of ‘deciding what to do’, then, the terrain I’m interested in is one sense narrower and in another sense broader than that which is sometimes implied by philosophers’ talk of ‘deliberation’ or ‘reflection’. It is narrower in the sense that the conscious attempt to decide what to do is, I think, a relatively infrequent occurrence in our lives, and it is therefore not obviously the

⁴ Interestingly, I think that this scene is also supposed to let us know something that the character of Miranda herself cannot just then know, which is that her decision is the *right* one; that their marriage can survive and they can grow back together. It is part of the world of the romantic comedy that such things can be determinately the case at the moment of choice, even if the protagonist does not know it, and I suspect that this is less often how things are in the real world. The example of Miranda and Steve and the Brooklyn Bridge is taken from the (not as good as the television show but still unfairly maligned) first *Sex and the City* movie.

⁵ See Chapter 3, section 4.4 and Chapter 5, section 3.3.

place to look for the marker of action or agency as such.⁶ On the other hand, it is broader in that I am allowing that this activity may take a very wide variety of forms; deciding what to do need not always appear as a process of ‘weighing reasons’ or of ‘stepping back’ from one’s emotions or desires. It is for these reasons that I have introduced my topic with the more colloquial phrase ‘deciding what to do’, rather than the philosopher’s preferred terms of ‘deliberation’ or ‘reflection’. However, now that I have clarified the scope of the phenomenon in question, I will switch to speaking more or less interchangeably of ‘deciding what to do’ and ‘deliberation’ for ease of expression when engaging with other philosophers making use of the latter term.

3. THE ROLE OF (REAL) VALUE IN DECIDING WHAT TO DO

3.1 ‘GETTING SOMETHING RIGHT’

I’ve suggested that the activity of deciding what to do can take a wide variety of forms, and I mentioned that one of these is the old trick of tossing a coin, seeing how you feel about the outcome, and then letting that immediate emotional reaction guide your decision. It’s important, however, that *merely tossing a coin* to settle what you are to do (without the additional attempt to catch oneself in the act of some presumably otherwise hidden feeling) does *not* count as an instance of deciding what to in the sense that I am interested in. In this case, what you have decided to do is to *stop* trying to decide what to do, and instead to allow the matter to be settled by chance.

⁶ To be clear: the infrequency of its occurrence doesn’t preclude this activity from being at least partly constitutive of agency or at least of the sorts of agents that we are, and in fact I think there is *something* to this idea. However, it certainly seems to indicate that to give a complete account of agency it will not be sufficient merely to give an account of the conscious activity of deciding what to do. And in any case, as I’ve noted, nothing in my argument requires that I be committed to any claim of this sort.

This observation points us towards an interesting feature that the otherwise widely varying forms of deliberation share in common. For when you are trying to decide what to do, it seems that one thing you *cannot* do is to suppose that any conclusion reached as a result of your deliberation will do as well as any other—or at least, you can't suppose this at the same time as sincerely continuing to deliberate. Instead, it seems, there is something that you are trying to get *right*. To be clear, I mean this in a fairly weak sense. In particular, I don't think that in trying to decide what to do you must assume that there is a single correct outcome that you trying to reach. Nor, of course, does 'right' here mean 'morally right' or anything of the sort. Nevertheless, your attempting to decide what to do means that you are supposing that some of the answers you might reach are 'better', at least in *some* sense, than some of the others, and furthermore that their being better is not itself determined merely by your ending up settling on that answer rather than another.

This feature of deliberation has been remarked upon before. Here is Nagel, writing some time ago in *The View From Nowhere*:

The ordinary process of deliberation, aimed at finding out what I should do, assumes the existence of an answer to this question. And in difficult cases especially it is often accompanied by the belief that I may not arrive at the correct answer. [...] In deliberation we are trying to arrive at conclusions that are correct in virtue of something independent of our arriving at them. (1986: 149)

Now, as we have seen, Nagel endorses a 'stepping back' model of deliberation. Does his observation still hold true of the broader range of deliberative activities that I have introduced? For instance, it seems clearly true of Miranda while she is making her list of pros and cons that there is something that she is trying to get right. But does it remain true of her when she glances up, notices her milk-mustache, and—abandoning the list—makes the decision that she is going to meet Steve at the Brooklyn Bridge? For there is one possible reading of the situation on which her giving up on the

activity of listing and weighing pros and cons is at the same time a giving up on the idea that what she is trying to do here is to get something right.⁷ However, this reading does not seem plausible to me. For Miranda's decision to go to the Brooklyn Bridge is her *answer* to the question that she asks herself when she sits down at the coffee shop to think. She asks herself: "What shall I do here? Shall I meet Steve and give our marriage another chance, or not?" It is clearly important to her to come to the right decision here; this is precisely why she takes so much effort in writing her list of pros and cons. Nonetheless, this strategy of attempting to reach the right answer does not work for her. But when she abandons it, it is because she *finds* the right answer through some quite different means. (In an alternative version of the story, we might imagine that Miranda, frustrated with the impossibility of discovering the right answer through her listing of pros and cons, decides to give up trying and instead to toss a coin and let *it* settle the matter. But this is not, in fact, what she does.)

Notice further that the significance of Miranda's noticing her milk-mustache is given precisely by the context within which this event is situated—namely, her having sat down in the coffee shop to try to decide what to do. Had she seen her reflection-with-milk-mustache under some other circumstances, then even if it had caused her to recall Steve in the same sort of way, this moment would not have meant what it means when she is thinking in the coffee shop. It would not have taken the form of providing, as it does here, an answer to a question asked.

Thus, the relationship between what Miranda is doing in her initial sitting down to think, and what she is doing when, on seeing her milk-mustache, she gathers her belongs and hurries out of the coffee shop to meet Steve, is an intimate one. It is a relationship between a question asked, and asked with the assumption that there is, or at least could be, a *right* answer, and an answer given as Miranda's good faith attempt to settle upon the right answer.

⁷ Thanks to Agnes Callard for pressing me on this.

To return, then, to the thread of argument: the feature of the phenomenology of deliberation that Nagel points us towards—that it involves an attempt to ‘get something right’—has been invoked more recently by David Enoch as the basis of an argument for (what he calls) ‘robust normative realism’. The argument as a whole comes in two parts, only one of which will be relevant for my purposes. In the argument’s first part (2011: 53–71), Enoch attempts to show that just as our best scientific theories license our belief in those entities (electrons, for example) that are essential to the theory, so too the activity of deliberation—being just as crucial to our lives as the explanatory project of science—can license belief in those entities essential to *it*. In the second part of the argument Enoch strives to establish “that irreducibly normative facts are indispensable for deliberation” (2011: 12). Putting these two parts together, Enoch claims, offers a novel defense of a strong form of realism. This first part of Enoch’s argument is not relevant to my own, and can be set aside. My interest is in the second part of the argument, for here Enoch seems to be making a claim that bears some similarity to the conclusion which I shall eventually be defending—namely, that what you do when you’re deciding what to do is to look outwards towards a value-laden world that you hope will guide you in your deliberations. Further, Enoch’s argument for this claim begins with the same point as that with which I opened this section: the observation that deliberation involves an attempt to ‘get something right’. Enoch and I, then, begin from more or less the same premise and reach what seems like rather a similar conclusion. However, I do not think that Enoch’s argument is satisfactory. What I shall try to do in what follows is to explain why I think the argument doesn’t work, in order to clarify what needs to be achieved in the following chapters to enable me to draw the conclusion that I believe Enoch is not entitled to draw.

Enoch's positive argument for the idea that deciding what to do involves some sort of commitment to realism is basically an elaboration of the 'getting something right' feature of deliberation that I noted above. Here is Enoch:

"All things considered", you ask yourself [in deliberation], "what makes best sense for me to do? When all is said and done, what should I do? What *shall* I do?"

When engaging in this deliberation, when asking yourself these questions, you assume, so it seems to me, that they have answers. These answers may be very vague, allow for some indeterminacy, and so on. But at the very least you assume that some possible answers to these questions are better than others. You try to find out what the (better) answers to these questions are [...]. You are not trying to create these answers. [...] When trying to make up your mind, it doesn't feel like just trying to make an arbitrary choice. This is just not what it is like to deliberate. Rather, it feels like trying to make the *right* choice. It feels like trying to find the best solution, or at least a good solution, or at the very least one of the better solutions, to a problem you're presented with. What you're trying to do, it seems to me, is to make the decision that makes most sense for you to make. Making the decision is up to you. But which decision is the one it makes most sense for you to make is not. This is something you are trying to discover, not create [...]. Or so, at the very least, it feels like when deliberating. (Enoch 2011: 72–73)

This analysis is the cornerstone of Enoch's argument. Following this positive claim, he spends some time answering a number of potential objections to the conclusion he wants to draw from this, at least one of which takes the form of a counter-proposal from a sort of anti-realist (I consider this objection, and Enoch's reply, below). But overall, Enoch says very little to explain why the feature of deliberation that he has highlighted—that the deliberator must consider some answers to her deliberative question to be 'better' in some sense than others—can be understood only in realist terms.⁸

⁸ To be fair to Enoch, he does acknowledge that his argument is somewhat lacking in this respect: "Rather than fully defend the indispensability premise my indispensability argument relies on, I hint at why I think it is probably true" (Enoch 2011: 83); elsewhere, he describes what he does in this part of the book as "hint[ing] at reasons why I suspect that some major alternative views cannot accommodate what is indispensable for deliberation" (Enoch 2011: 80). So Enoch is not exactly claiming that he has offered a water-tight case; at the very least, then, there is plenty of room for

As it turns out, I think that the anti-realist is in quite a strong position in relation to the kind of argument that Enoch is trying to give. For there are plenty of *prima facie* very plausible things that various kinds of anti-realist can say to make sense of this aspect of deliberation. Further, I shall argue that the response that Enoch *does* offer to one sort of anti-realist has the unfortunate consequence of feeding right into the hands of another. Ultimately, I shall suggest that the strength of the anti-realist's position here is not an accident: there is something that the anti-realist has gotten right, and that Enoch (and, as it turns out, other realists) have gotten wrong. In my own argument (in particular in Chapter 5) I shall try to show how the anti-realist's insights can be incorporated into what is nonetheless a thoroughly realist picture of value.

3.2 AN ANTI-REALIST REJOINER

Let me begin with a simple suggestion on behalf of the anti-realist. Perhaps when I deliberate about what to do, all I'm really up to is attempting to figure out which course of action would best satisfy some particular desire of mine (or perhaps the set of my desires taken as a whole). Call this proposal 'simple neo-Humeanism' about deliberation. Such an account of deliberation seems to be perfectly compatible with the basic anti-realist commitment to prioritizing human attitudes over the space of value, for according to the anti-realist people do indeed tend to come equipped with desires as a naturalistically unworrying matter of fact (see my discussion in Chapter 1, 3.1). The suggestion is that it is simply these (as opposed to some mysterious realm of value) to which I turn in deciding what to do. Further, since there may well be a 'right answer' to the question of which course of action is most likely to satisfy my desire(s)—or at least some answers that are better than others—

development and defense of Enoch's idea that only realism can account for the getting-something-right feature of deliberation.

there is no mystery, on such an account, to the observation that deliberation involves an attempt to get something right.

Indeed, when we return to Nagel's discussion (which Enoch himself cites as a key inspiration for his own analysis of deliberation), we find that Nagel himself immediately allows that one way of satisfying the 'getting something right' feature of deliberation is what he calls (following Parfit) an "Instrumental Theory"—which is to say, an account that more or less echoes the 'simple neo-Humeanism' I outlined above (Nagel 1986: 149–50). This suggests that Nagel, at least, doesn't seem to think that this feature of deliberation, all on its own, takes us very far towards the sort of realism that Enoch and I wish to establish (although he does allow that even the Instrumental Theory might function as "the thin end of an objective wedge" (Nagel 1986: 150)).

Enoch does raise and respond to an interlocutor who may be interpreted as holding something like the simple neo-Humean position. This imagined interlocutor says: "*Well, aren't our desires enough for deliberation? Why do we need normative truths to settle deliberation, when we are moved by desires?*" Enoch's response to the interlocutor is as follows (some footnotes are omitted):

[Desires aren't enough for deliberation] because when you allow yourself to settle a deliberation by reference to a desire, you commit yourself to the judgment that your desire made the relevant action the one it makes most sense to perform. This may be either because, say, there is a general reason to satisfy your desires, or because you take having a reason to simply consist in having the relevant desire. But on either of these options, and even with desires at hand, you still commit yourself to a normative truth. Had we been here in the explanatory business - trying to explain action, or perhaps even deliberation, from a third-person point of view - perhaps desires would have been enough (though I doubt it). But the whole point of the argument of this chapter is the focus on the first-person, deliberative perspective. And from that perspective, desires are not often relevant, and whether they are or are not, the normative commitment is - though perhaps implicit - inescapable. (Enoch 2011: 75–76)

Enoch says that even if you decide, on some occasion, to act on the basis of some particular desire or other, in doing so you are—implicitly, even if you don't acknowledge as much—giving that desire

a stamp of approval; that is, you are making an additional normative judgment to the effect that it would be good (or permissible, or appropriate, or whatever) to act on that particular desire in these particular circumstances. And this, Enoch seems to think,⁹ means that you have given up the game to the realist. Given the analysis of the realist/anti-realist distinction that I defended in Chapter 1, however, this is not at all the case.

What is perhaps most striking about Enoch's response to his interlocutor is its similarity to the 'stepping back' picture of deliberation that I discussed in section 2 of this chapter. Enoch describes his idea here as "very Kantian, of course" (Enoch 2011: 76), and cites the *Prolegomena* in support. Now, whether or not such a picture may properly be attributed to Kant himself, Enoch's account here resembles nothing so closely as Korsgaard's interpretation of Kant, and the problem with this is, of course, that Korsgaard is not—not on my account, not by her own account, and almost certainly not according to Enoch either—a realist.

However, the problem does not stop at the fact that Korsgaard could have written the above passage. Certainly, neo-Kantians and realists make uneasy bedfellows, and Korsgaard in particular is one of realism's most sophisticated critics. Nevertheless, there are still key points of commonality between the two sides, including in particular the conviction that at least certain normative propositions are indeed the case whether you like it or not (we share, that is, a perhaps somewhat po-faced *seriousness* about the whole thing, and a desire to do justice to that sense of seriousness)¹⁰.

⁹ Although see my caveat in the previous footnote; it also turns out (although this is not made terribly explicit in the text) that Enoch thinks that certain constitutivist or constructivist views *may* be able to account for the 'getting something right' feature of deliberation (see (2011: 82–83)). Enoch holds that such views, however, are untenable for other reasons (see e.g. his (2006) and (2009) for these other arguments).

¹⁰ None more so, of course, than Enoch himself, whose own book title declares his commitment to *Taking Morality Seriously*.

But Korsgaard is not the only sort of anti-realist who can allow Enoch's response to his interlocutor to stand.

Enoch's claim is that desires aren't enough to settle deliberation; one needs in addition a normative judgment or commitment of some kind. That we can and do make such judgments is, however, precisely the point upon which Sharon Street's determinedly anti-realist constructivism is based, and in this case there is surely no hope of *détente* with the realist. Street holds that it is the normative judgments that a person makes which determines what she does or does not have reason to do, and thus that "things are valuable ultimately because we value them, not the other way around" (Street 2012: 41).¹¹

What this highlights is that it would be a mistake to think of *desire* as the cornerstone of the anti-realist threat that Enoch and I both face. This may once have been so—in the case of the historical Hume, for instance—but modern-day anti-realists have sought to equip themselves with a much broader and more sophisticated set of tools. Desire is far from the only piece of psychic machinery that the anti-realist can make use of in the account she gives of what we're up to when we're deciding what to do.¹² Since this is so, responding to Enoch's imagined interlocutor by

¹¹ One of the most interesting things about Street's position—which she herself characterizes as a Kantian-Humean hybrid—is that it reveals just how close neo-Humeanism and neo-Kantianism can come to one another without seeming to lose all that much in their essentials. Street claims that she and Korsgaard hold positions that start from the same basic ideas about normativity and its source in our own normative judgments, and that differ only in terms of how much of content each thinks can be extracted from this starting point. If this is right as a reading of Korsgaard (on this point I am not totally convinced, but I imagine that Street, having been Korsgaard's student, would be in a better position than I to say), then for all their seeming opposition, neo-Kantianism (assuming that we can take Korsgaard as its exemplar) might be thought of as essentially just a *really optimistic* form of neo-Humeanism.

¹² Even Schroeder, whose anti-realist (on my taxonomy) account *is* based on desires emphasizes that it is not ultimately *this* fact that characterizes the account as 'Humean'—one might, that is, accept all of the rest of Schroeder's anti-realist account of reasons but merely wish to substitute some other psychological feature in place of 'desires'.

insisting that she must implicitly approve any desire on which she wishes to act with a ‘normative judgment’ will not do. From where the anti-realist is standing, a normative judgment is just another example of the sorts of psychological attitudes that we’ve already acknowledged people do in fact possess. This is why Street feels able to say that what she is offering is “an account of what a reason *is* in terms of what it is to *judge* or *take* things to be reasons, where our understanding of this attitude is prior to and fully independent of our understanding of what a reason *is* in the relevant sense—that is, a clear, naturalistically comprehensible sense” (Street 2008: 241). (Compare again Scanlon: “the state of taking something to be a reason, as I interpret it, is a purely psychological state [...] Even a belief in witches is a naturalistic psychological state in this sense” (Scanlon 2014: 62)). This is why I characterized the distinction between realist and anti-realist the way I did in the previous chapter: as a matter, that is, of the order of priority between value, on the one hand, and the whole gamut of human attitudes, on the other.

With this understanding of the distinction in mind, we can formulate the following schema, which the anti-realist is free to fill out in any number of ways in order to make sense of the getting-something-right feature of deliberation that Enoch holds can only be explained by realism. The schema is this: in broadest terms, what makes an answer to the question ‘what shall I do here?’ better or worse is *not* the value inherent in certain courses of action or their outcomes, but rather, ultimately, some psychological attitude or attitudes of mine—maybe a desire (or set of desires) as in the simple Humean proposal, but not necessarily; maybe a normative judgment or an evaluative commitment, or a non-cognitive pro-attitude that I express in stating such a judgment, or a plan or goal I’m pursuing, or any of a wide array of attitudes within what Williams calls my ‘motivational set’, and/or the logical consequence of such, and so on.

Of course, different anti-realists will want to fill out the schema in different ways. And there is plenty of existing literature highlighting the strengths and weaknesses of different forms of anti-realism that could be brought to bear on these different iterations of the schema. However, as I emphasized in the previous chapter, my strategy against the anti-realist is not one of divide-and-conquer; I am interested in treating the various anti-realist positions as much as possible *en masse*. The point of the anti-realist schema, then, is to indicate that as things currently stand, the (generic) anti-realist has *plenty* of options to play with—any one of which looks like it will allow her to make sense of the getting-something-right feature of deliberation that Enoch and I are interested in. For some options that I might entertain in my deliberations could be, as a matter of perfectly objective fact, more or less likely to satisfy the relevant desires, or more or less in accord with my values, or more or less likely to help me meet my goal, and so on.

As illustration of how the anti-realist might develop such an account, consider the following remarks from Sharon Street:

Quite apart from whether we think a given set of values is *correct*, in other words—indeed, even if we aren't clear yet on what it *is* for a set of values to be correct—we can nevertheless think about and discuss what *follows*, as a purely logical and instrumental matter, from a given set of values in combination with the non-normative facts. For example, we might say “Ann is badly mistaken to value counting blades of grass above all else, but it follows from within her (bizarre and mistaken) evaluative point of view that she has reason to buy a calculator.” “Moreover,” we might add, “this follows even if Ann doesn't recognize that it does.” Due to her ignorance of the non-normative facts or some other cognitive deficiency (perhaps Ann doesn't know or keeps forgetting what calculators do), Ann might never *see* that it is entailed from within her evaluative standpoint that she has reason to buy a calculator; nevertheless it *is* entailed. (Street 2010: 367)

Street's particular version of anti-realism would thus allow her to make sense of Ann's deliberation (such as it is) in something like the following way. Ann thinks: “I've come into some money. What shall I spend it on? Well, let's see. I value counting blades of grass above all else. So what sort of

purchases will best support this life goal of mine?” And, Street suggests, there is a right answer to this question. Indeed, presumably there are many purchasable things that would aid Ann in her endeavor, so there are plenty of ways that she might ‘get it right’ in concluding her deliberations. One of these right answers, according to Street, would be for her to decide to buy a calculator. (And one of the marks of this ‘rightness’ being objective in the way we need is Street’s observation that Ann might, for a number of reasons, get things *wrong*; she might settle on purchasing something that will not be helpful at all in her quest—an abacus, say, when it turns out that she does not know how to use it properly and has no means of finding out.)

In raising this first problem for Enoch’s argument, my interest is really in clarifying the nature of the challenge that *I* face in my attempt to argue that deliberation involves looking outwards towards realist value. What my discussion here indicates is that if my account is to be successful, I must show why *no* form of the anti-realist schema can ultimately work as an explanation of the getting-something-right feature of deliberation. I must show, that is, that psychological attitudes *in general*—and not only the attitude of desire—cannot play the role that the anti-realist wants them to play in the activity of deciding what to do, and that instead realist value is needed for you to make sense of what you’re up to in deciding what to do.

3.3 MAKING (THE RIGHT KIND OF) SPACE FOR DESIRE IN DELIBERATION

I want to raise a second problem for Enoch’s account that once again will help to clarify a constraint that I must meet if my own argument is to be successful. In the previous section, I considered Enoch’s response to an interlocutor who suggested that desires might do the necessary work to allow us to make sense of deliberation. I argued that Enoch’s “very Kantian” response to the (neo-)Humean merely opened the door to any form of anti-realism that shares Korsgaard’s and Street’s idea that the making of normative judgments on the part of the agent (as opposed to some

independent normative reality) is itself the source of normativity. I then offered a generalized schema on behalf of the anti-realist that can be filled out in any number of ways, not merely with the neo-Humean desire-based proposal. I want now, however, to return to the notion of desire, and come at Enoch's response from the other direction: in making so much of a 'stepping back' picture of deliberation, Enoch (and likewise Korsgaard and Nagel) fail, I shall suggest, to give desire (and, more generally, individual and idiosyncratic preference) its proper due.

Consider again Enoch's response to the *aren't-desires-enough?* objector to his argument that deliberation invokes realist value. Enoch says, recall, that:

when you allow yourself to settle a deliberation by reference to a desire, you commit yourself to the judgment that your desire made the relevant action the one it makes most sense to perform. This may be either because, say, there is a general reason to satisfy your desires, or because you take having a reason to simply consist in having the relevant desire. (Enoch 2011: 75–76)

Notice that Enoch is here *allowing* that his interlocutor may hold the view that "having a reason simply consist[s] in having the relevant desire"; he is not—not here, at least—objecting to such a position. Assuming that this should be read as the claim that *all* reasons 'are' (or perhaps somehow follow upon) desires, such a position would clearly count as a form of neo-Humeanism, and thus as anti-realist on my taxonomy. Despite appearances, then, Enoch is *not* actually challenging most contemporary iterations of neo-Humeanism here. Again, this may be fine from Enoch's point of view, since he has other arguments later in the book that can be marshaled independently against the neo-Humean. But it is, of course, a problem from my point of view.

It seems to me that a large part of the enduring appeal of the neo-Humean approach lies in its ability to account so comfortably for those actions and decisions that actually—if we are lucky, anyway—make up the bulk of our lives: the mundane and unimportant choices made for no

particular reason other than preference, or the fancy of the moment, or some idiosyncratic desire that I don't really expect anyone else to share. Likewise, when it comes to more significant matters that nevertheless seem inescapably 'personal', such as choosing the particular people with whom one wishes to be friends, or to date, or to marry, it has seemed to many philosophers that the source of one's 'getting it right' in these cases could only be understood as residing *in me*: in my idiosyncratic character and tastes and values and desires and so on. Indeed, as I shall show, plenty of realists (as well as neo-Kantians and others) have been happy to allow that *these* sorts of decisions and actions, both the mundane and the significant, can be understood in terms roughly amenable to the neo-Humean. Thus, there is a sense—widespread enough, although perhaps not fully the 'consensus' that we'll see a number of neo-Humeans claiming—that something like neo-Humeanism might be thought of as a sort of baseline metaethical position, with the contentious question then being whether there is anything *more* to reasons, or value, or whatever, than the neo-Humean account acknowledges. Consensus or not, this is an assumption that I shall here be resisting. I can only successfully undermine the neo-Humean's claim to be the 'common denominator' of metaethical positions, however, if I can offer a decent alternative to the neo-Humean's seemingly compelling account of these distinctively personal or idiosyncratic choices.

Let's see just how powerful the neo-Humean account seems to be in these sorts of cases.

Consider this example of Mark Schroeder's:

Tonight there is going to be a party, and everyone is invited. There will be good food, drinks, friendly chat, music—and dancing. Ronnie and Bradley, like everyone else, have been invited to the party. But while Ronnie loves to dance, Bradley can't stand it. Not only does he not like dancing, he prefers to stay away from where it is going on, lest he come under pressure to be shown up in his awkward maneuvers by those with fewer left feet than he. So while the fact that there will be dancing at the party is a reason for Ronnie to go, it is *not* a reason for Bradley to go. Far from it; the fact that there will be dancing at the party is a reason for Bradley to stay away. Ronnie's and Bradley's reasons therefore differ—each has a reason that the other does not. (Schroeder 2007: 1)

The example of Ronnie and Bradley and their differing reasons is used by Schroeder as a touchstone throughout his book's defense of a neo-Humean position.¹³ Schroeder goes on to assert that:

it's not hard to see *why* Ronnie's and Bradley's reasons differ, at least at a first pass—this is something to do with their respective psychologies. It is because of what they *like, care* about, or *want*. [...] It's largely uncontroversial—even among philosophers—that at least *some* reasons are like Ronnie's and Bradley's, in that whether they are reasons for some particular person depends on some feature of that person's psychology, such as what that person desires, wants, likes, or cares about. (Schroeder 2007: 1)

With this assumption in place, Schroeder is then able to characterize the neo-Humean position in the following way:

And according to one time-honored tradition often taken to be vaguely inspired by Hume, *all* reasons are at some basic level pretty much like Ronnie's and Bradley's in this way. They are explained in some way by the psychological features of the agents for whom they are reasons—by the psychological features that explain the difference between Ronnie and Bradley. (Schroeder 2007: 1)

On the very first page of his book, then, Schroeder has set up the dialectic between himself and his opponent in a very distinctive way. We all more-or-less agree, he claims, that *some* reasons arise from some particular attitude that the person possesses (though we may disagree, as he notes in a footnote, regarding exactly which attitudes are the relevant ones: it may be “intentions” or “ends” or “valuings” rather than, as he will claim, desires). The task he sets himself in defending neo-Humeanism is then that of showing that *all* reasons can be understood in this sort of way.¹⁴

¹³ To reiterate a point I made in the previous chapter: Schroeder classifies his position as reductive naturalist realism; on my taxonomy, he counts as an anti-realist.

¹⁴ He does so by arguing *against* all of the objections to such a claim. The assumption that at least *some* reasons are like that is thus playing an unobtrusive yet essential role in making such defensive work amount to a genuinely positive argument for neo-Humeanism.

There is something that Schroeder does not quite make explicit in these opening remarks, but which will be from my point of view crucial. In claiming that whether something is a reason for someone “depends on some features of that person’s psychology, such as what that person desires, wants, likes, or cares about”, what Schroeder means (it turns out; see the contents of his Chapter 4 if in doubt) is that the thing’s counting as a reason depends on the relevant feature of that person’s psychology and *not* in any way on any *further* evaluative fact about the inherent goodness (or similar) of the thing in question (compare my discussion of Williams in section 3.5 of Chapter 1). On Schroeder’s account, it is the attitude alone, in combination with the non-normative facts, that gives rise to value. Schroeder’s starting assumption, then, is that the majority of moral philosophers already think that *this* is how things work in at least some cases.¹⁵

I don’t know whether Schroeder is right to characterize this idea as a consensus—but maybe that is just because I shall be disagreeing with it. Some version of the thought is certainly common enough. Nagel, for instance, wants to make room in his view for two different kinds of ‘reasons’: first, ‘agent-neutral’ reasons, the acknowledgement of which is supposed to be what renders his view genuinely realist, and second, another sort of reason which depends entirely on our own idiosyncratic desires, commitments, and so on:

¹⁵ Schroeder is not alone in considering it uncontroversial that at least some reasons work like this. Here is Paul Katsafanas: “if you have an aim, you have a (pro tanto) reason to fulfill it.[...] this claim [...] is relatively uncontroversial; even the most minimal accounts of practical reason, including most variants of the Humean account, accept it” (Katsafanas 2013: 184). See also Korsgaard: “Most people suppose that the means/end relation is normative, in the sense that the fact that a certain action is a means to your end provides you with a reason to do it. Very few people have ever supposed that this requires an adjustment in the metaphysics of the Modern Scientific Worldview” (Korsgaard 1996: 36) (of course, this statement of Korsgaard’s should be considered in light of her (Korsgaard 1997)). Joseph Raz rightly notes that things might not be so simple, since we may well want to say that “a would-be murderer cannot create for himself a reason for poisoning his intended victim just by making it his goal to kill him” (Raz 2005: 2–3); he does acknowledge, however, the force of the intuition that Schroeder *et al* are drawing on.

Most of the things we pursue, if not most of the things we avoid, are optional. Their value to us depends on our individual aims, projects, and concerns, [...]; they acquire value only because of the interest we develop in them and the place this gives them in our lives, rather than evoking interest because of their value.” (Nagel 1986: 168)

This quotation from Nagel is in line with Schroeder’s claim that the real question the metaethicist must answer is whether value can be understood solely in neo-Humean terms, or whether rather we must *add* something to the neo-Humean picture in order to make sense of certain special forms of value, such as Nagel’s agent-neutral reasons.

I suspect it is not an accident that Nagel both endorses, as we saw in part 1, the stepping-back picture of deliberation, *and* is inclined to understand those pursuits which are “optional” in basically neo-Humean terms. For, as I shall explain, the stepping-back picture will tend to leave the deliberator with a certain problematic gap, a gap of the sort which neo-Humeanism seems very well-placed to fill.

Recall my example of choosing food from a restaurant menu, introduced in the first part of this Chapter, and consider a structurally similar example. Suppose I am deciding whether to buy a particular pair of sneakers in white or in grey. According to the stepping-back model, I must step back and take a look at my desires. I note that I have a longing for the white ones (the contrast with my black jeans will look striking and fresh, and will brighten up otherwise dreary outfits), and I have to decide whether or not to give this desire my endorsement. Now, suppose that prior to this point, presumably in a previous act of stepping-back, I’ve *already* taken into account all the other considerations I can think of, such as cost and practicality of upkeep and the ethics of manufacture and so on, and all of these reflections have left me with the two choices of the grey and the white pair. For *these* two, all these other factors are (let’s suppose) equal. Everything is equal, that is, except my longing for the box-fresh white ones. Given that everything else is equal, and thus that I’m in

equal measure 'permitted' (so to speak) to choose either the grey or the white, which shall I choose? In other words, how am I to understand the relevance, from the point of view of my neutral, rational, deciding self, of my desire for the white ones? It seems pretty clear both that I can (am permitted) and will (unless I have a very specific shoe-purchase-based form of masochism, I guess) choose to follow my desire in this case. But is it still the case that I am here trying to 'get something right'? As we might alternatively put the question, is there really a *reason* for me to go with the ones I (happen to) want, or not?

Here are two possible ways that a proponent of the stepping-back model of deliberation might think about what is happening here. First, she may suppose that what I do at this point in my deliberations is essentially to decide that I am permitted to *stop* deliberating, and to allow the sheer force of the desire for the white sneakers (held thus far at bay by the power of reason) to overtake me and drive me to the cash register with the white, and not the grey, sneakers under my arm. The other possibility is that my selecting the white sneakers remains under the aegis of the rational deliberative process; that there *is*—as the neo-Humean has suggested—at least a pro-tanto reason to satisfy my desire. And since, in this case, there is nothing standing *against* my satisfying it, this pro-tanto reason takes on the full force of a reason proper, and I am rational and justified in going ahead and buying the white sneakers.

The stepping-back picture, I am suggesting, leaves us with a gap to fill: the gap of explaining what is going on when I select, from among the shortlist of options that the stepping-back process has left me with, the one that I (merely happen to) *prefer*. The two possibilities I raised above may be thought of as representing, roughly, stepping-back-plus-bald-Humeanism, and stepping-back-plus-neo-Humeanism respectively. Either way, the gap that the stepping-back picture leaves us with is one that it seems natural to want to fill with something approximately Hume-shaped.

Why, then, is this a problem? Well, *perhaps* it will turn out not to be a problem for Korsgaard or Nagel or Enoch or any of the other stepping-backers.¹⁶ But it *is* potentially rather a large problem for me. For I am arguing, recall, that *what you do when you're deciding what to do* is to look outwards, towards realist value. And yet there seems to be a very broad and familiar space of cases in which a great many philosophers—including realists—have assumed that the best explanation of why it makes sense for me to choose *this* rather than *that* just is that I happen to have some particular attitude, namely a desire or preference. Further, this assumption is perfectly comprehensible. For plenty of decisions, it seems to me, my desires and preferences clearly do matter, and indeed matter a great deal, in making it the case that one outcome of my deciding what to do would be 'better' than another.

But all of this is just to say that it seems that one particular fleshing out of the anti-realist schema (the broadly neo-Humean one) *can* make sense of the getting-something-right feature of deliberation in at least a wide range of cases. As I noted at the end of the last section, however, my argument depends on showing that *only* realist value can do the relevant work. Thus, I face the following difficulty: I must explain how it is that we can make choices on the basis of desires and preferences while still maintaining that doing so can itself only be understood within a realist framework.

On the one hand, then, my approach faces a problem that does not seem to arise with quite the same urgency for other realists, and this might be thought a negative for my account. On the other hand, if I can resolve the problem, my account has a benefit over other forms of realism in so

¹⁶ Though note Korsgaard's anxiety about her own account, expressed in the context of a commentary on Jonathan Lear's *A Case For Irony*: "If our ordinary desires are just so much random debris blowing in through the causal network, we might wonder why it's rational to act on them at all" (Korsgaard, in (Lear 2011: 79)). Her answer to the worry is given in terms of the concept of "practical identities", as developed in her (1996) and (2009).

far as it will enable me to harness many of the important goods of anti-realism *without* allowing the neo-Humean the foot-in-the-door that Schroeder makes use of: I will be able to reject his starting assumption, supposedly so uncontroversial, that Ronnie and Bradley's reasons, even if no others, ought to be understood in anti-realist terms. On the account that I shall develop here, we cannot make sense even of Ronnie and Bradley's reasons without invoking realist value.

4. MOTIVATING AN INVESTIGATION OF 'THE FIRST-PERSON PERSPECTIVE'

I will be making my case via an investigation into what I shall call 'the first person perspective'. This is a concept that has seen increasing interest in metaethics over the last few decades. However, those writers who make use of it have rarely offered much elucidation on what is meant, and in fact on closer inspection it turns out that people have used it to mean rather different, and in some cases almost entirely opposed, ideas. As illustration of the extent of this problem, I'll give a rough overview of the way that the authors that we have been dealing with in this chapter—Enoch, Nagel, Williams, Korsgaard, and Street—have thought of and made use of this notion.

In the course of making his argument about deliberation and realism, Enoch mentions a number of times that there is something distinctively first-personal about deliberation. For instance:

Notice that deliberation is not merely the process of trying to find out what the decision that it makes most sense for one to make is. Rather, deliberation consists in trying to find out what the decision is that it makes most sense for one to make, and then (trying to) make it. It is this latter part that gives deliberation *its first-personal nature*. (We can, after all, rather easily try to find out what the decision is that it makes most sense for someone else to make, but we cannot try to make it.) [My emphasis] (Enoch 2011: 73, fn 54)

This remark is somewhat obscure, and although Enoch mentions several times that it is crucial to his argument that we are talking about something distinctively 'first-personal', he never spells out quite what he means by this. Still, the idea is suggestive, and if nothing else, Enoch is surely right to

think that deciding what to do is not identical with merely working out what course of action would be best given *someone's* position (even, perhaps, if that someone happens to be you).¹⁷ In addition to such remarks, Enoch emphasizes that “the whole point of the argument [about the deliberative indispensability of normative truths] is the focus on the first-person, deliberative perspective” (Enoch 2011: 76), and that he is concerned only with what is necessary if one is to “(first-personally) engage in deliberation” (Enoch 2011: 79, fn. 72). For Enoch, the ‘first-person perspective’ is, if not equivalent to the ‘deliberative perspective’, then at least the perspective from within which deliberation is undertaken. And as we saw in section 3.2, he seems to endorse what I called a ‘stepping-back’ model of deliberation, according to which one surveys one’s desires as though from a distance, and decides whether or which to endorse with a normative judgment. So, for Enoch, the ‘first-person perspective’ must be one that allows such stepping-back.

Now let’s turn to Nagel, whom Enoch cites as the inspiration for his own analysis of deliberation. Again, we have already seen that Nagel, too, endorses a stepping-back model of deliberation. But notice that he goes on to say this:

[T]he standpoint from which one assesses one’s choices after this step back is *not just first-personal*. One is suddenly in the position of judging what one ought to do, against the background of all one’s desires and beliefs, in a way *that does not merely flow from these desires and beliefs but operates on them*—by an assessment that should enable anyone else also to see what the right thing is for you to do against this background. (Nagel 1997: 110) [emphasis mine]

For Nagel, then, the process of ‘stepping-back’ is to be understood as a step *away* from (what he calls) the ‘first-person perspective’. Here, the ‘first-person perspective’ seems to mean (something like) the perspective constituted by all of the practical psychological attitudes I personally happen to

¹⁷ Enoch seems to hold that what one does in deliberation is the latter activity *plus* something else—the actual making of the decision *to do* what it makes most sense for this person (me) to do—and this doesn’t seem right to me. Some of my work in the following chapter will indicate why I think such a picture must be wrong.

hold, and what it is opposed to is the perspective from which one views things more ‘objectively’ or ‘impersonally’ (elsewhere, Nagel speaks of this as “transcend[ing] the subjective standpoint of our own desires” (Nagel 1986: 141). Nagel thus contrasts his model of how deliberation should be understood with that of Bernard Williams, who, as Nagel explains:

argues that reflective practical reason, unlike reflective theoretical reason, *always remains first-personal*. [According to Williams] one is always trying to answer the question “What shall (or should) *I* do?” and the answer must derive from something internal to what he calls one’s ‘motivational set’. (Nagel 1997: 112) [first emphasis mine]

And this, indeed, seems to be how Williams understands his own position, including the connection Nagel draws between the notion of the ‘first person’, on the one hand, and the person’s individual set of (practical) attitudes or ‘subjective standpoint’ on the other (see, for instance, (Williams 1985: 21 and 67)).

So far, then, we have Enoch who endorses a stepping-back picture of deliberation, thinks of this activity as characteristic of the ‘first-person perspective’, and believes value realism follows from within this perspective; Nagel, who also endorses a stepping-back picture of deliberation, thinks of this activity as stepping *out of* the ‘first-person perspective’, and is also (more or less) a realist; and Williams, who shares Nagel’s conception of the first-person perspective as constituted by a person’s individual attitudes, but thinks that deliberation must proceed precisely *from* this subjective material, and as a result ends up (or at least has often been read as being—see my Chapter 1, section 3.5) an anti-realist.

Now consider Korsgaard, who is described by Arpaly as holding a position “that focuses exclusively on the first-person conscious perspective” (Arpaly 2003: 19). Korsgaard, of course, endorses the ‘stepping-back’ picture of deliberation along with Enoch and Nagel, and like Enoch, this activity lies for her at the heart of how she understands what it means to speak and think about

the ‘first-person perspective’.¹⁸ However, while Enoch holds that this very stepping-back itself entails a commitment to realism, Korsgaard not only rejects realism but indeed does so in large part out of a concern that it cannot justify normativity’s demands for us *first-personally*: “the answer we need is really the first-person answer, the one that satisfies *us* when we ask *ourselves* the normative question” (1996: 17), and realism, she claims, cannot provide such satisfaction. For Korsgaard, then, we may say that the first-person perspective (construed as the perspective we take when we ‘step back’ from our desires) requires that we endorse anti-realism.

Street takes up this idea from Korsgaard, and claims explicitly that the first-person perspective (which she seems to use interchangeably with “the practical perspective”) entails anti-realism (Street 2010). And like Korsgaard, she takes the making of normative judgements, as opposed to mere desires, to provide the foundations of normativity and its authority. Despite these similarities to Korsgaard, it is not at all clear that we ought to attribute a ‘stepping-back’ picture to Street. She does not speak of any kind of neutral or desire-less or rational self who must perform some act of distancing from her own attitudes in order to decide what to do. Instead, as she herself notes ((2008), (2012)) her position turns out to be rather similar to Williams’ account. For Street, the ‘first-person perspective’ seems to mean something like ‘the particular and idiosyncratic set of

¹⁸ Given how frequently her position is characterized by others in terms of its focus on ‘the first-person perspective’ (along with Arpaly, see (FitzPatrick 2005) and the interview with Korsgaard in Pauer-Studer (2002)), it is interesting to notice how infrequently she herself is inclined to use exactly this terminology. The most prominent example of her doing so is perhaps that at the beginning of *The Sources of Normativity* that I’ll discuss below. In any case, since she has never (to my knowledge) challenged such a characterization of her view, and since it seems apt enough, I’ll join the mainstream in articulating her position in such terms.

normative judgments a person happens to have', and it is on the basis of these that one must deliberate about what to do.¹⁹

Between Enoch, Nagel, Williams, Korsgaard, and Street, then, we have a constellation of views, all of which claim to make some use of the 'first-person perspective'. We can compare these authors to one another with reference to three key topics: (1) how we are to understand what is meant by the 'first-person perspective' (is it constituted by our ground-level attitudes, or by the ability to step back from those very attitudes?);²⁰ (2) how we are to understand the relationship between deliberation and the first-person perspective (does deliberation take place from within or beyond that perspective?); (3) how we are to understand the relationship between our attitudes and normative reality (realism or anti-realism about value?). As we saw above, the answers from our authors display all kinds of fascinating pairwise similarities and differences. Despite this, the central notion of the 'first-person perspective' is often treated as though its meaning and significance for thought about deliberation and about metaethical questions is obvious enough to not require much, or any, elaboration.

The above overview suggests that this is not so. And this means that there is a general motivation for metaethicists to be interested in further investigation of what might be meant by the notion of a 'first-person perspective', in particular as it relates to deliberation, and to the question of value's reality. In the following chapters, I'll develop an account of one particular facet of the first-person perspective, which I'll call 'transparency', and show how this feature of mindedness has significant consequences for metaethical thought.

¹⁹ Roughly, we may say that although Street shares the neo-Kantian's interest in normative judgments, the actual role she gives these judgments in her account is ultimately basically the same as that given to desire by the neo-Humean.

²⁰ In Chapter 4, I'll argue that neither of these pictures is the right one.

5. CONCLUSION AND RECAP

In this chapter I have set up the question that will provide the basis of my argument for value realism: the question, that is, of what you're up to when you're deciding, on some occasion, what to do. The answer that I shall eventually give is that what you are doing is looking outwards, towards the world as itself the source of realist value, as something that will guide you in your deliberations.

I began by clarifying the scope of my topic, noting that it is one sense narrower, and in another broader, than some familiar philosophical discussion of the concepts of 'deliberation' and 'reflection' have suggested. On the one hand, I am thinking of 'deciding what to do' as a conscious and explicitly undertaken activity, and of all of our actions taken together, those that result from such a process are the exception rather than the rule; for the most part, we are perfectly capable of doing all sorts of things, including things of great importance and complexity, without undertaking any such conscious deliberation. On the other hand, I insisted that I wanted to allow that 'deciding what to do' in the relevant sense may come in quite a wide variety of forms. In particular, I rejected the idea that it ought to be limited either to a 'calculative' model (on which one identifies considerations for and against some course of action and performs some sort of weighing activity to work out which options is best, or good enough) or to a 'stepping back' model (on which one distances oneself from certain aspects of one's psyche—typically, but not always, one's desires—and makes a decision both in light of these psychic contents but at the same time immune, at least to some degree, to their power as a motive force).

Having clarified in this way the phenomenon in which I'm interested, I then identified an important feature of this phenomenon that I would like to be able to make sense of: the idea that, in deliberating about what to do, there is something that you are trying to get in some sense *right*. I

offered an analysis of a recent argument given by David Enoch, in which he claims (roughly) that value realism is needed to make sense of this feature of deliberation. In the course of explaining why I find Enoch's argument inadequate, I was able to identify the two key challenges that I must meet if I am to make my own argument successfully.

The first is the challenge posed by the anti-realist schema that I developed in section 3.2. I proposed, on behalf of the (generic) anti-realist, that we can make sense of the getting-something-right feature of deliberation with reference to some psychological attitude (or set of attitudes) of mine: it is simply the fact that I have certain ends, or desires, or that I have made certain normative judgments, or that I have a particular motivational set, that makes it the case that some ways of concluding my deliberations are better than others. On such a model, the 'rightness' that I am trying to achieve in deliberation *does* stand independent of the process of deliberation in the way we are after (it's not the case that any way of concluding my deliberations will do as well as any other, and getting things 'wrong' is a real possibility). And yet it remains the case, on this account, that the ultimate source of this rightness is not some value out there in the world, but rather, as we might want to put it, something 'in my head'.

Faced with this model and the plethora of opportunities it offers to the anti-realist, my task in what follows is to show that you *cannot*, in fact, understand your own process of deciding what to do purely in the terms laid out by the anti-realist schema. In general, you cannot suppose that what would ultimately make some course of action better than another is (merely) the fact that you have some psychological attitude or set of attitudes (nor such 'in combination with all the non-normative facts'). Instead, you must suppose that there exists an external source of value—value out there in the world—which is the ultimate source of some answers to your deliberative question counting as better than others.

If I can achieve this first task, it immediately gives rise to the second challenge that I face. For in establishing that you *cannot* suppose that some psychological attitude of yours is what renders some results of your deliberations better than others, I seem to face an obvious and broad set of counterexamples: those in which I make my decision, and am quite right to do so (perhaps even irrational *not* to) on the basis of some particular desire or preference that I happen to have. The significance of these sorts of cases—in which I choose the peanut butter cup ice cream because I happen to like it best, for example—have often been downplayed in the work of moral philosophers and metaethicists, but from my point of view they appear to threaten a *reductio* to my argument.

These, then, are the two tasks I face in making my case. My goal is to defend a thoroughly realist account that nevertheless is able to do justice to the compellingness of the anti-realist's idea that when it comes to deciding what to do, my own personal attitudes do matter, and indeed often matter a great deal. It is just that they do not matter in the *way* that the anti-realist would have them matter; they are not themselves the *source* of the rightness of some courses of action over others.

Finally, I offered a brief overview of the wide variety of different ways that the thinkers I've written about in this chapter each make use of the notion of the 'first-person perspective' within their views, and usually without making the details of their usage very explicit. This overview was intended to give a more general motivation for an investigation of the first-person perspective—one that takes place in light of the particular concerns that we have as metaethicists. This will be my topic in the following chapter.

CHAPTER 3: ALIENATION AND TRANSPARENCY

...in making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world.

Gareth Evans¹

“Oh, I'm sure that eventually someone will fall in love with me, even though my therapist has convinced me that unconsciously I believe it's impossible that anyone should.”

'Harry', via David Finkelstein²

1. INTRODUCTION

At the end of Chapter 2, I set a task for myself: I need to show that no version of the anti-realist schema for deliberation can be adequate from the point of view of the person doing the deliberation. I also gave a more general motivation for an investigation into the idea of 'the first-person perspective', one which is undertaken with a view to the concerns that we have as metaethicists in particular.

In this chapter I shall turn to a certain recent tradition in the philosophy of mind in order to address these two points together. The school of thought that I am drawing on is interested primarily in the so-called 'first-personal authority' with which we seem able to speak about our own mental states, or (as this is sometimes glossed) a distinctively first-personal form of 'self-knowledge' that we seem to possess. First-personal authority and self-knowledge are not themselves the locus of my concern here, and my aim is not to draw any definitive conclusions about these topics. However,

¹ (Evans 1982: 225).

² (Finkelstein 2003: 115).

as I shall try to show, this debate within the philosophy of mind, and in particular the work of Richard Moran, will allow me to bring into view certain key insights that have interesting metaethical consequences.

I shall proceed as follows. I'll begin with an overview of the context within which Moran is working and the questions he is seeking to answer. I then outline Moran's position alongside another account offered by David Finkelstein, who shares many of Moran's concerns and interests but draws different conclusions and raises some important criticisms of Moran's approach. (I'm using Finkelstein as my foil for Moran because I believe that his criticisms are the ones that are most relevant to the use that I want to make of Moran's work.) I'll then introduce the topic of 'alienation', which will provide the basis for my own arguments concerning the nature of the first-person perspective, and go on to show how we can extract from Moran's account the idea that I need to make my case for realism—the concept of 'transparency'.

2. TWO ACCOUNTS OF THE FIRST-PERSON

There are many different kinds of things that I may know about myself. I know, for instance, my height, my eye color, and my cell phone number. Within the broad set of things that we may speak of as my 'self-knowledge', there seems to be a subset that I am able to know in a distinctively first-personal way. These include such things as: whether I'm hungry right now; who I think will be the next Prime Minister of Great Britain; who I hope will be the next Prime Minister of Great Britain; and what I plan to do this evening.

Two interesting features of this sort of self-knowledge are what have come to be known as its *authority* and *immediacy*. The former points towards what Matthew Boyle speaks of as the "apparent entitlement to some sort of deference" that attaches to a person's own self-ascriptions of

such mental states (Boyle 2009: 136); as Finkelstein puts it, “If you want to know what I think, feel, imagine, or intend, I am a good person—indeed, usually the best person—to ask.[...] When people don’t accept my mental state self-ascriptions at face value, it is generally because they take me to be insincere rather than mistaken” (Finkelstein 2003: 9). And, perhaps surprisingly, this authority seems to go hand-in-hand with the so-called *immediacy* of my knowledge—the fact that I seem able to answer your question without having to look towards any of the sort of evidence that I might consider were I answering the question about someone else. I don’t have to look towards my own behaviors, or reflect upon my own previous remarks on the topic, for instance. Indeed, if I tell you what I think about some matter, although the question ‘how do you know?’ has a natural application to the *content* of my thought, it seems quite unclear what sort of response you would be after if you wanted to hear how I know that *such is* my thought.

The two accounts that will be the focus of my attention here—Moran’s ‘transparency-based’ and Finkelstein’s ‘expressivist’ approach—can be situated within the history of thought that is concerned with the question of how we should understand this distinctively first-personal form of self-knowledge.³ Within this history, one traditional sort of picture has a person’s mental states taking place within a sort of inner realm, or ‘inner theatre’, to which she (and she alone) has privileged access. On such an account, a person knows her own mind because she is able to do

³ My précis of Moran’s position is based primarily on his (2001), and to a lesser extent his (2003). My précis of Finkelstein’s position is based on his (2003). For Finkelstein’s criticisms of Moran’s position, I make use of the Postscript to (Finkelstein 2003), and (Finkelstein 2012). I have set these two accounts within the context of a history of attempts to understand self-*knowledge*, but it is worth noting that Finkelstein is less interested than Moran in holding onto the idea that what is at stake here is a sort of *knowledge* (see (Finkelstein 2003: 148–52)). In addition, the most significant criticisms of Moran’s account as a whole concern the question of whether or not it can actually provide the epistemic goods (see for instance (O’Brien 2003), (Gertler 2011: 180ff), as well as Finkelstein’s arguments towards the end of his (2012), which I do not make use of here.) This debate is interesting in its own right, but, as I shall explain, it does not impact on the use that I intend to make of Moran’s ideas.

something akin to seeing or otherwise perceiving her mental contents. In the most extreme version of such a picture, typically attributed to Descartes, this access is complete and infallible, with the contents of the mind assumed to be entirely transparent to the subject. This version of the account will find few defenders nowadays; quite aside from concerns about metaphysical extravagance, it stands in conflict with familiar findings of empirical psychology, and indeed with the ordinary experience of coming to realize, for example, that one's motives in doing something or other were not quite what one had thought they were. However, even where the account is modified, so that the mind's inward-directed gaze is allowed to be fallible and incomplete, it faces a set of criticisms developed by a number of philosophers, including Moran and Finkelstein.

Moran and Finkelstein are united in their aim to overcome all such pseudo-perceptual or 'inner sense' models of first-personal self-knowledge. Indeed, for both Moran and Finkelstein, the most fundamental problem with these sorts of approaches applies even more broadly than this, to any model of self-knowledge on which a person's mental contents are construed as a set of facts that she is able somehow to 'detect', or to which she has some form of special access (I'll follow Finkelstein in labeling such positions 'detectivist'; Moran speaks instead of a problematically 'theoretical' paradigm of self-knowledge, and I take it that these two terms cover at least roughly the same ground for these two authors). In short, the problem that both Moran and Finkelstein have with this broad class of positions is that they all end up portraying a person's mental life as fundamentally too external to her. However the details are worked out, the access a person has to her own thoughts, beliefs, intentions, pains, and so on, will still look like an access that she could in principle have to the thoughts, beliefs, intentions, and pains, of another person. As Moran puts it, any such model ends up portraying self-knowledge as a sort of "self-telepathy" (Moran 2001: 91).

Moran and Finkelstein are also united in rejecting a second sort of account, which stands in many ways at the opposite extreme to the detectivist paradigm. Where the detectivist has my mental contents existing as a realm of independent facts that I am able to access and report on, the second approach, which Finkelstein labels ‘constitutivism’, in effect closes any gap between what I say about my mental states and what the facts (in as much as this notion can get a foothold within such an account) about those mental states are. (To illustrate, one version of such an approach understands first-personal authority as something like a social concession that we grant to one another, so that what a person says about their mental states is just, as a rule of thumb, allowed to stand unquestioned; another suggests that the facts about my mental states are just constituted by what I say about them.)

Moran and Finkelstein thus both seek to make sense of the first-personal authority with which we speak about our own mental states while avoiding the pitfalls of both detectivist and constitutivist approaches; this is the philosophical context within which their accounts are originally situated. I want to be clear, however, that Moran’s and Finkelstein’s philosophical aims are not my philosophical aims, and their interlocutors are not my interlocutors. For Moran’s and Finkelstein’s critiques of detectivism and constitutivism, I refer the reader to their own work (especially (Moran 2001) and (Finkelstein 2003)); I shall not be assuming that these critiques are definitive nor that either Moran or Finkelstein’s own positive accounts, as outlined below, can necessarily do the best job of answering their own central questions about self-knowledge and first-person authority. Instead, my claim is that some of the ideas that Moran develops in the course of attempting to answer these questions—or, more precisely, these ideas as refined and clarified in light of some criticisms from Finkelstein—can provide the basis for my own account of a particular facet of the first-person perspective that will prove, I shall argue, to be metaethically fruitful.

Moran's attempt to make sense of first-personal self-knowledge begins from what he (following Roy Edgley)⁴ calls 'transparency'. Here is Gareth Evans' now-canonical illustration of the phenomenon in question:

[I]n making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?'. (Evans 1982: 225)

Moran's aim is to show how this structure—in which a question about a person's psychological state is, for that person, 'transparent to' a question that appears to have a different subject-matter, namely that of how things stand in the world—can help us to understand first-personal self-knowledge.⁵

Moran's claim is that my knowledge of and ability to speak with authority about (certain) aspects of my mental life, including my beliefs, intentions, fears, (some sorts of) desires, and so on arises from my ability to *make up* my mind—which is to say: to come to believe something on the basis of considering the evidence, or to come to want something on the basis of reflection on what's worth wanting, and so on. Finkelstein glosses this (I raise an objection to this way of thinking about transparency in section 4) as Moran's claiming that the question of *what* I think (believe, intend, desire, fear, etc) is, for me, 'transparent to' the question of what I *ought* to think (believe, intend, desire, fear, etc).

The notion of transparency, understood in this way, is what Moran thinks will allow us to understand the distinctively first-personal way in which I seem able to know the contents of my own

⁴ (Edgley 1969), cited in (Moran 2001: 60).

⁵ In taking this approach, Moran is following Shoemaker's attempts to show how transparency can be extended beyond the case of belief to cover other psychological attitudes. See for example Shoemaker (Shoemaker 1994); for some criticisms of Shoemaker's arguments, see Finkelstein (Finkelstein 1999) and (Finkelstein 2003: 157–58).

mind, without the need for anything like the inward-glimpse of the pseudo-perceptual model. Since we do not usually suppose that someone needs some special ‘inner sense’ to determine, say, whether some claim about the world is true, then if that person’s ability to speak about her own mind ultimately requires no *more* than this ordinary ability (to look outwards and make up her mind about some matter), self-knowledge should begin to look less mysterious.

While Moran’s account begins from the notion of transparency, Finkelstein’s hope is to be able to make sense of our ability to speak about our mental contents by thinking about the phenomenon of expressive behavior. Our ability to self-ascribe mental states should be understood, thinks Finkelstein, as on a continuum with, and fundamentally akin to, other ways that we may express those mental states. A child may express sadness or pain by crying; later, she learns also to express the inner state precisely by *saying* that she is sad or in pain. Finkelstein’s focus on expression is not supposed to undermine the idea that, in stating (say) that I intend to stay home tonight, I am saying something *true*—and thereby giving you some information that you can make use of in planning your own evening.⁶

Part of the appeal of this account is that in focusing on the connections between comparatively primitive expressive behavior (crying, smiling) and the more sophisticated ability to express a mental state by self-ascribing it, Finkelstein renders the latter less mysterious. In particular, since we do not usually suppose that a child (or indeed an adult) must ‘look inside’ to check whether

⁶ In his book, Finkelstein reserves the term ‘expressivist’ for accounts that understand first-personal authority in terms of expression to the exclusion of truth-aptness, and thus does not count as an ‘expressivist in this sense, though he later abandons this restriction and allows that he be counted as an expressivist in a broader sense. Notice that this means that Finkelstein’s ‘expressivist’ account of first-personal authority is not analogous to anti-realist metaethical expressivism (see my discussion in Chapter 1).

they are happy before spontaneously smiling, the pressure to suppose that this is how things work when a person speaks about their mental states is much reduced.

In rough outline, then, we may think of both Moran and Finkelstein as aiming to overcome the temptation towards a quasi-perceptual picture of the knowledge we have of our own minds by connecting this knowledge to some other phenomenon that is less worrisome and that does not seem to require the inward-directed gaze. For Moran, this is the ability to think about how things stand out in the world; for Finkelstein, it is expressive behavior.

This, then, is the philosophical context within which these two accounts are situated. The question that launches and shapes the tradition of thought to which they belong is something like: how should we understand the ability we appear to have to know, or in any case to speak with a distinctive authority about, the contents of our own minds?

3. MAKING SENSE OF ALIENATION

3.1 FIRST-PERSON, THIRD-PERSON, ALIENATION

As I have noted, *this* is not my question, and I do not pretend to be able to answer it here. Instead, my interest is in something that is certainly intimately connected to these questions, but that stands as it were a little off to one side from them: I am interested in a certain kind of contrast that can be drawn between the case in which one's relationship to one's own mental state can be characterized as 'first-personal' in the fully-fledged sense, and, on the other hand, the case in which this

relationship echoes in certain respects the ways one typically relates to the mental lives of other people (or indeed to other parts of the external world).⁷

When you meet another person in the world, you encounter them as coming already equipped, as it were, with their own complete set of psychological attitudes. The fact of these attitudes stands just as real and independent of you and *your* attitudes as does any other ordinary fact about the world.⁸ If you meet Tom, and he believes that global warming isn't happening, then whether you like it or not, this is a fact about Tom just as his height is a fact about him, and just as the location of the tea bags in his kitchen on the topmost shelf is a fact about his apartment—and all of these are things that you may have to take into account, and perhaps work around, as you negotiate your way through the world.

When it comes to Tom's attitude towards climate change, you can in a reasonably straightforward manner treat the fact of this attitude as separate from the question of how things stand in the part of the world that his attitude is about, thinking, for instance, that though Tom *believes* that global warming isn't happening, still the fact is that it *is* happening. Of course, you might try to change Tom's mind by arguing with him (or through some other more covert means). Still, the fact of Tom's belief prior to your attempting to change his mind is nonetheless something that you encounter as merely part of the way the world happens to be.

This is not, however, how things typically look when it comes to the matter of your own beliefs—nor, as I shall go on suggest, a great many of your other attitudes. Of Tom, you may think:

⁷ *One* aspect of this relationship is the knowledge (or lack of it) one has of, and the authority (or lack of it) with which one speaks about these mental states, but the question of how to understand the distinction that I am interested in is not equivalent to, and need not depend upon, the question of *how* it is that we are able to have first-personal self-knowledge or to speak with first-personal authority in relation to our mental states.

⁸ On the topic of concerns about the naturalistic acceptability of intentional states generally, see section 3.1 of my Chapter 1.

“it’s so tiresome that he won’t look seriously at the scientific evidence for climate change, and remains so stuck in his convictions” (you might add: “but he’s otherwise basically a nice guy, so we’re still friends”). To have this sort of attitude about a belief of your own, however, would be extremely odd—“yes, it’s really tiresome that I’m stuck believing that climate change is a liberal conspiracy, but I don’t seem to be changing my mind any time soon.” Similarly, while there seems to be no difficulty in thinking of Tom: “Tom believes climate change isn’t happening, but of course it is,” such a thought about oneself (“I believe climate change isn’t happening, but of course it is”) would be strikingly peculiar. (It is, of course, an instance of ‘Moore’s Paradox’.)

Indeed, it may be tempting to suppose that the thinking of such a thought would be not merely peculiar, but in fact impossible. On closer inspection, however, this seems to be not quite correct. There are contexts in which it *would* make sense to think (or say) something of this kind. Consider, for instance, the following case:

Imagine someone—call him Harry—who says: “my therapist tells me that I unconsciously believe no one could ever fall in love with me, and she’s generally right about such things, so I suppose I must have this belief.” [...] But we can imagine that Harry holds no such conscious belief. When asked about his future, Harry says, “Oh, I’m sure that eventually someone will fall in love with me, even though my therapist has convinced me that unconsciously I believe it’s impossible that anyone should.” This is a perfectly intelligible remark. (Finkelstein 2003: 115)

The case is Finkelstein’s, and it is introduced as part of his attempt to show that the distinction between what we may describe as ‘conscious’ and ‘unconscious’ attitudes can best be understood in terms of his own expressivist account: an attitude is conscious, thinks Finkelstein, to the extent that the person can express the attitude through self-ascribing it. Harry can successfully self-ascribe his state (he can agree or assert that he has the belief in question), and no doubt it is also expressed in various behaviors and remarks (something must have tipped his analyst off, after all), but what he is

not able to do—at least so long as the belief remains unconscious—is to express the belief *by* saying that he has it.⁹

Here, again, we find another interesting parallel between Finkelstein and Moran’s accounts, which is their treatment of ‘defective’ cases—cases in which first-personal self-knowledge or authority falls short in some way. Both Finkelstein and Moran make use of such cases to highlight the way in which their accounts (as opposed to, in particular, a quasi-perceptual or detectivist picture) can succeed in capturing a genuinely first-personal phenomenon. For each aims to present an account on which various defective self-relations can be understood *as* defective—and their opponents are revealed as attempting to analyze the ordinary case, the case in which all goes well, in a way that turns out to be more reminiscent of these defective cases.

Although it is not detectivism that I am seeking to overcome, I shall be taking on something of the same strategy in developing my own account. That is, I share with both Moran and Finkelstein the idea that we might be able to arrive at a clearer understanding of what is involved in something’s being ‘first-personal’ by thinking about what has gone missing or gone wrong in cases where the person’s self-relationship appears to have failed in certain respects at being fully first-personal, and resembles instead the sort of relationship that one might bear to something more external to oneself—the attitude of another person, or some other part of the external world. I shall refer to this phenomenon as ‘alienation’, and speak interchangeably of an attitude’s being alienated, and of a person’s being alienated from an attitude. The cornerstone of my argument, then, will be the claim that we need something like Moran’s concept of transparency in order to make sense of (certain cases of) alienation.

⁹ For a nice explication of the idea that such a phenomenon might come in degrees, see Finkelstein’s discussion of Golde and Tevye singing “Do You Love Me?” in *Fiddler on the Roof* (Finkelstein 2003: 122ff).

Moran himself makes a version of this claim regarding a case that bears striking (and, so far as I know, accidental) similarities to Finkelstein's case of Harry. Moran has us consider a psychoanalytic patient who unconsciously "feels anger at the dead parent for having abandoned her, or who feels betrayed or deprived of something by another child" (Moran 2001: 85). As in the case of Harry, this patient may come to know this fact herself—"[s]he might become thoroughly convinced, both from the constructions of the analyst, as well as from her own appreciation of the evidence, that this attitude must indeed be attributed to her"—without this sufficing for the attitude to count as an instance of the ordinary first-personal self-knowledge that Moran's account aims to capture. Now, according to Moran, what has gone wrong here is that the structure of transparency is not present. For although the patient may (like Harry) know and be willing to assert that she has the attitude in question (in this case, the feeling of betrayal), still:

at the same time, when she reflects on the world-directed question itself, whether she has indeed been betrayed by this person, she may find that the answer is no or can't be settled one way or the other. [...] So, transparency fails because she cannot learn of this attitude of hers by reflection on the object of that attitude. (Moran 2001: 85)

Moran, then, claims that what is going on in this case of alienation is a failure of transparency, and that transparency must therefore be given a central role within our understanding of the nature of the fully first-personal situation. And this suggestion is a very appealing one for my purposes. But I face a problem here, for Moran's explanation, although perhaps compelling enough on its own terms, is not the only possible explanation that we have available to us. For as we saw previously, Finkelstein has offered us a different analysis of what appears to be a structurally similar case. So more work needs to be done if I am to establish my own claim about the significance of (something like) Moranian transparency for our first-personal relationship to our own mental states.

3.2 TWO KINDS OF ALIENATION

As it turns out, I think that Moran is actually at least partly wrong about this case, and understanding why he is wrong will help us to make progress. For this purpose, I want to develop an objection raised by Finkelstein against Moran.¹⁰ In bringing this objection. Finkelstein is not thinking in particular of Moran's treatment of the case of the psychoanalytic patient, nor of the idea that transparency might help us to understand certain cases of alienation. Rather, he is thinking of Moran's broader claim that transparency can help us to make sense of the authority with which we are typically able to speak about our own mental states.

Finkelstein's objection, in short, is that we are able to speak with a distinctively first-personal authority about a great many features of our mental lives that do not seem to follow the structure of transparency. These include such things as sensations and pains, attitudes one held in the past, attitudes which are neither rationally prohibited nor rationally required (I'll return to this topic in section 4.2 of this chapter, and then in a different manner in Chapter 5), and—which will be my immediate focus here—attitudes that the reasons speak against ((Finkelstein 2003: 161–68); I also draw on Finkelstein's recap of this objection in his (2012)). Consider the following example:

Max has a spider phobia. He is terrified of spiders—spiders in general, and the one on the pillow beside him in particular. Max doesn't take this fear to be rational; he knows that he suffers from a phobia. Still, when he says, "Get that thing away from me; I'm really afraid of it!", he is speaking (or shouting) about his own state of mind with first-person authority. (Finkelstein 2012: 105)

¹⁰ In a later paper, written partly in response to defense of Moran's position made by Matthew Boyle, Finkelstein offers what I take to be a new objection to Moran. It is the original objection from the book that I shall be working with here, though I will also make use of Finkelstein's recap of that objection as provided in the later paper. I'll develop a second aspect of the book's objection later in this chapter. The further objection in the later paper applies to a particular reading of what Moran's position, one that I am not sure is quite right, and concerns in particular Moran's ability to answer his self-imposed question of how it is one knows one's own mental states. Since this is not the question I am claiming to answer here, Finkelstein's later objection can be put to one side.

This is surely right. And it suggests something about the case of Moran's psychoanalytic patient, which is that her unconscious mental state might become conscious (just as Max's fear is clearly conscious—though it may of course have its roots and causes in his unconscious life) *without* it following the structure of transparency that Moran had claimed was what was missing in this case. For certain sorts of therapeutic work may put the patient in a position to access the feeling of betrayal that will allow her to say, angrily—and with the distinctively first-personal authority that she lacked while the attitude remained unconscious—something like: “I feel so betrayed by my mother!”. Here, we need no longer think of this statement as a mere report of an attitude the patient knows, on the basis of evidence, that she happens to have; rather, the statement is itself an expression of that very feeling. And yet, for all this, the analysand may believe (or even know) that there was no betrayal here, nothing to feel betrayed about.

To emphasize this aspect, we can suppose that the feeling of betrayal is a response to something that could not possibly have been a betrayal—say, the accidental death of the mother, at a time when the daughter was especially vulnerable. It's not that we can make no sense of a feeling of betrayal here, for one can imagine experiencing the death as a sort of abandonment, and abandonment at a vulnerable time can certainly be a betrayal. So certainly the feeling of betrayal makes a kind of sense to us, although we do not thereby think of it as rational, and we will not be inclined to think of the accidental death as an abandoning or a betrayal of any kind. And the bereaved daughter may be fully aware of all of this, and yet *still* find herself with a feeling of betrayal—a feeling that was previously unconscious and known only indirectly, but is now fully conscious, and something about which she can speak with first-person authority.

What this means, I think, is that Moran's analysis of his case of alienation cannot be quite right. In the story I have told above, the attitude has transformed from something that is not

essentially first-personal (for prior to the attitude's coming to consciousness, her knowledge of it came via the sorts of assessment of evidence that may likewise be at work when one comes to know the attitude of another person) to something that *is* essentially first-personal; once the attitude has been brought to consciousness, the relationship she bears to the attitude seems to be of the kind that one can only bear to one's own attitudes. And yet this transformation has occurred *without* the attitude's coming to display the structure of transparency. Transparency, then, seems not to have been the essential component that was missing in this case. For the patient's attitude is now conscious; she is able to speak about it with distinctively first-personal authority (which may or may not amount to the same as its being conscious); and it seems that Finkelstein's expressivist account tracks, in a way that Moran's transparency-based account does not, the important shift that has occurred for the patient.

And yet there is, it seems to me, more to be said here. The patient's attitude is now conscious, but she remains in an odd position; things are surely somewhat out of joint for her. It makes sense to say, as I did above, that the patient 'finds herself' with the feeling of betrayal, a formulation which echoes the way in which one 'encounters' (cf. section 3.1 of this chapter) another person in the world, already equipped with their own set of psychological attitudes. This signals a sense in which the bereaved may still experience her own feeling of betrayal as something that is 'alien' to her.

Here, the 'alienation' that is at work may be characterized by an experience of the attitude as an external or oppressive force. The patient may experience her own possession of the attitude as distressing and nonsensical in various respects: instead of grieving her beloved mother, she is instead plagued by an inexplicable anger towards her, an anger that remains disturbingly unresponsive to her

own knowledge that it is unwarranted.¹¹ Or she may think, of her own attitude, that the fact of its existence is tiresome, just as you found the fact of Tom's beliefs about climate change tiresome. She may even be moved to utter such things as: "My mother's death certainly was not an act of betrayal—and yet at the same time it seems I can't help but think of it as such", and we may suppose now that the attitude she self-ascribes in the second part of this sentence is one which is conscious, and expressed through this very self-ascription.

What this suggests, I think, is that there is more than one way in which a person may experience herself as 'alienated' from some particular attitude—there are, that is, multiple ways in which she may fall short of a fully first-personal relationship to her own attitude. And these different forms of alienation can come apart from one another, such that a person may relate to her attitude in a distinctively first-personal manner in one respect, while at the same time relating to it as though it were not her own in another respect. We may say, provisionally, that a person is *expressively* alienated from her attitude when she is unable to express the attitude by self-ascribing it, and that she is *rationally* alienated from her attitude when the attitude seems to her out-of-joint with, and unresponsive to, her sense of how things stand in the world—when it deviates from or stands in conflict with the attitude that she takes the situation actually to call for or to warrant.

¹¹ Of course, it makes a difference here that our case is a sort of 'toy' case of the practice of psychoanalysis, one that has been simplified perhaps to the point of parody. In a real situation of this sort, things will be more complicated; there may be reason (for instance) for the ensuing analysis to explore whether the anger towards the mother has some other source that has shifted its object onto the mother's death. And if this is so, this alternative source for the anger may be one that does—or can be brought to—follow the structure of transparency. (Or, of course, it may not; it would surely be a mistake to suppose that the logic of the unconscious works in terms of a straightforwardly rational pattern of connections that is merely hidden from view until its symbolic code is deciphered within analysis.) In any case, the point is that my discussion here (and probably the same is true of Moran and Finkelstein) is not really intended to tell us anything much about *psychoanalysis*, nor about the notion of the unconscious as it is understood within psychoanalytic practice. Rather, these toy cases are intended to draw on our familiarity with a certain trope in order to model the structures of the first-person perspective that I am interested in.

4. FURTHER DEFENSE OF TRANSPARENCY

4.1 TRANSPARENCY ITSELF AS A FORM OF ALIENATION?

So far, I have suggested that although Moran is wrong to think that transparency is the key component missing for the psychoanalytic patient whose attitude is unconscious, nonetheless he is correct in thinking that the absence of transparency *does* constitute a falling-away from a fully first-personal relationship to one's own attitude—it constitutes, that is, a form of alienation. In the rest of this chapter, I'll continue to develop and defend this idea, while along the way reworking and clarifying the notion of transparency for my purposes.

To this end, let me consider another, later, criticism raised by Finkelstein against Moran's account (Finkelstein 2012). Finkelstein's new claim is that the concept of transparency itself actually describes a form of alienation. If Finkelstein is right about this, it raises a serious problem for my account, since I am claiming that transparency is a key component of one's bearing a genuinely first-personal—i.e. *not* alienated—relationship to one's own mind.

Finkelstein makes his case by drawing on a later characterization that Moran offers of the role of transparency in his account (Moran 2003). In responding to some (different) criticisms raised by Shoemaker, Moran casts his account as a way of trying to make sense of an aspect of the phenomenon of transparency that is otherwise, he claims, very puzzling. What is puzzling is that the question about my state of mind ('do you believe that P?') that I treat as 'transparent to' the question about the world ('does the evidence favor P?') is about a totally different subject matter—it's about the empirical fact of my psychological state. Reflection on this should lead us to wonder, thinks Moran, what on earth entitles me to answer the former question by answering the latter:

What right have I to think that my reflection on the reasons in favor of P (which is one subject-matter) has anything to do with the question of what my actual *belief* about P is (which is quite a different subject matter)? Without a reply to this challenge, I don't have any right to answer the question that asks what my belief is by reflection on the reasons in favor of an answer concerning the state of the weather. And then my thought at this point is: I *would* have a right to assume that my reflection on the reasons in favor of rain provided an answer to the question of what my belief about the rain is, if I could assume that *what* my belief here is was something determined by the conclusion of my reflection on those reasons. An assumption of this sort would provide the right sort of link between the two questions. (Moran 2003: 405)

This characterization of the position encourages Finkelstein to gloss Moran as claiming “that I can and do self-ascribe attitudes by: (1) reasoning about what my attitude ought to be and (2) avowing the conclusion of this reasoning as what my attitude *is*” (Finkelstein 2012: 108). On this interpretation of transparency, its role is to allow us to answer the question of how a person knows their own mental states in something like the following way: if I want to know what I think about something, I don't need to look *inwards* (as in the pseudo-perceptual model), for I can instead simply figure out what I *ought* to think, and then just assume that this is in fact what I *do* think.

Now, I actually think this is not what Moran is saying here, though it is true that some parts of his writings are suggestive of such a picture. For instance, in the context of the discussion of the psychoanalytic patient, Moran states that here “transparency fails because she cannot *learn* of this attitude of hers by reflection on the object of that attitude” ((Moran 2001: 85), my emphasis). The implied idea that transparency is the mechanism through which one comes to ‘learn’ what is going on in one's own mind is, I think, a mistake on Moran's part, suggesting as it does just the ‘theoretical’ model of self-knowledge that he is attempting to overcome. But in any case, since my concern here is to defend my own account of transparency, rather than to do Moran exegesis, it doesn't really matter whether Finkelstein's reading of this particular passage is accurate. What matters is that it gives us a particular model for understanding transparency that I would like to

avoid, namely one on which the point of transparency is that one comes to know what one thinks or feels by figuring out what one ought to think or feel, and then assuming that one's attitude matches this.

In the face of this model, Finkelstein claims that “insofar as someone needs to make this sort of assumption in order to self-ascribe an attitude, one is *alienated* from the attitude” (2012: 108).

As evidence, Finkelstein offers the following case:

Henry is a Navy SEAL, about to embark on a hazardous mission. His wife, who has gathered that he's being sent off to do *something* dangerous, asks him, “Are you afraid?”. He answers, in a flat tone of voice, “No”. She says: “Come on, Henry; you can tell me what you're feeling. We've talked about how dangerous your work is. Surely, you feel *some* fear”. Henry thinks about what he's liable to be doing for the next few days and says, “Well, I see that there are lots of good reasons for me to be afraid, so I suppose that at some level, I must be”. (Finkelstein 2012: 109)

If we suppose that Henry is right, that he *is* afraid at some level, then here we have a case in which someone seems to follow exactly the structure of thought sketched out above: he reflects on the reasons, draws a conclusion regarding which way the reasons point, and assumes that his attitude is in accordance with this. And, indeed, he turns out to be right; his feeling does reflect those reasons (and this fact, presumably, is not an accident). But this example, far from exemplifying the ordinary first-personal way that one may relate to and talk about one's own attitudes, seems instead to be a prime case of alienation.

Now it is worth noticing that Henry appears to suffer from what I have called ‘expressive’ alienation, and I have already allowed that this sort of alienation should not be understood as a failure of transparency. This may make it seem that we have no *new* objection against a transparency-based account here. But I think this would be a mistake, for Finkelstein's point here is not that Henry's case is alienated *despite* following the ‘transparency’ structure outlined above, but rather that

it is *precisely this structure* that indicates that what we have here is alienation. This becomes clearer in Finkelstein's next example:

It's 7pm on a Monday. I'm at my office, working, when a friend phones to see if I want to go with her to a movie. I need only engage in a moment or two of practical reasoning—I think about the class that I have to teach in the morning, the letter of recommendation that's two days overdue, and the dissertation that I still haven't finished reading—whereupon I say to my friend, "I'm sorry; I'm sure I'd enjoy the movie, but I want to stay at the office this evening". (Finkelstein 2012: 109)

Of this example, Finkelstein says:

This is the kind of case that Moran's account should be best suited to explaining, one in which a subject rationally deliberates about what he wants to do, arrives at a single conclusion, and avows it. But now: at the point in the story when I tell my friend that I want to stay at the office, am I *assuming* that what I, in fact, want is in line with my assessment of what I ought to want? Surely not. At that moment, I *know* what I, in fact, want; my desire is (unlike Henry's fear in the last example) fully conscious, and I have no need of any such assumption. (Finkelstein 2012: 109)

In this case there is neither expressive nor rational alienation. Yet nor do we see room for the 'assumption' step that Finkelstein takes to be the point of transparency. Further, if there *were* such a step—if someone said: "well, it looks like I have most reason to want to stay here, so that's surely what I do want"—this itself would be evidence that the person is alienated from the desire in question (if indeed she really has it at all).

I will leave aside the question of whether this 'assumption' step is an accurate reading of Moran's position. In any case, I think Finkelstein must be right in thinking that a situation in which I come to know the contents of my mind by first working out what they ought to be and then assuming that they match my conclusion does indeed look like a case precisely of my alienation from the attitude(s) in question. So I shall have to make sure that my own claim about transparency as an aspect of the first-person perspective does not amount to such a picture. I will say something about

the various different aspects of this picture that I would like to reject, and then something about the positive account of transparency that I am defending.

4.2 INTERPRETING TRANSPARENCY

I have said already that my own concerns are not with a question about how we are able to have knowledge of our own minds, and this means that I need not posit transparency as any sort of mechanism for the purposes of answering such a question. Instead, I have suggested that we need some such notion in order to make sense of certain kinds of alienation: we must understand the nature of our first-personal relationship to our own mental lives to involve something like transparency, because where transparency is *not* present it marks a falling-away from the ordinary first-personal case. This, then, addresses the issue of what problem or question transparency is introduced to solve.

But Finkelstein's criticism also raises the issue of how, exactly, we should understand what 'transparency' itself amounts to; what it means to say that (as I shall want to put things later on), the first-person perspective is 'transparent'. For I have so far said somewhat evasively that I think that 'something like' Moranian transparency is needed in order to make sense of certain kinds of alienation. Let me now take some steps to address this evasion.

When I introduced the concept of transparency, I did so in the way that it is standardly done within the literature I am working with, which is to say via the following quotation from Gareth Evans:

[I]n making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward—upon the world. If someone asks me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question 'Will there be a third world war?'. (Evans 1982: 225)

In the first part of this remark, Evans speaks of where a person's 'eyes are directed'—whether figuratively or literally—when they speak to the matter of their own belief. This part of the quotation, however, has generally been overshadowed by the gloss Evans gives to this claim in the second part of the quotation, and in particular to the idea of two *questions*, one of which is then, in the broader literature that has built up around the work of Moran (and before him Shoemaker) said to be 'transparent to' the other.¹²

Now, this construction—that of a question's being transparent to another question—has meant various things to various authors. For Edgley, from whom Moran takes the terminology of 'transparency', it means that "I cannot distinguish the question 'Do I think that P?' from a question in which there is no essential reference to myself or my belief, namely 'Is it the case that P?'" (Edgley 1969: 90). Moran finds this way of putting things somewhat unsatisfactory, since he (sensibly) wishes to make room for the idea that a person can coherently conceive of their own fallibility—which is to say, can acknowledge that there is some possibility of a gap between the question of what they, a fallible human being, in fact believe, and what is actually the case. Moran generally prefers to say that I answer the first question *by* answering the other (e.g. (Moran 2003: 405)).¹³ And, as we have seen, Finkelstein interprets this suggestion as meaning that in order to know or find out the answer to the first question, what I do is to answer the second question (presumably on the assumption—which Finkelstein calls into question towards the end of his later paper—that the latter is more easily answered?). But in general, the idea that *what* is 'transparent' in the phenomenon at issue is fundamentally a *question* (with one subject-matter), and that what is 'seen through' this 'transparency'

¹² Evans does not use this terminology—as far as I am aware, he does not use the terminology of transparency at all in relation to the phenomenon he is discussing. But, just as my concern is not with Moran exegesis, much less is it with Evans exegesis, and so it does not really matter what Evans 'really meant' by his much-cited remark.

¹³ Compare also (Boyle 2011).

is *another question* (with a different subject matter) is widely accepted. I don't want to suggest that this way of thinking about transparency ought to be abandoned. But it brings with it certain potential pitfalls, as I'll explain below. Further, and perhaps more importantly, it also prevents from coming into view the full usefulness of the metaphor of a thing's being 'transparent'.

If we start from the idea that transparency is fundamentally about the substitution of one question for another question (whether or not we suppose that the substitution takes place because of the 'indistinguishability', from the first-person perspective, of the two questions), then naturally a central concern will be with giving an accurate characterization of the form of the second question. The first question is, in the original account given by Evans, of the form: 'What do I believe (about X)?', and as we've seen Moran follows Shoemaker in suggesting that we can generalize this to cover at least some attitudes other than belief, so that the form of this first question is now "what do I think / hope / intend / fear / desire / etc (concerning X)?". Given this, how should we characterize the second question—the question to which this first question, about the subject's state of mind, is 'transparent'?

According to Finkelstein's account of Moranian transparency, the second question takes the form: 'what *ought* I to think / hope / intend / fear / desire / etc (concerning X)?' And then, combining this idea with the idea that transparency is the mechanism through which I can come to know what is going on in my own mind, the challenge raised by Finkelstein that I outlined in the previous section becomes pressing. For a case in which I come to know what is going on in my own mind through such a method—by figuring out what I *ought* to be thinking (hoping, intending, fearing, desiring) about some matter and then assuming that my thought (hope, intention, fear, desire) matches this conclusion—does indeed seem to paint a portrait precisely of my being alienated from the thought (hope, intention, fear, desire) in question.

I have already relinquished the idea that ‘transparency’, however we understand it, will give us the answer to a question about how one has knowledge of one’s own mind. But the other component of Finkelstein’s Moranian picture—the idea that transparency, whether or not it tells us something about how self-knowledge is possible, is fundamentally about there being a ‘match’ between the answers to two different questions—brings with it problems of its own.

To see this, let’s return to Finkelstein’s original objection to Moran, given at the end of his book. There, recall, the problem was that we seem able to speak with authority about a wide variety of mental states that do not follow the structure of transparency. In section 3 of this chapter, I discussed the case of, as Finkelstein puts it, “attitudes that the reasons speak against”, and made use of this case to argue that we ought to understand ‘alienation’ as something that can come in multiple different forms, one of which is characterized by a failure of transparency. I now want to consider one more set of attitudes to which Finkelstein thinks that transparency cannot apply, namely “all the attitudes that would be neither prohibited nor required by deliberative reflection” (Finkelstein 2003: 163). One of Finkelstein’s examples here is his fondness for his friend’s dog, who we later learn (Finkelstein 2012) is called Sadie. Finkelstein says:

I don’t view it as rationally incumbent upon me to be fond of Sadie. I cannot answer the question of whether I am fond of Sadie by addressing a question about whether I ought rationally to be. Nonetheless, I’m fond of Sadie, and I have no trouble speaking with first-person authority about my fondness for her. (Finkelstein 2012: 105)

This case is an interesting and important one from my point of view. For it is not intended as a case in which our protagonist (here, Finkelstein himself) is speaking with first-person authority about an attitude which conflicts with what he takes to be the reasons (as perhaps it might if, for instance, the dog happened to be especially bad-natured). There is not, that is, the sort of alienation at work that I identified in the case of Max and his spider-phobia, or Moran’s psychoanalytic patient who is

plagued by an irrational sense that her mother's death was a betrayal. Indeed, this case does not seem to involve *any* kind of alienation. It is just an ordinary instance of a person's possessing an attitude, in what appears to be a fully first-personal sense.

If Finkelstein is correct in thinking that transparency is not on the scene here, then this is once again a serious problem for my central claim. For I am claiming that transparency is a key part of what it is to have an ordinary first-personal relationship to one's own mind; its absence marks a case of a certain kind of alienation. So I must show either that the case of Finkelstein's fondness for Sadie does, despite appearances, involve a form of alienation, or that we ought to understand transparency (and the case) in such a way that we are able to see how transparency *is* present here.

I think that there might be a version of the example in which we might plausibly suppose a form of alienation to be at work (I outline this version below). But on the most natural interpretation of the case, this seems like it would be a strange thing to say. I also think that Finkelstein is probably right to think that his fondness for his friend's dog is indeed neither rationally demanded nor rationally prohibited. And so my suggestion is that transparency *is* on the scene here, but that we should not understand this concept along the lines of the model that Finkelstein attributes to Moran.

The sense in which the first-person perspective is 'transparent', then, is not fundamentally a matter of a match between the answers to two questions, one of which concerns my state of mind, and the other of which asks what my state of mind *ought* to be. The problem that has been brought to our attention is that quite often, there doesn't seem to be any determinate answer to the latter question; plenty of things are, as Finkelstein puts it, neither rationally demanded nor rationally prohibited, so that there is nothing, strictly speaking, that one *ought* to think or feel regarding some

matter, and yet in such cases we are perfectly capable of thinking or feeling *something*, and of doing so in an entirely non-alienated manner.

This problem arises because of the form that we have specified for the second of the two questions between which the relationship of ‘transparency’ is supposed to hold: it arises, that is, where we take the second question to be “what *ought* I to think (...)?” This means that there may yet be versions of the ‘question-to-question’ interpretation of transparency that are more successful. There may also be versions of this interpretation that are compatible with the picture that I am developing here, and this is why I said above that I do not want to say that we must entirely abandon such ways of thinking about transparency. However, I think that the notion of transparency can be developed more effectively in other terms, leaving aside the formulation of one question’s being ‘transparent to’ another.

Consider again the case of Finkelstein’s fondness for Sadie the dog. I have agreed that such a fondness need be neither rationally obligatory nor rationally prohibited. But the realm of rationality need not be exhausted by the question of whether a thing is obligatory or prohibited. I can ask Finkelstein *why* he’s so fond of Sadie, and it needn’t be the case that he just meets this request with a blank look. He may not have a fully worked-out answer to my question, but he can begin to make an attempt to answer. He might say, for instance: “oh, I don’t know. She’s kind of stupid, but she’s always so cheerful and upbeat...”, and such a gesture acknowledges, at the very least, the relevance of the question. I would like to say that here Finkelstein’s *attitude* is itself ‘transparent to’ the part of the world that it is about, namely Sadie. It is not that the attitude merely matches the answer to some other question (about, say, what might be rationally warranted here, even if it is not rationally

obligatory). Rather, the attitude *just is* Finkelstein's finding (part of) the world to be a certain way—finding Sadie to be goofy and adorable.¹⁴

Now suppose that when I ask Finkelstein why he is fond of Sadie, he is not merely inarticulate about his reasons, but instead insists that there just aren't any. I ask about Sadie's good qualities, and he has nothing to say. It's not that she necessarily has *bad* qualities; she's not an unfriendly or messy or dislikable dog. But the fondness does not seem to be experienced by Finkelstein as his *finding Sadie* to be any particular way at all; although the attitude is *about* Sadie, it doesn't really seem to be 'transparent to' Sadie. The attitude is experienced by Finkelstein primarily as a psychological fact about himself, rather than a fact about how Sadie is and the attitudes that her nature might rationally elicit. It exists within Finkelstein's psyche as a sort of opaque object, something that he can think and talk about, but not something through which Sadie herself is visible. The fondness simply sits there, lump-like, a part of himself that he is just inexplicably stuck with. And where this is the case, what we have is once again a case that it makes sense to speak of as a kind of alienation.¹⁵

What I have tried to do here is to draw out the idea of transparency as a useful metaphor, rather than simply a technical term. The first-person perspective is 'transparent' in the sense that a person's attitudes are experienced not as a set of psychic objects, but rather more like windows on to

¹⁴ Doesn't this imply that Finkelstein must suppose that *everyone* ought similarly to find Sadie goofy and adorable, rather than, say, hyperactive and annoying? And doesn't that sound too strong an implication? Surely Finkelstein might well allow that though *he* is fond of Sadie, and finds her goofy and adorable, he can well understand why others aren't and don't. Surely, that is, his fondness may well be idiosyncratic, without this idiosyncrasy amounting to any kind of alienation. Yes! Stay tuned for Chapter 5.

¹⁵ What does the limit case of such alienation look like? —Perhaps something that you can only grasp as an inexplicable behavioral disposition of yours (say, to always pet Sadie—though no other dog—when she's in your vicinity, even though you have absolutely no idea why). And in this limit case the attribution of 'fondness' becomes rather strained. We might usefully consider the case of Warren Quinn's 'radio man' here (Quinn 1993).

the world. And this aspect of mindedness can be brought into view, I have suggested, by thinking about the contrast between this ordinary case and the case in which the attitude does *not* function like a (transparent) window, but is rather more like an (opaque) object in its own right.

4.3 A CASE STUDY: ALIENATED DESIRES

So far, I've been making use of examples of alienation drawn from within the conversation between Finkelstein and Moran. I now want to demonstrate the broader applicability of the ideas I've been developing by showing how they relate to a new case, which is provided by Alec Hinshelwood:¹⁶

We can imagine a man who has received a strict religious upbringing and who later goes on to live in a monastery. Say that this person enters into a particularly severe order which encourages self-flagellation in response to thoughts or feelings of a sexual nature, but later becomes disillusioned and leaves the monastery. Now this person no longer believes in God, and neither does he maintain his previous evaluative outlook which condemned his erotic desires. Nevertheless, we can imagine that he still wants to hurt himself upon entertaining an erotic desire, even though he no longer considers there to be any reason for him to condemn them. [...] In this case, the lapsed believer may experience his desire to hurt himself as an “alien intruder”, or a force which moves him to act “in spite of himself”, as we might say. (Hinshelwood 2013: 244)

Hinshelwood's ex-monk suffers from what I have called ‘rational alienation’. His desires to harm himself are quite out-of-joint with what he takes the reality of his situation to warrant; the

¹⁶ Hinshelwood introduces this case to illustrate the phenomenon that he calls ‘alienation’, which he contrasts with ‘identification’. His concern in the paper is to reveal the connections between certain accounts of the distinction between alienation and identification, on the one hand, and certain accounts of the nature of action, on the other. He argues (to my mind convincingly) that an event-causal picture of agency requires that we accept a hierarchical model of the distinction between alienation and identification (such as that offered by Harry Frankfurt). However, the hierarchical account brings with it problems which Hinshelwood argues are insurmountable; this means, thinks Hinshelwood, that we shall have to find an alternative model of action to that of the event-causal model, and he finds the basis of such in Moran's account. Hinshelwood, like me, finds in Moran's account of self-knowledge something philosophically important that is not really about self-knowledge at all. For Hinshelwood, the key point at the heart of the Moranian approach—as opposed to both the event-causal model of agency, and the Frankfurtian hierarchical model of identification—is that it takes the agent and her rational powers as basic, rather than attempting to construct such out of certain arrangements of causal happenings or combinations of psychological states.

experiencing of erotic feelings is not something that calls for self-flagellation, and this is something that he can see quite clearly. Nonetheless, at the relevant moments, the man's desires arise, and when they do, far from being a transparent window *onto* the world, they instead function like an opaque object that must be navigated around. As Hinshelwood puts it:

The man is [...] forced to deliberate about the avoidance or satisfaction of his alienated desires in a way which renders them similar to facts about his height or even his external surroundings. His alienated desires constrain the projects he might sensibly undertake, or figure as obstacles which must be strategically overcome. [...] The attitudes from which the man is alienated figure in his deliberations in much the same way as straightforwardly external matters. (p. 245)

Now of this example, we may wonder whether or not the case also involves the other features of first-personal self-relation that we have been considering. For instance: can this man speak of his desires without looking towards evidence, and with distinctive authority? My guess is that, as Hinshelwood is imagining things, he can. If this man is our friend and we know about his difficulty with these desires, we may in certain circumstances ask him how he's faring: "are you struggling right now? We can leave the party if you like", to which he may say "no, everything fine right now, we can stay", or alternatively "ugh, yes, I'm starting to want to hurt myself, let's go somewhere else to take my mind off it", and in such a situation we would be unlikely to doubt that he is the person best placed to make such assessments about his current state of mind.

Although this may be the most natural reading of Hinshelwood's case, I think that strictly speaking, the case is under-described in this respect. For we can also imagine another version, in which the man's desires are not quite so first-personally accessible to him. Having reflected upon his own behavior and history, he may well know that he does indeed have these desires, and yet he might fail to be able to speak about them with the authority that (for instance) Max has regarding his fear of spiders. Perhaps, for instance, the lapsed believer notices that in certain circumstances he

may find himself inflicting pain on himself without having been previously aware of a desire or intention to do so; perhaps sometimes the presence of the desire is only visible to him in retrospect, or not at all.

In the first version of the story, the man is able to speak about his current state of mind authoritatively and without having to look towards evidence to support his assertion. In this second version, however, if you ask your friend how he is doing, he may look to his own current behaviors or experiences in order to infer whether or not his troublesome desires are coming on to the scene. For instance, rather than simply making an assertion about how his desires stand, he may say: “well, I feel mostly okay, but a bit jittery, and I’m fidgeting a lot, which is usually a bad sign; I guess we should probably leave to be on the safe side.” Of course, it’s likely that he knows himself well and may well be correct. In this case, however, there is much more room for the possibility that someone else may be able to make a more accurate inference than he. A friend who knows him very well, for example, may be able to say: “actually, I don’t think you need to worry right now—your fidgeting is just foot tapping, and isn’t that usually just a sign you’ve had too much caffeine?”, and the friend may well be right.

The possibility of the two different readings reinforces the distinction I drew between two kinds of alienation in section 3.2, and in particular supports my central claim that transparency is needed in order to understand one particular form of alienation. For *whether or not* the man suffers from forms of alienation that would impact his first-personal authority, still it is clear that his situation is one of alienation; it is not the ordinary way in which one experiences one’s desires. And this form of alienation is characterized precisely by the absence of transparency.

4.4 THE POSSIBILITIES OF VISION / LOVING THE ALIEN

I have developed the idea of transparency in terms of the metaphor of a window onto the world. Of course, no metaphor is perfect, and it would not be too difficult to put pressure on this one and cause cracks to appear. However, my hope is that in emphasizing the imagery of transparency—of the transparent as something *seen through*, as opposed to the opaque, which is something *seen*, itself an object of scrutiny—we can avoid some of the problematic consequences of the question-to-question interpretation.¹⁷ My account allows space, for instance, for us to understand transparency as present in cases for which the realm of the rational is not exhausted by simple obligation or prohibition, such as the case of Finkelstein’s fondness for Sadie the dog. In addition, thinking of transparency’s medium in terms of the metaphor of vision (the world seen) rather than language (the question answered) allows us to see it at work even where the person struggles to articulate the rationality of what is going on with her attitude. When I ask Finkelstein why he’s so fond of Sadie, he may, as I suggested, attempt to offer some sort of explanation. But he may find himself quite unable to do, and then he can throw up his hands and say: “oh, just *look* at her!” Here, the inability to find the words is not an indication that Finkelstein’s attitude is experienced as opaque. Indeed, quite the contrary: it is what he sees when he sees Sadie, and he is telling you that the answer to your question is to be found out in the world, and is something that can be captured by the subtlety and complexity of vision even if he cannot find, or does not possess, the words.¹⁸

¹⁷ It also seems to me that thinking of transparency in these terms allows the idea to resonate more closely with one’s everyday experience as a subject of attitudes, since the imagery of the transparent window seems more familiar and intuitive to me as a description of what it’s like to have an (unalienated) attitude. By contrast, the account given in terms of the answering of one question via the answering of another seems a level too abstracted to be really familiar—it seems primarily like a philosopher’s construct. But I accept that this might be an idiosyncratic response on my part.

¹⁸ “The moral agent, as well as the artist, may find himself unable to describe something which in some sense he apprehends” (Murdoch 1999: 90).

Finally, consider the way in which this shift to a metaphor of vision makes room for the important idea that in many cases of rational alienation, it may well not be what we might think of as ‘the calm voice of reason’ that is in the right. Some attitudes may be recalcitrant in the face of how we find the world to be because they express some form of pathology; such seems to be the case for Hinshelwood’s ex-monk. But other attitudes may be recalcitrant in the face of how we find the world to be precisely *because* there is something that the attitude is seeing that ‘I’ (the ‘I’ that considers the attitude ‘alien’ and not-I) am not seeing, or am refusing to see. (Here perhaps we may think of the attitude’s ‘opacity’ in terms of curtains resolutely drawn against a view that we are not willing to look at.)

Where an attitude is at odds with one’s sense of what the situation calls for, one hopes to be able to bring the two back to resolution. But this may be just as frequently a matter of coming to see that one’s explicit sense of the reasons was skewed, and thus an adjustment taking place on *this* side of the equation, as it is of having the ‘alienated’ attitude fall into line with what one takes, in the calm moment of reflection, to be what is called for. Often those attitudes that seem confusing, baffling, and alien may turn out to be in touch with reality in a way that our explicit articulation of how things stand may fail to be. Someone raised with homophobic beliefs may experience her desires to behave in certain ways as alienated, as in conflict with what she takes to be genuinely warranted or demanded here. In this case, I’m inclined to say, her desires are in touch with certain important aspects of reality—with the possibility of fulfilling relationships and experiences with people of her own gender—and it is her explicit sense of the reasons that is inaccurate.¹⁹

¹⁹ Notice that there is a deep connection between, first, one’s ability to articulate what one sees, second, how one experiences ‘the calm voice of reason’, and third, the experience of alienation. One may well mistake lack of ability to ‘make a case’ for the desire as its being opaque, and all sorts of things may limit one’s ability to make such a case. A person’s social setting limits and delineates the

In this way, I aim in my account of transparency to resist the image of the agent who, in the calm voice of reason determines how things stand in the world (alternatively: who answers the question ‘what ought I to think/feel/etc here?’) and then checks (or, worse, assumes) that her attitude matches the result of her reasoning. Another way of putting this is to say that I am resisting, again, (see Chapter 2) the image of the agent who ‘steps back’ from her attitudes and surveys them from some privileged distance.²⁰ Nomy Arpaly states that:

Much has been written about agents who experience themselves as choosing between their desires autonomously while being paradigms of self-control, and about agents who feel the pull of temptation and surrender to it. Almost nothing has been written about the agent who steps away from her desires and, as it seems to her, chooses calmly between them, feeling apparent mastery over temptation and emotion, while the very “I” that steps away from the desires is the unconscious dupe of other desires, emotions, or irrationality. (Arpaly 2003: 20)

I share at least some of Arpaly’s concern here.²¹ And I have tried to provide an understanding of transparency that allows for, and gives us some tools for talking about, such agents.

5. RE-UNIFYING THE FIRST-PERSON PERSPECTIVE

So far, I have attempted to develop the notion of transparency that I shall be putting to work in the following chapter, and defended my understanding of it against various challenges raised by Finkelstein against Moran’s transparency-based account of first-personal self-knowledge. I have tried to show that the notion of transparency, properly understood, is necessary if we are to be able to

concepts that are available to her, meaning that she may not be adequately equipped to articulate what she currently is aware of in a confused and confusing way. There are fascinating possibilities to explore here concerning the construction of this ‘calm voice’ through social, historical, and psychological forces, and the possibilities for expanding one’s potential for self-understanding and understanding the world through a critique of such forces.

²⁰ See (Russell 2017) for a reading of Moran that allies him closely—at least more closely than I would like—with Korsgaard’s model of the first-person perspective.

²¹ Arpaly offers a related example to mine of the person raised with homophobic beliefs at (Arpaly 2003: 14–15).

mark the distinction between, on the one hand, the fully first-personal case of the possession of a particular attitude, and on the other a certain form of alienation, and I have tried to develop the concept in ways that seem to me plausible and fruitful. In this section, I shall do some work to try to re-unify some of the various aspects of the first-person perspective that I have, for the purposes of my argument here, separated out from one another.

In section 4, I argued for a distinction between two kinds of alienation, which I provisionally labeled ‘expressive’ and ‘rational’ alienation. My discussions so far have pointed towards the conclusion that Finkelstein’s expressivist account of first-person authority can best account for the former kind of alienation, while transparency, understood along the lines of my own Moran-inspired account, is needed if we are to be able to make sense of the latter. If I am right about this, then we already have enough to establish that transparency is a key component of the first-person perspective. But I would like to show, further, that the picture of the first-person perspective that we have ended up with is not merely an unsatisfying hybrid of two views patched together in an ad hoc manner, but rather a fundamentally unified picture. There *are* two (at least two) different forms of alienation, each corresponding to a different facet of the first-person perspective, but these two forms of alienation and their corresponding forms of first-personalness are nevertheless intimately linked.

To see this, consider first an interesting ambiguity that is at play in a way of speaking that is frequently used (by Finkelstein, Moran, and others) in the characterization of transparency. Where transparency is at work, we are told, I can answer a question about what I think or how I feel about X simply by ‘turning my attention’ to X, ‘thinking about’ X, ‘reflecting on’ X, ‘looking to’ X, and so on. The ambiguity I see here is this: ‘turning my attention’ to X (and the other similar constructions) can be used, in the context of Moran’s account, to mean something like ‘turning my attention to X

in order to determine what is rationally to-be-thought or to-be-felt concerning the topic of X', but it could also mean simply 'calling X to mind' (or even just turning my eyes towards the object, if it is in the vicinity). The former of these signals the full-blown rational Moranian picture, and the failure of 'transparency' in this sense would be rational alienation. But the latter version, it seems to me, might point towards something more like Finkelstein's picture. My ability to speak with first-personal authority about (say) my irrational fear of spiders is marked, among other things, by my ability to access that fear just by thinking about—or indeed seeing—spiders themselves. I don't need to think about my own past behaviors (my tendency to jump out of bed if I see one on the pillow beside me), or search the inside of my mind for the box labeled 'arachnophobia'. To be able to truly say of myself 'I'm scared of spiders!', and in particular to say it as an expression of that very fear, I need not think about *myself* at all. All I need to do is 'turn my attention' to *spiders*, perhaps to the one in front of me, or just have my imagination linger a while on images of their hairy legs and scurrying movements. Where my attitude is expressively alienated, merely turning my attention to the object of my fear won't be sufficient to elicit the articulated expressive response. In the case of the patient with an unconscious feeling of betrayal in relation to her mother, turning her attention to her mother doesn't elicit the expression 'I feel betrayed!'. So this latter interpretation of knowing one's own attitude by 'turning one's attention' to its object looks more closely tied to sort of expressive authority/alienation that Finkelstein's account covers.

I've said what we have here is an ambiguity, and I think this is right. (I also think this ambiguity is partially responsible for various points at which Moran conflates together various different forms of alienation, as in the case of the analysis discussed above.) But I also think it makes sense to think of these two different ways that we may 'turn our attention' to something as related to one another, perhaps even as simply standing at different points on a spectrum of

fundamentally similar activities. At one end of the scale we have the explicit activity of deliberation, in which the question of what to do or to feel is raised consciously, and relevant considerations are taken into account (and perhaps—though see my suspicions raised in the previous chapter—somehow ‘weighed’), while at the other end we have something much more basic and inarticulate but not entirely different in kind. We begin to move from the more basic end of the spectrum towards the more explicitly rational end as we begin to pick out salient (and proto-reason-giving) features of object we have in view—as when I suggested that the arachnophobe might think about spiders’ hairy legs and scurrying movements in order to bring to immediate consciousness the fear that characterizes her phobia.

With this in mind, let us think again about the phenomenon of expressing a mental state by self-ascribing it that Finkelstein argues is key to understanding first-person authority. This phenomenon is itself, I’d like to suggest, intimately connected to the idea of transparency as I’ve been developing it. For expression *itself* seems to admit of the possibility of degrees of defectiveness that can best be understood in terms of transparency. There is a difference, for instance, between the ordinary case of expression by self-ascription (“I’m so happy!” I say, beaming, when the gift you’ve given me is *just* what I wanted) and a case in which the expression itself is experienced as a sort of possession by an alien force (“Take it away, I’m terrified!” I find myself yelping, much to my embarrassment, and even though I *know* it’s just my niece’s beloved and harmless pet mouse, and also that she’ll now never see me as the ‘cool aunt’ ever again). My discussion in the preceding two paragraphs suggests that we might wish to interpret this difference as one more of degree than of kind, but it *is* a difference. And further, it is a difference that not merely sits alongside the fact of expression-through-self-ascription that is present in the two cases, but instead alters the nature of the phenomenon of expression itself. In the first case, we may say that the person expresses *herself*,

but in the second case there is reason to resist this construction. I may prefer to say, for instance, that while *the fear* was present in the words, in a certain sense *I* myself was not. The expression, the utterance itself, is experienced as something that happened *to* me, rather than something I *did*. In both cases, the attitude is conscious, and we seem able to accord the person's words the special authority that is distinctive of the first person perspective,²² and it may well be that this is best explained, as Finkelstein suggests, in terms of expression. But if so, the nature of that 'expression' is different in the two cases, and the difference, it seems to me, can be explained in terms of the degree of transparency of the attitude expressed—the extent to which, that is, the attitude expressed just *is* how the person finds things in the relevant part of the world, as opposed to a bit of psychic baggage with which they find themselves in possession—even if that baggage is something that can be expressed through self-ascription.

6. CONCLUSION

In this chapter, I have developed an account of the 'transparency' of the first person perspective. The initial inspiration for my account came from Richard Moran's work on transparency in the context of his account of first-personal self-knowledge, but I have attempted to extract the notion from this context and to re-work it for my own purposes. The cornerstone of my argument is my claim that we need the concept of transparency in order to make sense of a certain form of alienation—a certain way, that is, that one may fail to relate to one's attitude in a fully first-personal

²² Although notice that in more extreme versions of the alienated case, we—and even the person herself—may wish to withdraw the ordinary assumption of authority. If I *really* just *find myself* saying "I'm afraid!" for no reason that is at all discernable to me—even the fact of my utterance is bafflingly to me—then even if my words are accompanied by the usual physical markers of fear (my heart is beating quickly; I'm shaky) I may well have reason to wonder whether I really am straightforwardly *afraid*, or whether I am not undergoing some other strange sort of psycho-physical upheaval that is not adequately described in terms of fear at all. (Again, compare Quinn (1993).)

manner. I defended this claim by separating out two different ways in which one may be said to be ‘alienated’ from an attitude, and using this distinction to defend the importance of transparency against the objection (which I developed out of some criticisms raised by Finkelstein against Moran) that it seems possible for first-personal authority to be restored to an agent without this restoration being accompanied by a restoration of transparency.

I also considered Finkelstein’s claim that transparency itself describes a state of alienation, and attempted to offer a re-interpretation of the notion that is not vulnerable to this complaint. My re-interpretation has two parts: first, I relinquished the idea that transparency allows us to understand how we have first-personal self-knowledge. Second, I showed how we can move away from the dominant interpretation of transparency which understands it primarily as a relation between two questions, choosing to emphasize instead the possibilities made available to us by taking more seriously the metaphor and imagery of a transparent window on to the world.

Finally, I attempted to do some reparative work, showing how the forms of alienation (and thus the associated facets of the first-person perspective) that I had previously separated out from one another should nonetheless be understood in terms of one fundamentally unified subject. Transparency and expression—the key concepts for Moran’s and Finkelstein’s accounts respectively—can be usefully understood as providing the basis for complementary understandings of the notion of the first-person perspective.

What, then, have we gained from all of this? In the following chapter, I will try to develop some of the metaethical consequences that I believe follow from the account that I have given here. For now, let me say the following. First, notice that the concept of transparency as I have developed it is not given adequate expression in any of the positions taken on ‘the first-person perspective’ that I canvassed at the end of the previous chapter. In particular, I noted that Enoch, Nagel, Williams,

Korsgaard, and Street, can be divided up along several axes, one of which was their answer to the question of whether ‘the first-person perspective’ is to be characterized in terms of our ground-level attitudes (whatever we actually happen to think, feel, intend, desire, and so on) or in terms of our ability to ‘step back’ from these attitudes. But as should now be reasonably clear (I’ll say a little more about this in the following chapter), my own account takes a third approach. The ‘first-person perspective’, or at least the facet of it that has been the focus of my interest here, is characterized in terms of a relationship of transparency between a person’s attitude and the world that she encounters. What is distinctive about a first-personal relationship to a psychological attitude is neither that it is a thing with which we happen to find ourselves in possession, nor that it is a thing from which we can ‘step back’; rather, it is our finding the world to be a certain way, and in particular finding it to call for a certain sort of response on our part.

When we are thinking of our first-personal relationship to attitudes in this way, certain aspects of the anti-realist’s model of the relationship between attitudes and value already begin to look uncomfortable. From the first-person perspective (which is just to say: from the perspective we are, in fact, occupying), the experience of our own psychological attitudes is an experience of already moving through a world which we perceive to call for and to warrant particular kinds of response. And we came to see this by thinking about an odd or defective sort of case—the case in which one’s attitude is *not* transparent to the world in this way, but exists merely as a psychological fact. In this light, the anti-realist’s commitment to the priority of psychological attitudes over a value-laden world that might call for or warrant such attitudes begins to look, itself, like it relies on a fundamentally alienated model of those attitudes.

CHAPTER 4: THE TRANSPARENT FIRST PERSON AND THE VALUE-LADEN WORLD

Attention alone—that attention which is so full that the ‘I’ disappears—is required of me. [...] The poet produces the beautiful by fixing his attention on something real. It is the same with the act of love. To know that this man who is hungry and thirsty really exists as much as I do—that is enough, the rest follows of itself.

Simone Weil¹

1. INTRODUCTION

I have argued that the first-person perspective is, in a sense elaborated in the previous chapter, ‘transparent’ to the world. The most important consequence of this account for our purposes is what it tells us about the way that a person relates to her own psychological attitudes, and how she takes those attitudes to relate to the world that lies beyond and prior to those attitudes. One thing we have seen is that a person’s attitudes are not, for her, something that can be treated merely as given facts. For *this* sort of relationship to one’s own attitudes characterizes not the ordinary form of possessing an attitude, but rather a case of alienation. What this means is that any anti-realist attempt to explain how the realm of value follows upon such facts, themselves treated as given, will not be acceptable from the point of view of the person herself—describing as it does only the defective case, the case of alienation. For her, the beginning of the story cannot be the attitudes themselves, since these attitudes themselves, in virtue of their transparency, invoke the concept of a world which warrants or calls for such a response. From the point of view of the person, the value-laden world—

¹ (Weil 2002: 118–19).

a world filled with things that are desirable rather than merely desired, frightening rather than merely feared, and worthy of doing rather than merely intended—comes first.

Another way of putting this idea is to say that the anti-realist approach, in attempting to understand human attitudes as something prior to evaluative reality, works with a picture of a person's attitudes which must look, from the point of view of the person herself, like a picture of alienation. One understands one's own attitude as a (mere) psychological fact, as something quite separate from questions about how things stand in the world (and thus whether such an attitude is true, or appropriate, or warranted) only where that attitude is a part of one's psyche from which one is alienated. The anti-realist model of the relationship between a person's attitudes and evaluative reality appears to make sense only if one is looking at it 'from the outside', as it were. When the person in question is you yourself, the anti-realist model portrays not the ordinary situation of wanting or intending or fearing, but rather the uncanny versions of these attitudes in the form of their alienation.

The above gives a rough sketch of what I take to be the metaethical significance of the concept of transparency as I have developed it. In the rest of this chapter, I shall develop this basic idea in more detail, first in relation to the question I raised in Chapter 2—that of what we're up to when we're deciding what to do—and second by considering how my accounts relate to broader questions about the relationship, in the context of practical thought in particular, between the mind and the world.

2. TACKLING TASKS FROM CHAPTER 2

In Chapter 2 I introduced the question of what we're up to when we're trying to decide what to do. I noted (following Nagel and Enoch) that trying to decide what to do involves the idea that certain

answers to the question of what to do must count as ‘better’ than others; in deliberating, that is, there is something that one is trying, in some sense, to get right. However, I disagreed with Enoch’s suggestion that this observation alone is sufficient to show that deliberation involves a commitment to realism. I argued on behalf of the anti-realist that there are plenty of different ways that one might account for this ‘getting something right’ feature of deliberation in anti-realist terms. The most straightforward of these is what I called the ‘simple neo-Humean’ proposal, according to which certain ways of resolving one’s deliberation count as ‘better’ than others, or as ‘getting something right’, because they are likely to (while other courses of action are not, or not so much, likely to) result in the satisfaction of (certain of) the person’s desires. I then suggested that the basic structure of this proposal could be extended beyond the neo-Humean case to allow the anti-realist a variety of different approaches to the problem. The ‘anti-realist schema’ that I introduced for this purpose states that what makes some particular resolution of my deliberation count as ‘getting it right’ is not any inherent goodness in some particular course of action or its outcome, but rather something in me: some attitude—some normative judgement or evaluative commitment or plan, perhaps—or set of attitudes, or the logical consequences of such.

My point in constructing this anti-realist response to Enoch’s argument for realism was not merely to accord with principles of philosophical good sportsmanship (giving one’s opponents the best possible shot before one attempts to demolish their position, say). Rather, my thought was that there really is something *to* the anti-realist schema, and that whatever it is that is right about this way of thinking about deliberation is something that the realist ought to be able to accommodate too. And so I concluded, in section 5 of Chapter 2, that I face two tasks: I need to show both what is wrong with the anti-realist schema as an account of deliberation, and also to show how the realist can accommodate whatever will turn out to be right about it.

One point that I think the anti-realist gets right—and which is at its most explicit in the simple neo-Humean iteration of the view—is that such things as a person’s desires *do* matter in deliberation, and furthermore they matter *as desires*, rather than (as in Enoch’s, and perhaps also Korsgaard’s, account) as mere motivational matter than is made normatively relevant only by having endorsement bestowed upon them by a further attitude in the form of a normative judgment. Enoch had attempted to see off the anti-realist by insisting that a desire would need to be *accompanied* by an extra thing (a normative judgment) in order to play its role within deliberation.² On my account, the problem that the anti-realist faces is instead that a desire is able to play the role that it does within deliberation only because and in so far as that desire is itself transparent, and thus invokes already the idea of a world beyond the desire itself—a world that calls for or warrants the desiderative response. The same is true of a normative judgment. And this understanding of desire (and of normative judgments, and indeed of any of the attitudes that the anti-realist might wish to have serving at the basis of deliberation) is not compatible with the anti-realist’s understanding.

The broad argument against the anti-realist schema for deliberation, then, is as follows. The anti-realist had hoped to explain the ‘getting something right’ feature of deliberation without calling on anything in the way of value out in the world. She attempted to do so by suggesting that what fixes the ‘rightness’ of certain outcomes of one’s deliberation is some attitude, or set of attitudes, that the person possesses. This basic idea was open to a wide variety of cashings-out, depending on the form of anti-realism that one prefers, but the foundational thought in each was that we might be able to begin from the fact of some such attitudes, grasped as something *prior* to any questions about

² And as I argued in Chapter 2, it’s very unclear that insisting on such a judgment actually helps the realist at all, given that anti-realists such as Korsgaard and Street are happy to incorporate the making of such judgments into their own accounts.

what the world might call for, and construct the ‘getting something right’ feature of deliberation on this basis.

But what we have learned is that attitudes, from the point of view of their possessor (who is, here, the person who is trying to decide what to do), are not amenable to such an interpretation. In the ordinary case, one’s own desire or intention or normative judgment (substitute your preferred anti-realist foundational attitude here) is not something that can be grasped, by you yourself, as something prior to questions about what the world calls for. In the ordinary case, your possessing the attitude you do just is, already, your finding the world to be calling for something (for desire, for an intention, for a normative judgment—or more directly for the actions that would follow upon such attitudes). In this way, the transparency of the first-person perspective means that the attitudes that the anti-realist had hoped to make use of already invoke the idea of a value-laden world, and it is *this*—not the bare psychological fact of the attitude—that the deliberator must take to render certain courses of action ‘better’ than others.

Indeed, where the deliberator relates to some attitude of hers as a ‘bare psychological fact’, as something that is not her already finding the world to be a certain way, what this signals is a form of alienation. And, I claim (I defend this claim in more detail in section 3), an attitude experienced as alienated in this way cannot serve as a basis for deliberation in the way that the anti-realist had hoped. What is clear even prior to a full defense of this claim is that the illustrations of the anti-realist schema that had made it seem initially so attractive and plausible were *not* based on alienated attitudes, but rather on the ordinary non-defective cases of desiring, intending, valuing, and so on.

I’ll now go on to show how this argument looks in detail when applied to one particular iteration of the anti-realist schema for deliberation, which I construct out of Mark Schroeder’s neo-Humeanism.

3. DESIRE AND DELIBERATION

3.1 DECIDING WHETHER TO GO TO THE PARTY

Let's return to Schroeder's central example, that of Ronnie (who loves dancing) and Bradley (who hates it) and the party (which will involve dancing) to which they are both invited. Schroeder claims, quite plausibly, that the fact that there will be dancing at the party is, on the one hand, a reason for Ronnie to attend, and, on the other, *not* a reason for Bradley to attend (indeed it may be a reason for him to stay away, if, for instance, people are likely to try to pressure him into dancing). And, says Schroeder, "it's not hard to see *why* Ronnie's and Bradley's reasons differ, at least at a first pass—this is something to do with their respective psychologies. It is because of what they *like*, *care* about, or *want*" (Schroeder 2007: 1). Schroeder continues:

It's largely uncontroversial—even among philosophers—that at least *some* reasons are like Ronnie's and Bradley's, in that whether they are reasons for some particular person depends on some features of that person's psychology, such as what that person desires, wants, likes or cares about. [...] And according to one time-honored tradition often taken to be vaguely inspired by Hume, *all* reasons are at some basic level pretty much like Ronnie's and Bradley's in this way. They are explained in some way by the psychological features of the agents for whom they are reasons—by the psychological features that explain the difference between Ronnie and Bradley. (Schroeder 2007: 1)

As I noted in Chapter 2, there is something crucial at work in this passage that Schroeder does not make explicit. Schroeder begins from the claim that the *difference* between Ronnie's and Bradley's reasons is to be explained by pointing to a psychological difference between them. There is some sense in which this claim really must be uncontroversial, provided we agree that Ronnie and Bradley's reasons do differ. For all that we've been told about them is, precisely, a difference in their respective psychologies; we literally haven't been given any other information about these people and their circumstances that *could* explain the difference. But by the end of this passage, as he outlines the neo-Humean's idea that "*all* reasons are [...] pretty much like Ronnie's and Bradley's",

Schroeder is no longer talking about how we are to explain the fact of a *difference* between Ronnie's and Bradley's reasons, but rather about the question of what explains the fact that Ronnie and Bradley *have* the reasons that they do at all. He is has now moved, that is, to the question of in what Ronnie's and Bradley's reasons consist; the question of what reasons *are*.³ Importantly, then, what Schroeder is outlining at this part of the text is really the claim that reasons depend on people's desires *and not* on any further normative features of the world.

Thus, when Schroeder tells us that it is philosophically uncontroversial that "some reasons are like Ronnie's and Bradley's", what he means (or ought to mean, given the structure of the rest of the argument here, and in the book as a whole) is that it is philosophically uncontroversial that some reasons depend on the psychological facts regarding what a person wants, likes, or cares about, *and not* on any further normative facts about (for instance) what is valuable, good, or worthy of wanting or caring about. It is much less clear that *this* claim is uncontroversial. In any case, I think it is wrong—it is wrong, that is, to think about even the case of Ronnie and Bradley in these terms. And I think that my account of transparency can help us to see why it is wrong.

Do you like dancing? Then put yourself in Ronnie's shoes. There's a party tonight to which you are invited and at which there will be plenty of dancing. Shall you go? Maybe you've a paper to finish, but you'll still have time to do that later in the week, so there's no real urgent need to work tonight. And, for once, your arch-enemy (who often attends such parties) is out of town, meaning that a factor that would usually count against attending doesn't apply in this case. So you can reason: "I have a strong desire to dance, and the party will be a good opportunity to do so. My paper can

³ And, of course, for Schroeder, 'reason' is the most basic normative notion, so that all talk of goodness, value, oughts, and so on is interpreted in terms of reasons.

wait—and this time Moriarty won't be there, so there's nothing much that counts against my going. I'll go!"

Do you hate dancing, and being around people encouraging you to dance? Then put yourself in Bradley's shoes. Shall you go to the party? Ronnie is likely to be there (especially since Moriarty won't be), and it would be nice to see him. But at the same time, he's sure to try to pressure you into dancing, which will be very tiresome and end up making you feel embarrassed. So perhaps you reason: "I'd like to see Ronnie. But I have a very strong desire not to feel deeply embarrassed on the dance floor, and deep embarrassment on the dance floor is sure to be the outcome if I do go. I'll stay home, and arrange to meet Ronnie for coffee later in the week."

In each situation, I had the deliberator take into account certain psychological features that they possess (the desire to dance; the desire to avoid peer pressure) that are, as Schroeder suggests, relevant to their attempts to decide what to do in the face of their respective party invitations.⁴ But now my account from the previous chapter allows us to elaborate on the first-personal relationship that each deliberator bears to these psychological features. What we have learned is that your desire to dance is not merely some psychic baggage you find yourself in possession of, but is rather 'transparent' to the world. It is your finding *dancing* itself to be something desirable. I can ask you about what you see in dancing, and though you may be more or less eloquent on the matter, the question is one that makes sense, and one that you can make some attempts to answer: "it's fun, and liberating, and allows me to let off steam and to express myself in ways I can't in the rest of my day to day life". In this way, your love of dancing is your response to something about how you find *dancing* to be—namely, something worth doing.

⁴ There is a complication in Schroeder's view that means that this is not, actually, how he would have the deliberators proceed. I deal with this complication in the following section.

Thus the story of your deliberations, although it does involve reference to your desires, does not ultimately terminate with the mere psychological fact of the desire. From your point of view, your desire is itself a response to something prior, something out in the world that means that dancing is something that it makes sense for you to spend time doing. And *this*—not the mere psychological fact of your desire—is why it makes sense for you to conclude your deliberations by deciding to go to the party.

3.2 THE ALIENATED CASE

So far, I've argued on the assumption that your desire to dance is transparent. But my account of the transparency of the first-person perspective did not involve the claim that *all* attitudes are actually transparent. Indeed, quite the contrary: my argument depended on the thought that we need to be able to make a distinction between the ordinary first-personal case and a case in which one experiences the attitude as in some way 'other' or 'alien'. So I am acknowledging, of course, that one can have attitudes that are not transparent (in the previous chapter, I spoke of such attitudes as 'rationally alienated' or 'opaque'). Since this is so, we ought to consider the question whether or not one can deliberate on the basis of an alienated attitude.

In theory, of course, one could be rationally alienated from a desire to dance. But it seems to me that this case is somewhat difficult to imagine oneself on the inside of—largely, I think, because the good features of dancing are quite readily apparent to most of us (even those of us for whom its downsides outweigh these good features). So let us consider instead a character we met in the previous chapter—Alec Hinshelwood's ex-monk, who is beset by desires to harm himself when he experiences erotic thoughts or feelings. This man, recall, no longer endorses any of the ideas that might once have made such desires make sense to him. He understands, we may presume, the causal origins of his desires—in his practices with the religious community of which he used to be a

member. But as things stand now, he doesn't take them to be warranted or rational in any way, and in this sense they are baffling to him. He *knows* that his experiencing erotic feelings does not warrant self-flagellation, and yet the desire to do so comes upon him anyway.

Now how might such desires fit into his deliberations about what to do? Hinshelwood himself considers a number of ways in which these desires might feature within such deliberations. For instance, the man may decide to avoid situations in which his desires are likely to arise, or to attempt to distract himself when they do arise, or he may decide to submit to the desires in a relatively minor way (say, giving himself a little scratch) in order to prevent their becoming unmanageably powerful (Hinshelwood 2013: 244–45). All of these scenarios seem plausible as ways that the man's desires might play a role in his deliberations, but none of them seem to look quite like the ordinary case of deciding to go to the party on the basis of your desire to dance. Instead,

[The man's alienated desires] feature as one more part of the world with which he has to contend, or perhaps better, as the demands of another to which he must listen but which do not seem legitimate to him. His wanting to hurt himself when he entertains erotic thoughts thus figures, in a particular way, as a fact about himself he has to take into account when considering what to do. [...] The man is [...] forced to deliberate about the avoidance or satisfaction of his alienated desires in a way which renders them similar to facts about his height or even his external surroundings. His alienated desires constrain the projects he might sensibly undertake, or figure as obstacles which must be strategically overcome. (Hinshelwood 2013: 244–45)

This is not how desire functions within deliberation in the ordinary case. And when we try to set up *this* desire in a role parallel to that of the ordinary case, it refuses to conform to the model. For instance: suppose the man, experiencing an erotic thought, has a desire to harm himself. The desire does not seem to him responsive in any way to what is actually desirable; rather it strikes him as an alien intruder within his psyche. But nonetheless, he *does* now have such a desire ('just as', we might try to say, 'just as Ronnie has a desire to dance'). And (as with Ronnie) various courses of action will

allow him to satisfy the desire, while others will not. Leaving the company of his friends, for instance, will mean that they will not attempt to stop him from satisfying the desire; going home to be alone is one route he can take that will make it more likely that his desire can be satisfied. So he tries to decide what to do—whether, that is, to leave the company of his friends. In this case (and here the structure of deliberation departs radically from that of Ronnie’s deciding whether to go to the party) it seems like the fact of the desire actually gives him a reason to *stay* with his friends, precisely because doing so will *prevent* him from satisfying the desire.

Now, of course, the man may in any case decide to go home in order to satisfy his desire to hurt himself. If so, isn’t at least *this* scenario structurally similar to Ronnie deciding to go to the party on the basis of his desire to dance? I think it is not. There are two different ways that we might understand the man’s reasoning here. On the first, the man’s ‘decision’ to go home does *not* really feel to him like a decision that he is making at all. He *tries*, in fact, to ‘decide’ to stay with his friends, since this (it seems clear to him) would be the course of action that counts as ‘getting it right’ in this instance of deliberation. And yet, in spite of this, he finds himself—here also we may usefully speak of him ‘watching himself’, as though from the outside—getting up, fetching his coat, making his excuses, and leaving. It certainly does seem like a decision was made—but the man does not feel that he himself made it. Instead, he was swept up by an alien force that, as it were, made the decision for him.

The second way that we might imagine the man ‘deciding’ to go home *does* feel to him like a genuine instance of he himself deciding what to do. In this case, we may suppose that the increasing tension of the unassuaged desire becomes deeply uncomfortable. He may decide that the discomfort of resisting his compulsion is just too much, and so he will go home and do what is needed to relieve the discomfort (and also, perhaps, to schedule an extra session with his therapist). In this

case, to the extent that the man feels that the decision is really his own, the desire to hurt himself does *not* ground his decision in the way that the desire to dance grounded Ronnie's decision to go to the party. Instead, it is *another* desire that is doing the relevant work: a desire to avoid the pain and discomfort that his unassuaged alienated desire brings with it. And here we need not suppose that *this* desire is similarly alienated. We may suppose that his desire to avoid pain and discomfort is perfectly in line with how he finds the reasons to point. Indeed, more than this: the desire is not merely *in line with* how he takes things to be (how he takes pain, for instance, to be—namely something that warrants avoidance); rather, the desire just *is* his finding things to be such (finding pain to be avoidance-worthy). And so here we have a case of deliberation that the man resolves by deciding to go home, but the desire that grounds his deliberation is *not* the alienated desire to hurt himself; instead, it is the transparent desire to avoid the discomfort that arises from a compulsion unassuaged.

What these considerations suggest is that a desire can serve as the basis for deliberation in the way we've been supposing only because and in so far as that desire is transparent. And what *this* means is that the supposedly 'anti-realist' story about deliberation is, on closer inspection, not actually anti-realist at all. For the desire that renders certain courses of action 'better' than others is not, from the point of view of the deliberator, merely a psychological fact about herself. Rather, it is her grasp of something beyond herself that calls for or warrants such a response on her part. Deliberation on the basis of one's desire thus invokes the idea of a value-laden world as something prior to the psychological facts about her attitudes.

3.3 ONE MORE COMPLICATION

There is one more complication that I must deal with before the application of my argument to Schroeder's position will be complete. In the discussion above, I had our deliberator explicitly

reason with reference to his own desire (the desire to dance, for instance). Schroeder, however, would not accept this characterization of deliberation, for slightly complicated reasons. (Dealing with this will require that we go into the weeds of Schroeder's position, but it will I hope be worth it.)

Schroeder aims to show how his version of neo-Humeanism can avoid what he calls the "Self-Regarding objection" (2007: 25–27). Roughly, the objection is that the deliberating agent as understood by the neo-Humean is concerned with himself and his own desires in a manner or to an extent that is, apparently, unappealing. Schroeder aims to undermine this objection by first making explicit its underlying assumptions, and then showing how the neo-Humean need not accept these assumptions. The Self-Regarding objection arises, according to Schroeder, if we assume (1) that the neo-Humean must hold that the desire that explains a person's having a particular reason is itself *part of that reason*,⁵ and (2) that when a person is reasoning well, "the kinds of things about which he should be thinking are his reasons" (Schroeder calls this the "Deliberative Constraint" (2007: 26)). With these two assumptions in place, we will conclude that a person who is deliberating well must think about his own desires when deciding what to do. And this is supposed to render the deliberator "objectionably self-regarding" (Schroeder 2007: 27):

It is *objectionable* to think that when Ryan is thinking about whether to help Katie [who is in need of help], his reasoning is somehow flawed or even enthymematic, if he does not take a moment to reflect on whether he wants Katie to have what she needs. Such a view commits agents who are deliberating well and non-enthymematically to taking what Mark Johnston

⁵ This assumption follows from what Schroeder calls the "No Background Conditions View", which is the view that everything that is a necessary condition on a thing's being a reason is itself a part of that reason. This is a position Schroeder finds made more or less explicit in (Nagel 1970: 90–95) and (Raz 1999: 228), and which he accuses many other (mostly unnamed) philosophers of holding but not making explicit: "The No Background Conditions view is not the kind of theory that philosophers usually make explicit. It is the kind of view which sympathizers are apt to take for granted" (Schroeder 2007: 24). See (Schroeder 2007: ch. 2) for further discussion.

calls the *pornographic attitude*: they are moved only by considerations about the satisfaction of their own desires—even when those desires happen to be other-directed.⁶

Schroeder aims to avoid the Self-Regarding objection by denying that the desire on which a person's reason is dependent is itself *part of* the reason. Instead, it is a background condition on some other fact's being a reason (for that person). According to Schroeder, then, Ronnie's desire to dance is part of what makes the fact that there will be dancing at the party into a reason for him, but it is not itself part of the reason. This means that Schroeder can accept the Deliberative Constraint while maintaining that the agent can deliberate about what to do *without* having to think about his own desires. As a result, Schroeder claims that his own version of neo-Humeanism avoids the Self-Regarding objection. I am not sure that Schroeder is right about this (though I confess that I am not exactly sure what the real point of the objection is—I am not sure that I am clear on what it means to be “objectionably self-regarding”, nor what the significance of the “objectionableness” is). But in any case, Schroeder's suggestion that the person's desire itself is not part of the reason which it generates—and thus need not feature in the person's deliberation—does not allow him to dodge the problem that transparency raises for the neo-Humean.

On Schroeder's view, your desire to dance is not *part of* the reason that you have for going to the party. Instead, it is part of what makes it the case that another fact—the fact that there will be dancing at the party—is a reason for you to go to the party. (And, to reiterate the crucial point I made in section 3.1, there are no additional normative facts at work here; it is not, say, the desire in concert with the inherent goodness of dancing that gives rise to the reason.) So when you are deliberating about whether to go to the party, you can merely look towards your *reasons*; you need

⁶ (Schroeder 2007: 27); the reference to Johnston is to his (2001).

not cite the fact of your desire to dance. You can think, for instance: “Shall I go to the party? Well, it will be a great opportunity to dance. And Moriarty is out of town and won’t be there. I’ll go!”

If this is so, perhaps my argument about the transparency of the desire can be avoided. After all, the desire itself seems now to have slunk away from view; it isn’t part of your reason, and you don’t have to make reference to it in order to deliberate well. However, even if we allow that the desire is not in any way part of your reason, still it remains the crucial part of the full story about *why* your reasons are as they are. And it seems to me that if this is so, then your activity of trying to decide what to do ought to be compatible with your knowledge of this full story. If it is true, as Schroeder claims, that part of why your reasons are as they are, then you ought to be able to acknowledge this true thing at the same time as you consider those reasons for the purposes of your deliberation. So suppose you reason as I sketched above, noting that the party will be a great opportunity to dance. You are *right*, on Schroeder’s account, to have considered this fact—for the fact *is* a reason for you; it is part of what makes one particular course of action (going to the party) count as ‘getting it right’ in this instance of deliberation. And, further, you don’t *have* to make explicit reference to your desire to dance. Nonetheless, your reasoning here ought to allow for a certain kind of supplementation—namely thought about what makes it the case that your reasons are as they are—without thereby undermining what had originally made it seem attractive and intuitive, in Schroeder’s account, that your reasons *are* as they are. And now the issue of the transparency of desire again arises.

On Schroeder’s account, the full story about what makes it the case that certain courses of action count as ‘getting it right’ depends on the facts of the deliberator’s desires. Going to the party counts as ‘getting it right’ *because* you have the desire to dance. But this idea itself only makes sense from your point of view to the extent that the desire to dance is something more than an opaque bit

of given psychic baggage. To the extent that you are alienated from the relevant desire, its ability to render certain courses of action ‘better’ than others is undermined—or, as I argued in section 3.2, can play this role only through dependence on another attitude, one which is aimed (for instance) towards the avoidance of the discomfort of a compulsion resisted, and which is itself transparent.

The problem for Schroeder, then, is this: deliberation about what to do may not *require* that one explicitly think through the full story of what it is that makes a certain course of action count as ‘getting something right’. But it must, I am claiming, *permit* such a thought. And as soon as we begin to flesh out this full story, we run up against the significance of the transparency of attitudes.

Schroeder had hoped to avoid certain problems arising from the idea of the deliberator’s inward glance (as he spells out the problem, it is an issue of this inward glance being “objectionably self-regarding”) by taking the desire out of the *reason* itself. However, I did not set up my tasks in terms of the contents of a person’s *reasons*, but rather in terms of the idea that deciding what to do involves a commitment to the thought that there could be something that counts as getting it right. And on Schroeder’s account, it is the desire that makes it the case that some course of action counts as getting it right. But this fact itself only makes sense from the point of view of the person trying to decide what to do when the desire points beyond itself—to the *desirable* out in the world. Where the desire is opaque, Schroeder’s story about why we should suppose that there is a reason here at all ceases to make sense to the deliberator.

4. ON WHAT COMES FIRST

4.1 FURTHER GENERALIZING THE ARGUMENT

So far, I have made use of my account of the transparency of the first-person perspective to provide a generalized argument against the anti-realist explanation for the ‘getting something right’ feature of deliberation. I’ve shown how this argument works in detail in relation to one particular anti-realist account—that of Schroeder’s neo-Humeanism. It is worth noticing that my arguments above do not rely in any significant way on the idea that the attitude that fixes it that a certain course of action counts as ‘getting it right’ is a *desire*. For any attitude that admits of the distinction between transparency and opacity will face the same issues. In this way, then, we can see how the argument above might be generalized to cover a wider range of anti-realist accounts of deliberation.

What I would like to do next is to explore a broader set of implications that I believe my discussions, both in this chapter and the previous chapters, point towards. In this project so far, I have been considering questions concerning deliberation about what to do, the role that a person’s attitudes might play within such deliberations, and the relationship between such attitudes and the idea of a value-laden world. And now I would like to demonstrate how the preceding discussions of these questions might help to shed some light on some deeper questions which, though they have not always been treated as very central to the metaethical question of realism and anti-realism, are I think ultimately very important to that debate.

To this end, then, I’ll discuss first the concept of a ‘directions of fit’ distinction between two types of attitudes, and how this idea naturally leads towards a certain conception of the distinction between practical and theoretical reason. I will then discuss the claim that practical deliberation can only take place against a pre-established background of normative commitments, and that in this sense commitments of this kind must be understood (at least by the deliberator) as ‘coming first’, as

something fixed or given prior to any practical deliberation. Finally, I'll discuss Korsgaard's idea that the origin of the normative authority of some claim can only be found within me—in my own endorsement of that claim. My treatment of these issues will be of necessity far too brief and inadequate. My hope is only that I will be able to gesture towards some of the broader usefulness of the arguments that I've been offering so far.

4.2 'DIRECTIONS OF FIT'

It has seemed to many philosophers that there is an important distinction between two sorts of mental states—beliefs and other belief-like states on the one hand, desires and other desire-like states on the other—and that this distinction can be characterized in terms of a difference in the 'direction of fit' between the state and the world. John Searle is responsible for popularizing this idea within the philosophy of mind. Here is his statement of the view:

If my beliefs turn out to be wrong, it is my beliefs and not the world which is at fault, as is shown by the fact that I can correct the situation simply by changing my beliefs. It is the responsibility of the belief, so to speak, to match the world, and where the match fails I repair the situation by changing the belief. But if I fail to carry out my intentions or if my desires are unfulfilled I cannot in that way correct the situation by simply changing the intention or desire. In these cases it is, so to speak, the fault of the world if it fails to match the intention or the desire, and I cannot fix things up by saying it was a mistaken intention or desire in a way that I can fix things up by saying it was a mistaken belief. Beliefs [...] can be true or false, and we might say they have the "mind-to-world" direction of fit. Desires and intentions, on the other hand, cannot be true or false, but can be complied with, fulfilled, or carried out, and we might say that they have the "world-to-mind" direction of fit. (Searle 1983: 8)⁷

⁷ Searle claims to find this distinction in Elizabeth Anscombe's work, in particular in her example of the shopper and the detective ((Anscombe 2000: §32); see (Searle 1979: 3–4) for this claim). Likewise Mark Platts: "Miss Anscombe, in her work on intention, has drawn a broad distinction between two kinds of mental states, factual belief being the prime exemplar of one kind and desire being a prime exemplar of the other [...]. The distinction is in terms of the direction of fit of mental states with the world." (Platts 1979: 256). That Anscombe meant her case of the shopper and the detective to show what these authors take it to show is disputed; see for instance (Vogler 2001).

The idea that there is a fundamental difference in ‘direction of fit’ separating two types of attitudes leads very naturally to a certain model of how the mind and the world interact in the realms, respectively, of theoretical and of practical thought. In the theoretical arena, according to the model I have in mind, the proper beginning of the story about this interaction lies the state of the world. So-and-so is the case out in the world, and—so long as all goes well in the exercise of my theoretical reason—my mind will come to reflect that fact, which is to say that I shall come to believe that so-and-so is the case. And this is where the story ends. In the practical arena, by contrast, the story begins from how things stand in my mind: I desire (or intend, or some other attitude with a ‘world-to-mind’ direction of fit) that such-and-such be the case, and then—so long as all goes well in the exercise of my practical reason—the world will be brought to accord with my mind, which is to say that such-and-such will be brought to be the case, and then *this* is where the story ends. According to this model, theoretical and practical reason each serve to mediate between *what is the case in the world* and *what is the case in my mind*, and to bring the two into accord. But in doing so, the two forms of reason operate in opposed directions, and take as the relevant mental contents different types of attitudes suited to this directionality. In the theoretical case, the content of the world is what is given, and the mind, in particular the person’s belief, is reshaped in light of this given content. In the practical case, the content of the person’s mind, in particular some desire or intention, is what is given, and the world is reshaped in light of *this* given content.

Despite the fact that such a model does not say anything explicit about the nature of value, its connections to my concerns here should be fairly apparent. In particular, the idea that the ‘story’ (as I have put it) of the operation of practical thought is something that *begins* from (certain kinds of) attitudes has deep affinities with what I have developed as an anti-realist approach to the topic of deliberation (and, indeed, with the anti-realist’s basic claim that attitudes are prior to the realm of

value). And, as such, this idea is one that runs into difficulties in the face of my account of the transparency of the first-perspective. For what we have seen is that the attempt to begin any such story from the fact of a person's attitudes cannot survive the transition to the first-person perspective; from within that perspective, the picture of an attitude as something merely given is a picture of alienation.

What the concept of transparency shows us is that from the point of view of the person whose desire or intention it is, the fact of her possessing that attitude cannot be the true beginning of the story about the relationship between mind and world that plays out in the exercise of her practical reason. Certainly, it may well be true that she will, as a result of her desire, go on to act in ways that bring the world into accord with that desire. But—at least in the ordinary non-alienated case—her very possession of the desire already invokes the world, in particular that part of the world which she finds to be desirable. Thus, when she acts to bring the world into accord with her desire, she does so because she *first* found some aspect of the world itself to be a certain way, namely desirable. The story thus does not begin with the desire as a given fact, but instead one stage before this, with the part of the world onto which the window of desire faces.⁸

4.3 'PULLING ONESELF UP BY THE BOOTSTRAPS'

I would now like to consider a line of thought that crops up in various areas in philosophy, and that seems to push in the opposite direction to the arguments that I have been making here. What I have in mind is the claim that a person can deliberate about what to do (or, more generally, about some

⁸ The literature on so-called 'besires' (mental states that combine the direction of fit of both beliefs and desires), and the attempt to make use of this concept for metaethical purposes only serves to underscore my point here. One needs this strange hybrid concept only if one first takes for granted the directions-of-fit distinction as primary; my account gives us reason to avoid this, and thus to avoid talk of 'besires' altogether. (The term 'besire' is coined in (Altham 1986); see also (Smith 1994: ch. 4), (Zangwill 2008), and related discussions.)

practical or evaluative question) *only* from *within* a particular normative standpoint—that is, only by taking a certain set of normative judgments or commitments as ‘given’. Some thought of this kind has seemed intuitive to many people, and yet it seems to suggest that I have, in insisting on the primacy of the value-laden world for the deliberator, gotten things entirely back-to-front. I’ll work here with versions of the thought that are presented in the work of Frankfurt, Street, and Schroeder.

Frankfurt opens section 10 of his *The Reasons of Love* with the following announcement: “It is by caring about things that we infuse the world with importance” (2004: 23). Although does not explicitly present himself as in dialogue with metaethicists here, we can note in Frankfurt’s statement the characteristic marks of the anti-realist position: the priority given to certain sorts of psychological attitudes, those that come under the banner of “caring,”⁹ as something which can then “infuse” the (previously value-less) world with a particular sort of value, namely “importance”. The person who cares about things—who comes already equipped with his basket of cares and concerns—comes first in Frankfurt’s account. Now, says Frankfurt,

⁹ This is a concept that Frankfurt has explored at length throughout his career. Frankfurt’s treatment of the concept stands in a sort of uneasy tension with his work on ‘identifying’ with a desire, which he wishes to explain either via a higher-order volition (Frankfurt 1988a), or (following Watson’s critique in (Watson 1975)), “a particular kind of decision” (Frankfurt 1988b: 68). The tension—which Frankfurt is consciously struggling with—concerns, on the one hand, the idea that one may make some sort of ‘free’ choice about which desires one will identify with, and, on the other, the idea that the things that are most important to us are experienced in terms of necessities that bind us, and concerning which we don’t seem to have much choice at all (along with the above-cited texts, see for instance (Frankfurt 1988d, 1988c, 1999a, 2004)). Interestingly, I think we can trace a parallel struggle in Korsgaard’s work, with the image of the self as pure rationality, as something that can step back from all of the contingencies and matter of her psyche and *decide* amongst them, vying with the (perhaps more human?) image of the self as something that is tethered down into, and must act on the basis of, the mundane necessities of some contingent practical identity—mother, American, teacher, friend... (see (Korsgaard 1996) and (Korsgaard 2009)). For both Frankfurt and Korsgaard, my diagnosis of the problem is the same: each leaves out the crucial role of *the world* as something that can make demands of us. The world calls, and we respond; this is the way in which we are at one and the same time ‘bound’ and ‘free’ in the respects that Frankfurt and Korsgaard each attempt to explain in terms only of the person herself and what is going on within *her*.

suppose he wonders whether he has got it right. That is, suppose that somehow he becomes concerned about whether he really should care about the things that, as a matter of fact, he does care about. (Frankfurt 2004: 23)

Frankfurt thinks that this form of self-reflection will ultimately turn out to be quite futile:

Once we begin asking how people *should* live, we are bound to find ourselves helplessly in a spin. The trouble is not that the question is too difficult. Asking the question tends to be disorienting, rather, because it is inescapably self-referential and leads us into an endless circle. [...] It is not hard to see why. In order to carry out a rational evaluation of some way of living, a person must first know what evaluative criteria to employ and how to employ them. [...] The trouble here is a rather obvious sort of circularity. In order for a person to be able even to conceive and to initiate an inquiry into how to live, he must already have settled upon the judgments at which the inquiry aims. (Frankfurt 2004: 24)

And, of course, “Nobody can pull himself up by his own bootstraps” (Frankfurt 2004: 26). Thus, Frankfurt concludes:

We need to understand [...] that the ambition to provide an exhaustively rational warrant for the way in which we are to conduct our lives is misconceived. The pan-rationalist fantasy of demonstrating—from the ground up—how we have most reason to live is incoherent and must be abandoned. (Frankfurt 2004: 28)

Frankfurt is admirably explicit about where this line of thought leads. As he explains, “[t]his means that the most basic and essential question for a person to raise concerning the conduct of his life cannot be the *normative* question of how he *should* live. That question can be sensibly asked only on the basis of a prior answer to the *factual* question of what he actually *does* care about.” (Frankfurt 2004: 26)

The thought is, roughly: *some* normative commitments must be taken as fixed or given prior to any deliberation, *because* one cannot call all of one’s commitments into question at once without losing any measure by which to check the correctness of these commitments. This line of reasoning crops up in a similar manner in both Street’s and Schroeder’s work. Street claims, for instance, that

deliberation must take place from within a particular evaluative standpoint, and that any attempt to question whether one's (actual) set of evaluative commitments is the right one must be "pos[ed] from the standpoint of some further set of values".¹⁰ Street continues:

The one thing that one cannot do, coherently, is to step back from every last one of one's normative judgments at once and try to pose such questions from nowhere—asking, while suspending one's acceptance of any value that might be capable of settling the matter, whether one should endorse one's own set of values, or some other set, or none at all. (Street 2012: 52)

As a result,

there is no escaping the fact that you have to start *somewhere* as an agent—with the acceptance of some values or other—and that this starting point cannot itself be chosen for a reason, since there is no standpoint prior to agency from which one could do this. (Street 2012: 58)

Notice what has happened here: Street's own view—that one must take some normative judgments as 'given' in order to deliberate—has been presented as in opposition to the view that one can 'step back' from all of one's judgments and assess them from some neutral perspective "prior to agency".

Schroeder, too, makes a very similar move: when faced with the challenge that desires are the sort of thing that one may have on the basis of reasons (and thus that his attempt to explain reasons in terms of desires might be problematically circular) he chooses to argue against "the idea that one ought to be able to step entirely outside one's desires and deliberate about which desires to have". (Schroeder 2007: 190)

¹⁰ There is an important difference between Street's claim here, and Frankfurt's. For Street at least sometimes qualifies the claim, asserting that such must be the case only on the condition that we have already rejected realism as a metaethical option (which she does on the basis of a separate argument; see her 2006). Nonetheless, some of the things she says do suggest that, like Frankfurt, Street thinks that any attempt to make some kind of practical or evaluative decision *must* rely on certain other evaluative principles or commitments being treated as 'given'—regardless, that is, of whether realism has been taken off the table or not.

Here, then, the notion of ‘stepping back’, first introduced in Chapter 2, has returned in a new guise, and playing a new argumentative role. Now, however, we are in a better position to think about this idea and its problems, and in particular to think about the opposition that appears to be forced upon us here. I have been arguing that an attitude (whether a desire, normative judgment, or whatever) that is treated as merely given *cannot* function as the basis for deliberation about what to do. Now we face the counter-claim that, in fact, one *must* take (at least some) such content as given if one is to deliberate at all.¹¹ And this claim is supported on the basis that one cannot ‘step back’ from and call into question all of one’s evaluative attitudes at once.

But I do not think we ought to accept the dichotomy that we are being offered here, between having to treating some attitudes as ‘given’, and being able to ‘step back’ from an evaluative all of one’s attitudes at once. For it seems to me that *neither* of these stances towards one’s own attitude(s) looks like what I have defended as the ordinary first-personal form of self-relation. What it being proposed is that, for some given attitude, one can *either* open the attitude up to question, by ‘stepping back’ or ‘distancing’ oneself from it, *or* one treats it as ‘given’, as something for which the question of its correctness, its relationship to how things stand in the world, is considered *closed*, not-to-be-asked. But *both* of these stances treat the attitude as opaque, as an object that I have to decide how to relate myself to: shall I just take it as given that in fact I value x? Or shall I ‘distance myself

¹¹ In case one is concerned that ‘givenness’ means something different in these two claims, notice that for Frankfurt, Street, and Schroeder, the ‘givenness’ is put to a more or less explicitly anti-realist use: they each make these arguments in order to defend the idea that at least some attitudes must, at the end of the day, come first, such that the normative notions of reason/value/importance/goodness and so on ultimately follow upon the prior facts about the person’s attitudes. See for instance Frankfurt’s remark, cited above, that ultimately the factual question of what a person actually cares about must be prior to any normative questions that she may ask herself about what to do or how to live. And Street: “if one accepts [Street’s position], then [...] one regards the attitude of valuing as the more fundamental explanatory notion, and understands value itself as a ‘construction’ of that attitude. Things are valuable ultimately because we value them, not the other way around” (Street 2012: 41); and for Schroeder, see (2007: ch. 4).

from that value, suspend my commitment to it, in order to ‘call it into question’ (and perhaps remove it from my mental economy)?

Neither of these stances, that is, captures the ordinary form of relating to an attitude that I have characterized in the previous chapter in terms of ‘transparency’. When an attitude is transparent to the world, it just is your finding the world to be a certain way. By contrast, in the case in which one treats the attitude as ‘given’, the question of how the world stands is not raised at all. And in the case in which one ‘steps back from’ the attitude in order to assess it, if the question of how things stand in the world *is* brought onto the scene, it will be in order that the attitude be held up and checked against that reality—to be checked for an adequate match, perhaps, between the two. But again, this is not the ordinary case in which your attitude just is, already, your finding the world to be a certain way.

In this way, then, my account of transparency can help us to see why the worries about ‘pulling oneself up by one’s bootstraps’, or ‘stepping back from all of one’s evaluative attitudes at once’ are wrong-headed—*not* because one can, in fact, do such a thing, but rather because the proposed dichotomy between the possibility of such stepping-back, on the one hand, and the suggestion that one must take *some* evaluative commitments as ‘given’, on the other, is a false one. For neither option properly captures the role of the world in the relationship one bears to one’s attitudes when all goes well; neither option properly captures the idea of the attitude as a window onto the world, as a mode of seeing.

4.4 ‘THE AUTHORITY OF REFLECTION’

In this section I will discuss a pair of claims that together lie at the heart of Korsgaard’s rejection of realism and her commitment to neo-Kantianism. The claims are these: first, that realism, since it understands ethics as a branch of knowledge, cannot provide an answer to what Korsgaard calls ‘the

normative question'. The normative question is one that arises for the agent when she finds herself doubting the authority that some purported ethical demand has over her—when she is inclined to ask, for instance, *why* it is that she must do her duty. The second claim is that the only thing that *could* provide the relevant authority—authority that can be understood to be such from the point of view of the agent herself who is trying to decide what to do—is her own reflective endorsement.

This pair of claims is especially interesting to me because of how explicit Korsgaard is that what matters to her is how things look from the point of view of the person deliberating about what to do. And I think that on both counts—the claim about the impotence of realism in the face of the deliberator's worries, and the claim about the power that her own neo-Kantianism has to assuage these worries—Korsgaard is quite wrong. Korsgaard's pair of claims, as I shall indicate below, turn out to have certain resonances with the ideas that I have already discussed in the preceding two sections ('directions of fit', 'pulling oneself up by the bootstraps'). And as with these two previous ideas, I believe that my account of the transparency of the first-person perspective can help us to understand why Korsgaard is wrong.

Korsgaard notes that 'the scientific worldview' is often raised as the central problem for realism, and cites Mackie's 'argument from queerness' as the *locus classicus* of this worry (Mackie 1990: 38–42). However, as far as she is concerned, realism faces a deeper and more serious problem than this:

What is really wrong with substantive realism is its view about the source of normativity. [...] According to substantive realism [...] ethics is really a theoretical or epistemological subject. When we ask ethical questions, or practical normative questions more generally, there is something about the world that we are trying to find out. (Korsgaard 1996: 44)

"But", asks Korsgaard, "isn't ethics supposed to be a practical subject, a guide to action?. And the difficulty here, according to Korsgaard, is that in "conceiv[ing] ethics as a branch of knowledge,

knowledge of the normative part of the world” (Korsgaard 1996: 37), the realist cannot arrange things so that the demands of ethics have the proper practical import. In particular, realism cannot explain the normative authority that ethical demands have over us—it cannot, as Korsgaard puts it, give a satisfactory answer to “the normative question”. She explains:

The normative question is a first-person question that arises for the moral agent who must actually do what morality says. When you want to know what a philosopher’s theory of normativity is, you must place yourself in the position of an *agent* on whom morality is making a difficult claim. You then ask the philosopher: must I really do this? Why must I do it? And his answer is his answer to the normative question. (Korsgaard 1996: 16)

The realist, according to Korsgaard, cannot (or, as she sometimes puts it, refuses to) answer this question. This is because:

realism is a metaphysical position in the exact sense criticized by Kant. We can keep asking why: ‘Why must I do what is right?’—‘Because it is commanded by God’—‘But why must I do what is commanded by God?’—and so on, in a way that apparently can go on forever. This is what Kant called a search from the unconditioned—in this case, for something which will bring the reiteration of ‘but why must I do that?’ to an end. [...] The realist move is to bring this regress to an end by fiat: he declares that some things are *intrinsically* normative. [...] Having discovered that he needs an unconditional answer, the realist straightaway concludes that he has found one. (Korsgaard 1996: 33)

Faced with the question ‘but *why* must I do so-and-so?’ (or perhaps, ‘but why *must* I do so-and-so?’), the realist, working with the model of ethics as a form of knowledge of how things stand in the world, will answer simply that so-and-so is intrinsically to-be-done. And this, thinks Korsgaard, is no help at all.¹² Realism points us towards “the bare fact” (Korsgaard 1996: 38) that so-and-so is the person’s duty, but bare facts cannot, thinks Korsgaard, have the relevant authority over us.

To understand why Korsgaard thinks this is so, we must turn to the second of the pair of claims that I highlighted at the beginning of this section. The realist tells us that something out there

¹² I discussed this issue in Chapter 1, section 4.

in the world has some sort of normative authority over us.¹³ And it soon becomes clear that it is not (as perhaps it is for Mackie) the idea of something's having 'normative authority' over us that is really worrying Korsgaard, but rather the *out there in the world* component. The external world is, as Korsgaard sees it, fundamentally too alien a place to provide a source for the right sort of authority, an authority that I cannot but acknowledge *as* authoritative over me. The only thing that *could* provide this, thinks Korsgaard, is my own endorsement of some external demand: "obligation derives from the dictate of the agent's own mind" (Korsgaard 1996: 31).

[T]he reflective structure of human consciousness establishes a relation here, a relation which we have to ourselves.[...] And it is a relation not of mere power but rather of *authority*. And *that* is the authority that is the source of obligation. [...] It is not the bare fact that it would be a good idea to perform a certain action that obligates us to perform it. It is the fact that we *command ourselves* to do what we find it would be a good idea to do. (Korsgaard 1996: 104–05)

By "the reflective structure of human consciousness", Korsgaard means the now-familiar 'stepping back' picture (Korsgaard 1996: 92–93). We are assailed (from 'without', or as though from without) by various impulses, perceptions, and so on, but luckily we have "the capacity to distance ourselves from them and call them into question" (Korsgaard 1996: 93). And from this position of distance, we can decide how we are to "*command ourselves*". Only then does the thing that (as the realist might uselessly tell us) "would be a good idea to do" actually have normative authority over us; only then have we answered "the normative question".

Let me briefly highlight some connections between Korsgaard's view, as sketched above, and the two general ideas that I discussed in the previous sections ('directions of fit', 'pulling oneself

¹³ Korsgaard is concerned in particular with the notions of obligation and demand, whereas I have been speaking instead of things 'calling for' or 'warranting'. The difference is an important one, but I have tried to set things up here so that I can avoid the more serious consequences of this difference in our respective concerns.

up by the bootstraps'). Korsgaard is concerned that the realist wrongly conceives ethics on a 'theoretical' rather than a 'practical' model, and the problem here seems to be that the 'theoretical' model places the source of normativity in the external world, while a 'practical model' (Korsgaard's own account) would properly situate it within the individual herself, in particular within her stepping back from and choosing to endorse, or not, her particular attitudes.

Now I have discussed in the previous section some of the difficulties that arise from the 'stepping back' model in general. But now I would like to consider why Korsgaard thinks that the act of my own endorsement is something that can *bestow* normative authority where there was none before. If the fact that it would be a good idea to do something is not enough, on its own, for the thing to have any kind of normative authority over me—to make a call on me that I consider myself actually bound to answer—then why *would* I endorse it?

Here we face a dilemma: is the fact that it would be a good idea to do the thing just entirely irrelevant to my decision to endorse it? Or does this fact in some way make my endorsing it the right thing to do (or at least one of the possible right-things-to-do under the circumstances)? The former option sounds very strange (why does Korsgaard use an example in which the thing that I ultimately end up endorsing would be 'a good idea' if this fact is irrelevant to my deciding to endorse it?), while the latter seems to contradict her claim that it is really only the act of endorsement that does the work—as well as falling afoul of the various complaints that she has already leveled against the realist.

Consider, in light of this, the way that Korsgaard's conception of the distinction between the theoretical and the practical relates to my previous discussion of the idea of 'directions of fit'. Korsgaard's arguments draw on the idea that in the theoretical case, the relationship between mind and world must begin from something in the world, while in the practical case it must begin from

something in the mind. In section 4.2 I attempted to undermine in particular one side of this picture, arguing that the transparency of the first-person perspective shows us that the ‘practical’ story must begin in something out in the world, something to which the person’s attitude (the ‘in her mind’ part of the story) is formed as a response. Now I think that what Korsgaard’s anxieties about the normative inertness of (merely) theoretical knowledge show is that the relationship between theoretical and practical reason must be shown to be yet more intimate even than this. It is not only (as I argued above) that practical reason must begin from a place of receptivity to the world. The other side of the coin is that theoretical reason’s receptivity to world—its acquisition of what Korsgaard calls ‘bare facts’—must be understood as something that already presents itself as fundamentally *relevant* to practical reason. Certain kinds of ‘bare facts’ must be the sort of thing that are, already, enough to have practical relevance for me without needing some additional stamp of approval in the form of my reflective endorsement.

4.5 CONCLUSION

Let me recap what I have done in this chapter. I have taken the transparency-based account of the first-person perspective that I developed in the previous chapter and tried to draw out its metaethical significance in various respects. Returning to a task I set myself in Chapter 2, I showed how my concept of transparency helps to clarify the role that a person’s attitudes play in her deliberations about what to do, and how this clarified picture undermines the anti-realist schema for deliberation. I then demonstrated how this works in detail in relation to one particular anti-realist position, that of Mark Schroeder’s neo-Humeanism.

In Part 4, I tackled some more general metaethical ideas and debates, and began to sketch out ways that I think my account allows us to rethink these topics in ways that will be helpful to the cause of realism. I discussed three broad lines of thought: first, the idea of a ‘directions of fit’

distinction between two types of mental state, and a related model for thinking about the distinction between practical and theoretical reason, arguing that the concept of transparency encourages us to reject this model of the distinction. Second, I considered the argument (offered in various forms by Street, Frankfurt, and Schroeder) that one can only deliberate by beginning from at least some evaluative commitments or practical attitudes which are taken as ‘given’—a claim that stands in direct opposition to the claim that I argued for previously in the chapter, namely that one can only deliberate on the basis of attitudes that are *not* taken as ‘given’ but are rather seen-through to a world that calls for them. And third, I turned to Korsgaard’s idea that only something ‘in me’ can render a claim genuinely authoritative over me, and that this is why the realist—who understands value as something discovered out in the world—cannot ultimately account for this authority.

Very roughly speaking, these three topics all circle around a set of questions concerning (1) the relationship between the mind and the world, (2) questions of priority between the two—of which (mind or world) is to come first in our account of that relationship, and (3) the distinction between the practical and the theoretical case in relation to each of (1) and (2). And in raising such questions, clearly I am moving beyond anything that I could hope to tackle adequately here. My goal in this part of the chapter was to show how my ideas here might be situated in relation to these much larger issues.

I can, I think, say the following in conclusion. My anti-realist opponents are united in thinking that one can begin from the fact of the person and her attitudes, and with this entity in hand *then* go on to ask questions about the reality (or otherwise) of value, and its location (or not) out in the world. (I set up and argued for this framework in Chapter 1.) For my opponents, the person and her attitudes comes first. But what the transparency of the first-person perspective tells us is that from the point of view of the person herself, this conception of she herself dissolves. The

idea that she is a *thing*, an object of knowledge and scrutiny, in possession of a set of attitudes which are themselves objects of knowledge and scrutiny, drops away, and in place of these ideas the world itself fills her vision and takes over as the relevant object of scrutiny.

It is this image—that of a substantive self dissolving in the face of the world to which she is, when all goes well, transparent—that first suggested to me the quotation from Simone Weil that I set as an epigraph to this chapter:

Attention alone—that attention which is so full that the ‘I’ disappears—is required of me. [...] The poet produces the beautiful by fixing his attention on something real. It is the same with the act of love. To know that this man who is hungry and thirsty really exists as much as I do—that is enough, the rest follows of itself. (Weil 2002: 118–19).

I acknowledge that Weil’s remark may seem to stand in rather a tenuous connection to my arguments here. After all, what Weil has in mind is a spiritual-religious requirement that one fix one’s attention on what is real, and I certainly do not take the spiritual-religious to be my business here. Likewise, she is speaking of how the artist is able to produce great art, and about love, and about other people existing as much as you do, and again these are not topics which have arisen in the course of my arguments here. Still, I see in these words of Weil’s a sort of gesture towards the modes of *ethical* thought to which I take my (here strictly *metaethical*) account to be especially amenable. (I shall have something to say about love and attention to the real in the following chapter.) Finally, Weil’s assertion that “[t]o know that this man who is hungry and thirsty really exists as much as I do [...] is *enough*” (emphasis mine) provides, I hope, a vivid and concrete sort of illustration of what I had in mind at the end of the previous section when I said—speculatively, optimistically—that an adequate realism ought to understand theoretical reason’s knowledge of ‘bare facts’ as something that can *already* speak to practical reason, that does not require mediation via some additional attitude of mine in order to be rendered practically relevant.

5. A PUZZLE REMAINS

My arguments in this chapter leave us with a problem. I have argued that we should understand a person's attitudes as something like transparent windows onto a world that calls for certain sorts of response—a world, that is, that is already perceived as value-laden. But now, we might plausibly wonder, what are we to make of the simple fact the Schroeder started with, namely the *difference* between Ronnie's and Bradley's reasons? If Ronnie likes to dance, and Bradley hates it, then the metaphor of the window seems to run into an interesting sort of difficulty. For if Ronnie and Bradley's attitudes are each windows onto the ('the') value-laden world, then *why aren't they seeing the same thing?*

More generally, recall that in Chapter 2 I argued against Enoch and Korsgaard's 'stepping back' pictures of deliberation in part on the basis that it was difficult to see, on their accounts, how we are to understand what is happening when a person is faced with multiple options, all of which are rationally permitted, none of which is rationally demanded, and *one* of which they simply happen to like best. In such cases, I suggested, we ought to be able to understand some sense in which the desire itself is doing some sort of 'rational' work in warranting the choice of *that* option over the other (equally permissible) options.

But now I have argued—or so it seems—that a mere desire as such cannot do this kind of work. A desire can stand as the basis of deliberation about what to do, making certain outcomes of one's deliberation count as 'getting it right', only to the extent that this attitude is transparent—only to the extent, that is, that it implicates a value-laden world beyond the attitude itself and to which the attitude is a response. And now we run up against a version of the problem concerning the

difference between Ronnie and Bradley. We run up against what I shall call the problem of 'the idiosyncrasy of the practical'.

Where the attitude in question is a belief, the problem in question does not seem to arise (or in any case certainly not with the same force). The attitude of belief is, in the non-alienated case, a window onto the world. Where I am rationally alienated from a belief, I will not experience it as a window in this way. And if you and I both have ordinary non-alienated beliefs about the same thing, then to the extent that our beliefs are in conflict with one another, *at least one of us is wrong*. A 'difference' between what we each see in this case suggests that there must be, for at least one of us, something like a distortion in the glass, or a smear of dirt that obscures the true picture. (In the extreme case, it may be that one of us is staring at a painting hanging on the wall rather than out onto the world at all...) I have been arguing, however, that practical attitudes too should be understood as window-like in roughly the same sort of way. My desires, in the non-alienated case, are my finding some part of the world to be desirable. And yet if you and I differ in some of our practical attitudes, it would seem strange to have to conclude that such differences must amount to at least one of us facing some sort of 'distortion in the glass'. What are we to make of this puzzle?

CHAPTER 5: THE IDIOSYNCRASY OF THE PRACTICAL

[T]here are many and various forms of human excellence which will not all fit together into one harmonious whole...

Bernard Williams¹

We differ not only because we select different objects out of the same world but because we see different worlds.

Iris Murdoch²

1. INTRODUCTION

In the previous two chapters, I have argued that our first-personal relationship to our own practical attitudes is not compatible with the use that the anti-realist needs to make of these attitudes in her account of deliberation. I cannot understand my own attitudes as merely a set of pre-existing data upon which I can base my decisions about what to do; rather, they function more like a (somewhat, and sometimes) transparent window onto the world, a world which I understand as calling for a certain sort of psychological response on my part. To the extent that I *do* find myself merely saddled with a particular attitude, that attitude is one that I experience as alienated, and which as such cannot ground my deliberations about what to do in the ordinary way. Where my attitudes are not alienated in this way, they can serve as a basis for my deliberations only because I understand them as revealing or reflecting some value-laden or reason-giving part of the world. Thus, my activity of

¹ (Williams 1985: 153)

² (Murdoch 1999: 82)

deciding what to do does not terminate in the mere psychological fact that I possess some particular practical attitude, but rather in the value-laden part of the world to which that attitude is a response.

The preceding part of my argument, then, is my attempt to bring the value-laden world into the picture by showing how our attitudes should be understood as transparent windows onto that world. However, in presenting a person's attitudes as window-like in this way, it may seem that I have lost sight of the *person* herself. When we picture the individual and her attitudes as essentially, even if imperfectly, simply a clear and colorless window onto the world, then it may seem that anything that is specific to *her* as a person (her particular desires or preferences, the things that are important to her and around which she chooses to structure her life, indeed any notion of her character or personality) must be interpreted as imperfections or distortions in the glass.

In its general form, some version of this problem has troubled moral philosophers of various stripes. Korsgaard, for instance, discusses the “worry about what’s left of the individual” in relation to her own view. She connects this worry to Williams’ critique of utilitarianism—“According to Williams, utilitarianism deprives the moral agent of her integrity or individual character, because it does not allow her actions to be guided by commitments to a set of people and projects that are distinctively her own. But these are the very commitments that make us who we are as individuals and give us reasons for caring about our own lives” (Korsgaard 2009: 207); see (Smart and Williams 1973: 108–18).³ Indeed, this concern has been a prominent theme throughout Williams’ writings (and the Kantian is not let off the hook any more than the utilitarian in this regard)—see for instance (Williams 1981b), (Williams 1981a), (Williams 1985: ch. 4).

³ Korsgaard adds that this worry is what “moved Thomas Nagel [...] to reverse his own earlier defense of the publicity of reasons” (Korsgaard 2009: 208); in Chapter 2, we saw how this idea plays out in Nagel’s work as an endorsement of what is essentially neo-Humeanism about those reasons which are not “agent-neutral” (Nagel 1986: 164–74).

The general form of the problem is thus clearly interesting in its own right. But, as I explained at the end of the previous chapter, the problem also takes on a particular shape within the context of my argument here. For I am trying to give an account of what you're up to when you are deciding what to do, and my claim so far—that, ultimately, what you do is to look outwards towards a value-laden world rather than inwards towards a set of psychological facts about yourself—seems to stand in tension with a point that I emphasized back in Chapter 2, namely that in plenty of cases, your own particular practical attitudes *do* seem to play a crucial role in your attempts to decide what to do. In that chapter, the cases that I focused on were those involving the sort of thing that we may describe as mere taste or preference, as when I decide to buy myself a particular flavor of ice cream because it happens to be my favorite. These sorts of cases, I suggested, were the ones that have typically been understood by realist and anti-realist alike to be explicable (even from the first-person point of view) in neo-Humean terms (Chapter 2, 3.3). That is: my desires happen to point a certain way, and this is what makes it make sense for me to make the decision in the way that I do; the psychological fact of desires come first, and the rightness of my settling the decision a certain way follows upon this. But the arguments of the preceding two chapters are supposed to undermine this possibility. For an attitude to ground decision-making in this way, I have argued, it must point beyond itself to some source of value in the world. Unless I can make sense of these cases of deciding to do something on the basis of a mere personal preference within the context of the realism I am advancing, these examples thus have the potential to function as powerful counter-examples to my account.

The possibility of choosing to do something on the basis of preference makes for an especially vivid example of the way in which things that are personal or idiosyncratic to an individual can play an essential role in allowing her to settle her deliberations. But the phenomenon that I am

interested in and hope to make sense of here is ultimately much broader than this. For it is not just when it comes to the (supposedly) trivial matters of choosing a flavor of ice cream or a color of shoe that things that are personal or idiosyncratic to someone can be important to the question of what to do. Things such as a person's love for her children, or commitment to social justice, or deep respect for the value of art, may each make it the case that 'getting things right' in some instance of deliberation depends upon precisely those particular commitments.

The previous two chapters, then, have been devoted to arguing that deciding what to do depends essentially on my looking towards something 'out there', some source of value in the world. But now I am suggesting that it seems to be true that something 'in here'—whether at the level of a preference for peanut butter cup ice cream, or a dedication to creating a just world—can *also*, at least in some cases (perhaps in all cases), play a crucial role in deliberation. In this chapter, I shall attempt to show how these two ideas are, despite appearances, not in conflict with one another; I shall argue that what I'll call the 'idiosyncrasy of the practical' is fully compatible with value realism.

Before I do so, let me make a (far-too-brief) remark about my choice of this particular phrase to signal the broad problem that I am interested in. One of the most pressing objections that have been raised against realism is the idea that the realist conceives the role of value in our lives on too much of a 'theoretical' model. As Korsgaard puts it, the mark of realism, and her main objection to it, is that it "conceives the procedures for answering normative questions as way of *finding out* about a certain part of the world, the normative part"; it "conceives ethics as a branch of knowledge" (1996: 37). Since this seems fair enough as a characterization of realism, the realist has to account for the fact that when it comes to value, matters seem unavoidably more *personal* than they appear to be in the (purely) theoretical realm. When it comes to deciding what to do in some situation, the fact that you and I are different people, with different goals and desires and guiding

values, seems to matter deeply in a way that it does not, or certainly not obviously, when it comes to the task of deciding what to *believe* about some issue. The sense in which the realm of the practical (as opposed to that of the theoretical) is ‘idiosyncratic’ is potentially a very deep and structural one, and it is thus a topic about which all realists ought, I think, to have at least something to say. I will not here be able to address this problem with any real adequacy, but for those readers who find such a line of thought interesting, my account in this chapter (in combination with my brief remarks on the relationship between practical and theoretical thought in Chapter 4, 4.2 and 4.4) might be thought of as a first step towards doing so.

2. MY PROPOSAL

2.1 PLURALISM AND EXPERTISE

The idea that I shall be developing and defending has two components that work together to render the idiosyncrasy of the practical compatible with realism. The first is value pluralism: the world contains many different forms of value. The second is an ‘expertise’ model for understanding individual people’s relationship to the (plurally) value-laden world. That is, I shall argue that different people possess different forms and degrees of expertise when it comes to observing, experiencing, understanding, and engaging with different forms of value found in the world, and it is these differences in expertise that will allow us to explain the idiosyncrasy of the practical within a fully realist framework. Value is out there in the world, but we each are—for a wide range of reasons, which I’ll explore below—able to access certain pockets of that value more clearly than others, and some will inevitably remain entirely opaque to us. I’ll discuss the modal status of that latter claim, and the nature of my pluralism, in section 4. First, however, I’ll explain what I mean by ‘expertise’ in

general, and develop the idea of value-expertise by way of analogy to more familiar examples of expertise.

If a person possesses a special depth of familiarity with, and/or skills regarding, some particular topic or part of the world, we can say that she has ‘expertise’ in that area. The knowledge and skills she possesses may be especially interesting or important or useful to others, and these are the cases in which her expertise may be reflected in her official title (Professor of Molecular Biology) or her career (electrical engineer), or for which she may be introduced in a panel discussion as an ‘expert’ in so-and-so. But one may also have expertise regarding more localized parts of the world which are less interesting to other people: I have special knowledge of, and abilities to make use of, my own embarrassingly baroque digital filing system, for instance. I am not going to be invited to any panel discussions on this topic, because this particular expertise is relevant only to myself. I am nonetheless quite likely the world’s leading ‘expert’ on the topic. One may, then, have expertise to greater and lesser degrees, and in all sorts of areas, from the minor to the grand.

2.2 EXPERTISE AND WHAT A PERSON SEES: A STORY

The expertise a person possesses often makes a difference to how she experiences the world that she moves through. It can affect what she notices, what she sees as salient, and even to some extent what she can see at all. For illustration, let me set up a scene: imagine a room containing a desk with piles of books and papers, a chair, a bookcase, a waste-paper basket, plants in pots on the windowsill. We may suppose it is someone’s study. If a (normally-sighted, awake, non-hallucinating, etc) person enters the room, we know where they will find themselves and what they will see (the desk with its piles of books and papers, the chair, the bookcase, the waste-paper basket, the plants). But depending who the person entering is, we may be able to say more about what they observe, and perhaps even *much* more:

- enter the building's janitor, who sees instantly that the room's occupant has (once again, after countless requests that he not do so) absent-mindedly put his banana-skin into the waste-paper basket, so that the janitor will have to fish it out himself if he is to be able to put the paper into the recycling.

- enter the professor whose study it is; he sees no banana-skin (though it's not that he would be *unable* to see it if it were pointed out to him—again—by the janitor), but notices that the papers on his desk are askew, for his absentmindedness is, as is often the way with academics, highly selective, and when it comes to keeping his desk organized he is meticulous. The window is closed and locked as always, since the plants don't like draughts, so it's likely, notes the professor, that someone has been into the room since the last time he was here and has (accidentally, or—though the professor himself is unlikely to register this possibility—in a minor act of banana-related revenge) brushed up against a stack and knocked it over.

- enter the professor's old friend, who like the professor is a botanist and specialist collector of orchids. Her attention is caught by the windowsill, which stands in the shade of a neighboring building, thus receiving the perfect amount of indirect diffused sunlight bouncing off its pale walls. What the botanist sees is that the windowsill is home to some extremely beautiful examples of *Phalaenopsis Corningiana*, a rare orchid species native to Borneo.

- enter a private detective: he sees first with his nose (they call him 'the bloodhound'), and beneath the heady and tropical scent of hot-house vegetation and over-ripe banana, there is stale coffee, and, faintly, something sharp and antiseptic. The girl's been missing for a week now, and while the professor himself isn't an obvious suspect, the detective is interested in his mysterious colleague, whose passport (inspected quickly and silently by flashlight two nights ago) showed recent stamps from Malaysia, Brunei, and Indonesia, and who was spotted in a heated argument with the

girl's boyfriend the night before she disappeared. The room looks well-cleaned, too, and probably not by the janitor, who—if the rest of the offices under his watch are anything to go by—tends to ignore the top of the bookcases and the gaps in the radiators for as long as he can get away with. What the detective sees, narrowing his eyes, is a possible crime-scene.

We can imagine, that is, that even though each of these people enters the same room, nevertheless what they experience or observe on entering depends on many things that are personal to those individuals—their familiarity with the room and its contents and history, the reasons for which they have entered the room on this occasion, the background knowledge they have about some particular topic, and so on. My interest here is in those features of the person that can be gathered under the broad heading of 'expertise' that I outlined above. In the examples of these various characters entering the study, we can see how their specific areas of expertise affect their experience of the space. It affects what they observe, how they conceptualize what they observe, what sorts of courses of action they see the space as open to, and so on. Similarly, in other cases, expertise will affect for example whether a person looking out of a window sees *a tree*, or *an oak tree*, or *my mother's beloved old oak tree*, or *an oak tree growing unhappily in what is likely to be overly alkaline soil*; it will affect whether she sees hanging on the wall *a painting of a pastoral landscape* or *an oil painting with some damage from sunlight in the left corner*, or *ob-my-god-surely-not-but-it-really-might-be-that-missing-Turner*.

The final character I imagined entering the professor's study was the private detective, as portrayed in a certain genre of fiction, film, and TV. The special expertise of the private detective, and in particular his or her distinctive experience of the world, is a familiar trope of the genre. It is nicely represented in the recent BBC adaptation of the Sherlock Holmes stories: when Sherlock and John Watson meet for the first time, we are as viewers as baffled as John is when Sherlock knows, seemingly on the basis of nothing, that John has just returned from war ("Iraq or Afghanistan?").

For, watching over Sherlock's shoulder, all we see is a man with a limp entering the room. *We* know him to be John Watson, but Sherlock has not met this man yet, and we see nothing that could possibly be enough to give him reason to deduce the various things he claims to know ("I know you're an army doctor and you've been invalided home from Afghanistan; I know you've got a brother who's worried about you but you won't go to him for help because you don't approve of him, possibly because he's an alcoholic, more likely because he recently walked out on his wife; and I know that your therapist thinks that your limp's psychosomatic, quite correctly I'm afraid"). Later on, when John asks Sherlock how he knew these things, we return to this scene, and are shown it through Sherlock's eyes. Sherlock's introduction to this segment is telling: "I didn't *know*, I *saw*."⁴ In illustration of the nature of Sherlock's 'seeing', the camera now moves and cuts dynamically, circling the objects in the space, zooming in on and freezing specific details, all while Sherlock guides us through what he saw via voice-over. Now, instead of seeing simply a man entering a room, we are shown the details of his haircut, his tan-lines and the scratches on his cell phone, and the meaning of these seemingly irrelevant details is explained to us. We are shown, furthermore, how what Sherlock sees is a function of his active and physical engagement with his environment (the scanning of the room, the spinning of the cellphone in his hands so that he can see it from all angles). The expert, albeit here in the larger-than-life form of the fictional private detective, knows where to look, and what he is looking for.

Of course, this trope has its home in fiction, and perhaps the kind of expertise it portrays is fictional too (although even this is probably best understood as a dramatic exaggeration of skills that are themselves more familiar and ordinary). But this particular fiction highlights the idea, which I am

⁴ Compare Holmes' famous first line to Watson in the Conan Doyle stories: "You have been in Afghanistan, I perceive." (*A Study in Scarlet*, (Conan Doyle 2001: 11))

suggesting is *not* a mere fiction, that expertise of certain kinds shapes and structures a person's experience of the world, so that what she sees on entering some part of the world will be different to that which is seen by someone whose expertise is of a different kind.

2.3 THE CASE OF VALUE

Applying these observations about expertise to the case of value, we can develop a model on which different people's relationship to value can take on idiosyncratic forms, without undermining the idea that the value itself is to be found out in the world as something that exists prior to, and calls for, our various practical attitudes. Such a model, I will argue, fits well with the phenomenology of certain interesting aspects of our relationship to value, and in particular it helps us to account for the familiar experience of coming to 'get' something we didn't previously 'get' (learning to hear the wondrousness of Bach; developing a taste for good coffee), as well as trying to help someone else to 'get' something that *you* already 'get' (pointing out the cleverness of the camera effects, or having someone taste the really good stuff immediately alongside the cheap plonk so they can notice the difference).

Some people are especially good at experiencing and understanding the value of classical music. The reasons that they are good at this can be many and varied, and a good deal of them are likely to be accidents of that particular person's history or physiological make-up: exposure to classical music as a child, say, or the happenstance of a particularly acute sense of hearing. But one can also (so I'm told, and to a small degree I have some personal experience of this) come to appreciate the value of classical music. And often, the way this happens is that someone who can see what you cannot, or what you can as yet glimpse only dimly, guides you in developing your expertise. A friend may invite you round to listen to particular piece of music, carefully chosen, perhaps based on what he knows about you and the areas of value-expertise that you already

possess. Perhaps, for instance, your friend knows that although you don't typically enjoy classical music, you are a person who is deeply aware of the preciousness of being alive. He suspects you might be able to come to hear and appreciate the beauty of the musical expression of overwhelming gratitude for life that he can hear in Beethoven's *Heiliger Dankgesang* ("sacred song of gratitude").⁵ And so when he invites you over to listen to this piece of music, he will arrange things such that there are no distractions, allowing you to listen carefully to the music, and as it plays he points out particular moments that he finds particularly expressive or moving. And perhaps under these conditions, and under the expert guidance of your friend, you may begin to see what he can see. Even if only nascently, your vision of the world of value begins to expand and to clarify; the value of classical music comes ever so slightly more clearly into focus for you.

This process of being taught to see some particular form of value, or of helping someone else to see a form of value regarding which you have expertise, is, I think, a deeply important aspect of human relationships. Our interest in, and respect or love or fondness for, other people, is often shaped by and through these kinds of interactions. Of course, there are kinds of respect (for instance) that one may have towards another person just because they *are* a person—a sort of Kantian respect for their humanity as such. But when it comes to our feelings towards and relationships with particular individuals, often our sense of what makes them special as the particular person that they are involves the idiosyncratic areas of value expertise that they possess, and their ability to open up valuable parts of the world to us. And building and fostering a relationship with another person typically involves a reciprocal sharing of such expertise, so that the value-laden world that the two people live in and experience comes, over time, to be a more fully shared one.

⁵ Beethoven, String Quartet No. 15 in A minor, Op. 132, Third Movement.

The nature of the value-expertise that a particular person possesses is an important part of what makes them *that* particular person, and likewise an important part of what makes that particular person important or interesting or attractive to us. Above, I've highlighted the way that relationships are often built by initiating one another into particular areas of value, so that we are able to experience a more shared world of value. But it is also possible to find a person special or interesting in so far as they possess a form of value expertise that you do not possess, where there is no expectation that this expertise will be passed on to you. Sometimes, that is, particular people are interesting to us in so far as they experience parts of the world that we do not, and perhaps cannot. Not all appreciation of another person's special value-expertise must point or tend towards one acquiring that expertise oneself; sometimes, the sense of distance between yourself and another person, the awareness that they see and move through a somewhat different world, is part of what is attractive about that person.

2.4 THE BASIS OF EXPERTISE

I have suggested a model on which different people possess different forms of expertise regarding different kinds of value, and explored one way in which a person may come to develop expertise in a particular area—namely, under the guidance of someone who already possesses the expertise in question. In general, however, the explanation of why a particular person has a particular form of expertise will make reference to a variety of aspects of that person's history and physical constitution. These are aspects of the person that obey a different sort of logic to that which applies to what I've called the first-person perspective. One may find oneself possessing certain physical characteristics, and also (although this case is more complicated) being saddled with a particular history, where these aspects of oneself are experienced as 'given' *without* thereby counting as defective or alienated. In this way, a person's having a particular kind of value-expertise can have its

source in a wide variety of things to which that person does not bear any distinctively first-personal relationship. One may end up with expertise as a result (in whole or in part) of accidents of history, or biology—being born into a particular class structure or a particular part of the world or a particular family; turning out to be a certain height or in possession of certain sensory abilities;⁶ or merely happening to be in the right place at the right time. Of course, this is also true of expertise more generally, and we do not suppose that this fact undermines the claim of expertise actually to *be* expertise: to be, that is, a special ability to know how things stand in some particular part or aspect of the world.

My account, then, aims to make room for the idea of an interplay between, on the one hand, the person's attitudes as *responses* to an external value-laden world—the first-person perspective as a transparent window onto that world—and on the other, those aspects of the person that can shape and affect her attitudes *without* themselves being similarly transparent responses. Continuing with the metaphor of the self as a window onto the world of value, we can think of a person's idiosyncratic areas of expertise as something like the frame and wall that surrounds the window, giving it its particular place and shape, and determining which part of the world is thereby visible.

Thus a person is not merely, as I had worried at the opening of this chapter, a transparent pane of glass, more-or-less identical to any other such pane of glass, with any idiosyncrasies appearing only as distortions or imperfections in the glass. Instead, part of what marks the person out as being *that* particular person is, we might say, the nature of the gap into which the transparent

⁶ It's interesting to note here that those features of a person's constitution that are still commonly thought of as *disabilities* do not always serve only to limit her possibilities in regards to value-expertise, but also to expand them. The literature on disability studies offers extensive resources for this line of thought—see (Davis 2013) as a jumping-off point; the work on deaf communities in particular has provided a rich source of insight here.

pane of glass is set—it's shape, the direction it points, and so on.⁷ And my point in this section is that the question of how this gap came to have its particular shape (how the person came to end up with her particular forms of value expertise) need not affect the window's nature *as* a window. Expertise can come about through all kinds of accidental, non-rational processes without this undermining the idea that value-expertise is still a view onto a pre-existing world of value.

2.5 A CASE STUDY: LOVING ONE'S CHILDREN

One of the central thoughts motivating Harry Frankfurt's work is, as we saw in the previous chapter, a commitment to 'the importance of what we care about'. It is important, for Frankfurt, that this phrase is supposed to highlight not just the significance of *caring*, but also the significance of the individuals involved in the 'we'. It matters what 'we' *as individuals* happen to care about—the projects and people that are personally important to each of us, but which need not be taken as similarly important by other people. My account in this chapter is supposed to allow me to make sense of this phenomenon in a different (and, I think, more appealing) way than that proposed by Frankfurt. To illustrate the difference, consider what Frankfurt has to say about loving one's children:

It is not fundamentally because I recognize how valuable or important to me my children are that I love them. My love for them is not derivative from their value or their importance to me. On the contrary, the relationship goes essentially the other way. My children are so valuable and important to me just because I do in fact love them. (Frankfurt 1999b: 173)⁸

As was the case in my discussion of Frankfurt in Chapter 4, 4.3, Frankfurt's account here is not primarily pitched as a metaethical position. Nonetheless, its general anti-realist appeal is evident (and

⁷ We could continue, for it's a belaborable metaphor: whether or to what degree the glass magnifies what it faces on to, or whether it has some filter applied that brings out certain aspects or colors of the landscape more clearly...

⁸ There is a slight complication here, since part of what Frankfurt is talking about is the value of the attitude or activity of *loving*, and I don't necessarily want to disagree with the idea that loving is itself a valuable thing. My objection is only to Frankfurt's (seeming) claim that the loving is not at all itself a response to some independent source of value that the beloved person possesses.

note Korsgaard's own statement (Korsgaard 1996: 99, fn. 8) of the affinities between her account and Frankfurt's in (Frankfurt 1988a). The key Frankfurtian idea for our purposes is that in loving something—one's children, say—one bestows upon them a special significance for oneself. And, at least for an anti-realist who wishes to make use of Frankfurt's picture, this is (at some level of metaphysical description, anyway) the end of the story. One's love makes it the case that one's children are now a normatively significant part of one's world, providing one with various reasons and prohibitions and necessities, without our having to find some independent source of value beyond that love to make sense of these reasons and prohibitions and necessities.

My value-expertise account allows me to explain the distinctive love one has for one's children in a different way. The realist component of the account means that we can begin from the posit that a human being is typically an especially wondrous source of value. And certain accidents of history and nature—namely, coming by whatever means into the relationship of parenthood—turns out typically to give you a distinctive window into the particular wondrousness of those human beings who are your children. Being a parent involves a distinctive form of love in part because it involves a distinctive ability to see—to see, that is, something that is really there. Of course this does not mean that parents cannot be wrong or deluded about their children, and cannot (for instance) mistakenly view as charming a trait that is really rather horrible. But in general, on my account, parents find their children to be amazing and fascinating and lovable because their children *are* amazing and fascinating and loveable. It's just that being a parent tends to give you special expertise regarding the value possessed by your children in particular; their value is bright and visible to you because you are especially capable of seeing it.

I don't have children myself, and so I cannot speak with much authority on this matter. But the experience of loving other people in general does seem to me more like having found something

genuinely special in the world, rather than like my rendering something special to me *by* caring about it. Of course, we also must acknowledge that everyone *else* claims to find similarly special certain people (their own snotty children or boring spouses) who strike us as not so special at all. And this has made it seem as though we are faced with two options: either I am right, and everyone else is deluded (*my* partner is genuinely wonderful; theirs tiresome), or we're all in some sense similarly deluded (we each happen to love certain people, and this either—depending on the nature of one's anti-realism—renders those people special or special-to-us, or alternatively makes us project an image of value or specialness where in fact there is none). My account opens up a third option, which is that my love for specific other people is a result of my having special insight into the particular pocket of value that those people really do possess.

3. THE BULLET I AM BITING

3.1 TASTE AND PREFERENCE AS VALUE-EXPERTISE

So far my instances of idiosyncratic value-expertise have been such things as an ability to appreciate the wonders of classical music, or the special depth of value that one is able to see in one's own children. These are certainly significant and noble examples of value-expertise. However, recall that the problem I originally set for myself was cast in terms of the need to explain the rather more mundane phenomenon of personal taste and preference, in particular when it comes to such matters as choosing a flavor of ice cream, or the color for a new pair of sneakers. I now need to make explicit that part of my account that is likely to seem the most strange and implausible. For I am suggesting that we should understand *these* everyday idiosyncrasies of taste and preference on the model of value-expertise as well.

What does this mean? It means that my preference for peanut butter cup ice cream should be understood as exemplifying a particular area of gastronomic value-expertise that I possess (albeit a localized, small, and unimportant one). My preference for peanut butter cup ice cream is like a tiny window that looks out onto that part of the world that is the special value possessed by peanut butter cup ice cream in particular. Of course, I might be wrong about the value of peanut butter cup ice cream; nothing guarantees that this isn't the case (though I don't think so). But if I am wrong, this will be because of a particular localized mistake I have made, not because in general there is no such thing as realist value in the realm of such things as different flavors of ice cream. On my account, the world contains all sorts of real value, great and small, including that possessed by peanut butter cup ice cream. My preference is not, therefore, a 'mere' preference, in the sense of an opaque given fact about my psyche; rather, it is my awareness of a real source of value in the world. And this claim sounds, I freely admit, somewhat implausible.

There is widespread evidence of philosophers finding such a view not only implausible, but even bordering on the ridiculous. For instance, Don Loeb's 'Gastronomic Realism: A Cautionary Tale' aims to reveal worrying parallels between moral realism and gastronomic realism (worrying, that is, for the moral realist). The specter is raised of a *reductio ad absurdum* against moral realism, since "[o]n the face of it, [gastronomic realism] seems highly implausible, even silly" (Loeb 2003: 31).⁹ Likewise, Nagel warns against "the dangers of false objectification, which elevates personal tastes and prejudices into cosmic values" (Nagel 1986: 143) and suggests that "one would have to be dotty" (Nagel 1986: 170) to suppose that the value lies in the preferred object itself. Plenty of

⁹ Interestingly, Loeb in the end admits to having changed his mind part way through the project, and claims to find gastronomic realism more plausible than he had at the start (Loeb 2003: 47). He therefore ends up presenting a different set of conclusions to those suggested by his ominous 'cautionary tale' title.

philosophers from all across the metaethical spectrum are likely, then, to find my position on this matter *prima facie* implausible. Let me try to make it seem more plausible.

My response comes in two parts. First, I shall offer a defusing explanation for the initial sense of implausibility that one may have in relation to my view. And second, I shall offer a set of considerations that serve as positive support for the position. Before providing these negative and positive components of my response, however, it is worth clarifying where I currently stand in relation to my opponent, and the weight of the argumentative burdens we each face.

3.2 CLARIFICATION OF THE CURRENT STATE OF PLAY

The claim I am defending at this point in my argument is that, contra Nagel (and others), such things as personal preferences about (for instance) certain kinds of food should not be viewed as merely given psychological facts about the individual, regarding which no further evaluative questions can be raised. Rather, on my account, what these preferences reflect is a specific sort of value-expertise that the person in question possesses. Of course, people can be wrong, and think they find value where in fact there is none. But they can also be right, such that a taste for whisky (for instance) represents not merely a psychological fact about the person, but also a fact about her relationship to an independent value-laden world, namely that she has special expertise in accessing the odd and specific values possessed by whisky in particular.

I am in agreement with Loeb that my position here is, when considered all on its own, likely to strike people as implausible. But the view is not presented for consideration ‘all on its own’. Rather, the arguments of the previous two chapters serve to support it: I’ve argued that deciding what to do necessarily involves looking outwards towards a value-laden world. And the form that my argument took applies just as well to cases of deciding what to do on the basis of taste or preference as it does to more seemingly noble pursuits. Now, if ‘gastronomic realism’ (to adopt

Loeb's terminology) constitutes a completely unacceptable position—if it is genuinely absurd—then what this means is that my prior arguments face their own *reductio ad absurdum*. As a result, my opponents would be able to dismiss them without having to identify any specific flaw in my reasoning. But I take it that this is not how things stand; gastronomic realism may be *prima facie* implausible (I am granting this much), but it is not self-evidently beyond the bounds of possibility. This means that my arguments from prior chapter have not yet been undermined, and thus the implausibility of the result that follows from them must be weighed against the strength of those arguments considered in their own right. And so although I do have a challenge to answer here, the *prima facie* implausibility of gastronomic realism thus does not mean that I begin my arguments in this chapter entirely on the back foot. My prior arguments still count in support of my position. With this starting point in mind, let me begin my further defense of the position.

3.3 DEFUSING EXPLANATIONS FOR THE SEEMINGLY IMPLAUSIBILITY OF THE VIEW

In this section I offer three explanations for why the view I am defending here will tend to strike us as implausible. These are presented as defusing explanations of various kinds—that is, they are intended to explain *away* the sense of implausibility, by showing how we are likely to mistakenly find implausible something that is in fact perfectly plausible.

First, we should note that matters concerning which flavor of ice cream or color of shoe a person chooses are not typically terribly important. Even if there is something to 'get right' here, it usually does not matter very much whether or not we *do* get it right. This relative lack of importance may easily be mistaken for the absence of there being any such thing as 'getting things right' in this arena, since for practical purposes it won't make much of a difference whether you suppose that there's no such thing as getting it right, or whether you suppose that there is or might be such a thing, but that it doesn't matter much that you *do* get it right. The unimportance of many matters of

taste and preference thus explains why we might be disinclined to think in realist terms about such things.

This suggestion is strengthened by the observation that as we move from the case of ice cream and footwear to the more (supposedly) serious end of the aesthetic spectrum, the *prima facie* implausibility of my view appears to fade accordingly. When first introducing my view, I gave the appreciation of classical music as an example. And here it is less obviously odd to suppose that a person's preference for a Mozart sonata over the Peppa Pig theme music reflects their grasp of something actually valuable in the former. Thus it is likely that the relative unimportance of the value of such things as ice cream flavors may have confused our judgments about whether or not such value can plausibly be understood as real.

A second possible source of confusion is that in matters of taste and preference we may find that there is not much that we are able to say if someone asks us *why* we like something. If asked why you prefer savory food to sweet, it seems perfectly reasonable to respond with a shrug, or "I just do". But recall that I argued in chapter 4 that inarticulacy with regards to *why* or *how* the world calls for some particular attitude is not necessarily the same as the attitude failing to be, or even failing to be experienced as, a called-for response. In some cases, developing expertise will go hand-in-hand with developing a vocabulary for and an explicit and articulable understanding of the part of the world that one is learning about. But this is not always the case.

Returning briefly to my beloved world of fictional detectives, note that though Sherlock Holmes is able to explain to Watson and to us exactly what he has seen and what it means,¹⁰ this

¹⁰ In the books he claims that his reasoning takes place at an unconscious level; he has become so good at picking up the relevant facts and making deductions that his conscious mind doesn't have to do any work. Still, he is able to reconstruct the thought processes after the fact for Watson (and us as readers) to marvel at.

isn't a hard-and-fast rule of the genre. Jonathan Creek, another BBC amateur sleuth, spots key details without being able to say what it is that he has spotted nor what significance it has for the case. In one particular locked-room mystery, Creek spends much of the episode being inexplicably troubled by the plumbing that has been laid for a toilet to be installed in the corner of the sealed underground bunker in which the body has been found. The memory of this plumbing bugs him, but he doesn't know why. (Eventually, he realizes what it is that he has seen: the outlet pipe is not far enough away from either wall to make space for the toilet's tank. And this turns out to be because—spoiler alert—the killer, after taking his revenge and setting up the fake suicide of his victim, then builds a second wall in front of the bunker wall, with just enough space to bury himself, the victim of a real suicide, between the two. The killer who *couldn't* have left the room in fact *didn't* leave the room, and his corpse is later found behind the fake wall.)¹¹

Fictional detectives and their fictional expertise aside, it seems likely that the aforementioned unimportance of many matters of personal taste explains why we may not possess an effective vocabulary or structure of thought with which to express the expertise that our preferences, on my account, represent. If my special insight into the sensory delights of peanut butter cup ice cream is not of much importance to anyone else, then there isn't much need for me to be able to explain the value that I have seen to someone who doesn't get it. And the lack of any need for an extensive

¹¹ Holmes' explicit knowledge and self-understanding, and by contrast Creek's baffled inarticulate flashes of insight, are each part of the charm of the characters. Sometimes the appeal comes from our uncertainty regarding just how much explicit calculation is going on in the mind of our expert detective: Miss Marple, for instance, presents herself as entirely guileless, and tends to attribute her sleuthing success to a generalized "knowledge of human nature" combined with the fact that people and situations simply happen to remind her of other people and other situations that she has seen before in her small village of St Mary Mead (see for instance (Christie 2011)). Is Miss Marple struck at first by some inexplicable sense of familiarity between the case and some particular other event, a sense that she must then work through in order to figure out whodunit, or does she know at once just what has happened and merely keeps quiet until it can be proved? As readers, we're never quite sure, and the unreadability of the sweet little old lady is part of what is so much fun.

practice of talking about such things is likely to leave me unpracticed, without a vocabulary, and tending to respond to questions about why I like this ice cream so much by saying that I “just do”.

Notice, however that in some areas we *do* have a cultural practice of talking to each other about what we see in certain foods, and in these areas people will be more able to say something about the value that they have seen. Consider wine connoisseurs, who may speak of a certain balance of sweetness and acidity, or a depth and complexity present in one vintage but not another. Or to take another sensory example about which I have rather more expertise (more, that is, than zero): I was very disappointed by the 2017 release ‘Mon Guerlain’, a fragrance from the storied perfume house of Guerlain. Taste in perfume, like taste in food, is of course deeply personal, and rightly so. But if someone wants to know why I don’t rate ‘Mon Guerlain’, I can explain myself as follows: The opening is attractive enough, but takes such care to be inoffensive that it ends up verging on banal.¹² And after an hour or so the scent becomes pretty much indistinguishable from any other reasonably nice vanilla-based department store fragrance, aimed very explicitly towards mass-market appeal, with seemingly more time and money spent developing the ad campaign than the scent itself. The disappointed is heightened by the perfume coming from Guerlain, which has made some of the great all-time classics of perfume, fragrances that still smell extraordinary and like nothing else, and which yet themselves continue to inspire tributes and imitators. This is the house that brought us ‘Shalimar’, ‘Samsara’, and my favorite, ‘Mitsouko’, which after almost a hundred years still stands as the exemplar case of the chypre category of fragrances. Against this background, ‘Mon Guerlain’ manages actually to *smell* like a cynical money-grab.

¹² This is a complaint that can’t be made of the official face of the fragrance, Angelina Jolie, who is known for—among other things—once wearing a vial of her husband Billy Bob Thornton’s blood around her neck; if nothing else, attaching Jolie’s image to this powdery-polite fragrance seems an irritating piece of false advertising.

The fact that we can develop a great deal of articulacy in talking about the nature of the value (or disvalue) that we find in certain sensory objects suggests that our lack of articulacy in other areas may be more accidental than integral to the nature of the object. There may, of course, be grey areas in which it is unclear whether some particular preference is *so* baffling or opaque to the individual that it has moved into the territory of an alienated desire—which is to say a desire that is not experienced at all as a called-for response to something in the world. But what I have suggested here is that the bare fact of inarticulacy—of an inclination to shrug when asked why you like something, or to say that you “just do”, does not automatically mean that the attitude is not in fact a window onto the world of value (compare my discussion of David Finkelstein’s fondness for Sadie, Chapter 3). And nor does it mean that you yourself do not experience it as such, even if your initial reaction to such a suggestion is to consider the idea absurd, or “silly”, or “dotty”. (I take it that one of the incidental lessons of this investigation as a whole is that giving the correct characterization of the phenomenology of a perfectly familiar experience can be oddly difficult.)

Consider again Nagel’s claim that “one would have to be dotty” to suppose that personal interests such as climbing Kilimanjaro or learning to play all the Beethoven piano sonatas “matter[s] impersonally” (Nagel 1986: 170). Recall that something’s mattering impersonally is, for Nagel, contrasted with its mattering *only* because and in so far as the individual cares about it. When it comes to climbing Kilimanjaro or learning to play all the Beethoven piano sonatas, the value of such things

depends on our individual aims, projects, and concerns [...]; they acquire value only because of the interest we develop in them and the place this gives them in our lives, rather than evoking interest because of their value. (Nagel 1986: 168)

This claim is clearly directly opposed to the account I am giving here. But immediately before the above quotation, Nagel states: “Most of the things we pursue, if not most of the things we avoid, are optional.” The idea that the individual’s commitments are the source of value, and not the other way around, is thus presented as following upon the idea that the thing in question is *optional*. It is easy to see how these ideas, and related ones, might naturally be connected together. The thought goes something like this: if value is real, if it is ‘out there’, then surely it must make its demands equally on anyone. Since there is surely no universal requirement that one learn to play all of the Beethoven piano sonatas—not even, it seems, a requirement that can potentially be overridden by other requirements—we must conclude that it would be a mistake to think of the value of learning to play all of the Beethoven piano sonatas as residing out there in the world; instead (so the thought goes) we must see any value in this activity as merely a product of someone’s having the relevant interest or desire.

One of the benefits of the account I am giving in this chapter is that it allows us to separate out questions about the *reality* of value from questions about whether the claims it makes on us apply *universally*. I argued in Chapter 1 that we should conceive the point at issue between metaethical realist and anti-realist in terms of whether a person takes value or human attitudes to come first. As I noted at the time, this is not the only way of characterizing the realist/anti-realist debate, and certainly for some philosophers their commitment to realism (for instance) may be largely *because* they are interested in establishing the universality of (some) normative demands. However, just as neo-Kantian anti-realists have sought to show how universality might be had without realism, here I am seeking to show how realism might be had without universality. Realism, I am arguing, is compatible with various different forms of partiality, with certain real values being visible and

relevant to one person but not to another—with what I am calling here the idiosyncrasy of the practical.

The third source of the sense of implausibility connected to the view, then, is a tendency to conflate together questions about the reality of value with questions about its universal applicability. Since my preference for peanut butter cup ice cream does not seem to involve any commitment to the idea that everyone else ought to agree with me about this, it can seem natural to suppose that the value of the ice cream cannot have its source in the ice cream itself, and that instead it is my preference that bestows value on the object. But in fact this need not follow: realism about the value of peanut butter cup ice cream need not imply the claim that everyone ought to share my preference. Peanut butter cup ice cream is genuinely good, and it is *my* favorite. But I think that vanilla ice cream is also genuinely good, good in a different way, and if *your* special skills lie in the appreciation of vanilla,¹³ then you will likely prefer vanilla ice cream, and that is fine. I have been told by a number of people that rum and raisin ice cream also has its distinctive charms, and though I absolutely do not see it myself, I can allow that this may be a result of blindness on my part rather than a mistake on theirs. (The question of whether any of, or perhaps all of us, are in *some* respect making a mistake in holding the preferences we do is a complex one, and depends to some extent on how we conceptualize our preferences; I consider this issue in section 4.2 below.)

We've now seen three respects in which we are liable to be misled in our intuitions about the plausibility of the idea that such things as 'mere' tastes and preferences can be responses to real

¹³ Which is, I do think, a much more interesting flavor than stereotypes suggest. Different varieties of vanilla pod, and methods of treating them, can give rise to flavor profiles ranging from fruity, to smoky, to rich and liquor-like. Again, my knowledge here is drawn primarily from the world of scent, so: notice the different facets of vanilla that are brought out by the smooth milkiness of Annick Goutal's 'Vanille Charnelle' compared to the boozy-woody 'Spirtueuse Double Vanille' from Guerlain.

value. In the next section, I'll present a positive argument in favor of the view, drawing on aspects of the phenomenology of one's tastes and preferences changing over time.

3.4 THE PHENOMENOLOGY OF COMING TO 'GET' SOMETHING

Having given a set of defusing explanations for the initial sense of implausibility surrounding my position, I now want to show how in fact the position fits certain aspects of the phenomenology of taste and preference very nicely.

According to the anti-realist account of taste and preference (as adopted by anti-realists, and also, as we've seen, a number of realists), the psychological fact of the person's preference precedes and gives rise to the sense in which the preferred object can be said to be 'valuable'. How, on this account, should we understand what happens when a person's tastes and preferences change? Well, if I begin by not liking sushi, and then I come to like it, then my psychology has changed from being one way (state A, say) to being another way (state B). When I am in state A, we may say that sushi is bad-to-me, or that I have reason to avoid sushi, or however one wants to put the normative or pseudo-normative facts resulting from my dislike of sushi. And when I am in state B, we may say that sushi is good-to-me, I have reason to eat it, and so on. These two states of affairs, although different from one another, nevertheless possess an important sort of equivalency or symmetry. At base, what has happened is just that my psychological state has *changed*.

However, this is not how we actually experience such changes in taste and preference, or at any rate not always. When I developed a taste for sushi, it did not seem to me that my tastes had merely shifted from being in one state to being in another, different, state. Rather, I experienced the shift as one of coming to 'get' something that I hadn't 'got' before. Sushi had always been interesting and complex and subtle, but I had not been able to see that before, and now I can see it. And this is, I take it, a common enough experience. When one learns to appreciate wine, or whisky, or dark

chocolate, the experience is not one of an arbitrary shift in one's preferences, but rather as of coming to see something that one had previously missed. It is not just that one has *changed*, for the change is as well an *improvement* of a certain kind, specifically an improvement in one's abilities to see or experience certain sources of value.

Do we always experience a shift in preferences as an *improvement* in one's vision in this sort of way? If so, this might give us reason to be suspicious of the appearance. After all, if I am right in suggesting that we can come to develop expertise in relation to new areas of value, then it ought to be possible to lose value-expertise as well. Expertise in general is something that can be both gained and lost: perhaps it is true that one never forgets how to ride a bike, but other kinds of expertise (mathematical knowledge for instance) are famously vulnerable to speedy decay if not exercised regularly. Since I am calling on the phenomenology of experiencing one's changing preferences as a coming to 'get' something, there ought to be some phenomenological support too for the experience of losing one's ability to see what had before been visible.

Happily, it seems there is such support. For instance: in my twenties, I was very interested in film. I liked to spend my weekends going to independent cinemas on my own to see odd art-house movies, usually European foreign language films, and I was especially keen on the works of Jean-Luc Godard and Pedro Almodóvar. My enjoyment of these films was sincere, and I believe was a reflection of their genuine value. Nowadays, my knowledge of the value of these kinds of experiences comes as though by testimony from my past self.¹⁴ When I have down-time these days, I tend to enjoy watching crime procedurals and Shonda Rhimes shows. I am not exactly alienated from my current preferences: I do think that there is something good and worthwhile about

¹⁴ See the end of this chapter for a sketched suggestion of how the notion of self-testimony might be made use of in thinking about the phenomenon of akrasia.

spending forty minutes watching *CSI: Miami*, or *Grey's Anatomy*. But I also feel that these preferences reflect a loss of something important in my own abilities—a certain patience and intellectual stamina, perhaps, which is in short supply due to the demands of my job and my PhD. The shows I currently like to watch are comforting rather than stimulating; they give me time to shut out the world, rather than enriching my life more generally. In this sense, I have experienced this particular change in preference as a loss of a particular form of value-expertise. And, again, I take it that some such experience might be common enough.

These aspects of the phenomenology of changing preference fit much more comfortably with realism than with anti-realism. If the preferences come first, and any sense of the 'value' of their object follows upon the fact of these preferences (as the anti-realist has it), then the anti-realist must come up with some explanation for why it often seems to us that a change in preference is not *merely* a change, but rather a clearing—or, sometimes, a clouding—of one's vision of something outside of oneself. The realist, by contrast, has the benefit of being able to say that here things just are how they seem to be.

4. DISAGREEMENT AND VALUES IN TENSION

4.1 A CASE STUDY: OPERA AND SOCIALISM

The value-expertise model allows me to make space for, and to make sense of, what I have called the idiosyncrasy of the practical. One person may be especially talented at seeing the value in good whisky, while another's skills may lie in the appreciation of fine wine. Given this account, and in particular given the realism that stands as its basis, we may naturally wonder whether or to what

extent different forms of value can stand in tension with one another.¹⁵ Might there, for example, be certain forms of expertise such that possessing one precludes possessing the other, where this preclusion is not merely a contingent matter but somehow integral to the nature of the value or the expertise itself?

It seems to me that this is possible, and that this possibility is also still compatible with realism. For example: I have it on good authority that there is a distinctive value to be found in opera. My evidence for this does not solely take the form of testimony, since in the past a patient and generous friend took me to Covent Garden to see *Turandot*, and spent some time helping me to glimpse some of what he saw in the performance. I believe that with the right sort of practice and inculcation, I could grow to develop some measure of expertise in the value of opera, and if I did so my experiences when attending the opera would be much enriched. However, it also seems to me that developing such expertise would *of necessity* mean that certain other parts of the world of value would fade from my view. In particular, there is a certain sort of political consciousness that I suspect is, in its nature, incompatible with the sort of immersion in a distinctively bourgeois aesthetic world that the proper appreciation of opera would require. Opera really is valuable, but to develop the expertise that would allow me a full and fluent experience of that value would mean lessening my grip on my socialist commitments, which are themselves a particular form of value-

¹⁵ Here again we approach the deeper question of how we conceive of the relationship between the realms of the practical and the theoretical. Given the accusation against the realist that her model of value is too ‘theoretical’, the question of whether the world of value is fundamentally ‘unified’ becomes especially interesting, since it is generally agreed that in the theoretical case things *are* so unified—or at least that our knowledge of things ought to strive towards unity. For what it’s worth, it seems to me that the complaint that the realist thinks of value in overly ‘theoretical’ terms is damning only if we work with an unnecessarily impoverished conception of what ‘theoretical’ means. My discussion of ‘directions of fit’ in the previous chapter is relevant to this issue, but as I’ve said in the introduction to this chapter, this topic is not something to which I can hope to do justice within the scope of this project.

expertise. Switching now to speaking of a tension between the values themselves, we may say that the value involved in fully automated luxury communism is in tension with the value involved in opera.

Now I have said that I believe that there *is* value in opera, and that those who spend time and money experiencing it are therefore not (not all of them, anyway) doing so foolishly. But not everyone who lacks value-expertise in the realm of opera as a result of their political commitments agrees with me here. At the beginning of *What is Art?*, Leo Tolstoy offers a vehement and very funny take-down of opera based on his experience of watching a rehearsal of a particular performance.¹⁶ This spectacle, says Tolstoy, is “one of the most gigantic absurdities that could possibly be devised”;

An Indian king wants to marry; they bring him a bride; he disguises himself as a minstrel; the bride falls in love with the minstrel and is in despair, but afterwards discovers that the minstrel is the king, and everyone is highly delighted.

That there never were, or could be, such Indians, and that they were not only unlike Indians, but that what they were doing was unlike anything on earth except other operas, was beyond all manner of doubt; that people do not converse in such a way [...], and do not place themselves at fixed distances, in a quartet, waving their arms to express their emotions; that nowhere, except in theatres, do people walk about in such a manner, in pairs, with tinfoil halberds and in slippers; that no one ever gets angry in such a way, or is affected in such a way, or laughs in such a way, or cries in such a way; and that no one on earth can be moved by such performances; all this is beyond the possibility of doubt. (Tolstoy 1979: 6)

What Tolstoy is highlighting is opera’s dramatic *ridiculousness*, its lack of truth-to-life. But the context of this passage makes clear that Tolstoy’s aesthetic objection here is driven by his having seen the huge machinery of time and money and human suffering that has gone into the creation of this absurdity. It is the human cost of the production—so many bodies tired out, all the angry demands shouted at underlings, the money spent on the dancing master “whose salary per month exceeded

¹⁶ It is unclear whether Tolstoy’s account is a pastiche, or whether he is faithfully recounting his experiences of the rehearsal of some particular opera.

what ten laborers earn in a year” (Tolstoy 1979: 3)—that renders the opera, as seen through Tolstoy’s eyes, not *merely* silly or unrealistic but instead a “repulsive sight” (Tolstoy 1979: 5).

Is Tolstoy wrong here? I have said that I think opera *is* valuable, and it seems fair to say that Tolstoy disagrees on this point. To this extent, then, I think he is wrong. And similarly, I think he is wrong to claim that “no one on earth can be moved by such performances”. (Tolstoy’s explanation for the fact that plenty of people *appear* to be so moved is, essentially, that what we have here is an Emperor’s-New-Clothes situation; patrons of the opera have bought into a false ideology regarding the value of ‘high art’, and are all deluding themselves into thinking that they are experiencing value where there is in fact no such experience.) Still, it seems to me that there are things about opera that Tolstoy is seeing aright—things that perhaps must necessarily fade from one’s view if one is to become a lover of opera. For opera *is* absurd, and furthermore this is in some respects a dark and not an innocent absurdity. The truth of the matter, it seems to me, is that opera is good and it is bad; it is a locus of distinctive forms both of value and of disvalue. And to develop one’s expertise in relation to one side of this equation will be, at the same time, a lessening of one’s grip on the other side.

These sorts of tensions are not merely a matter of the finitude of one’s time or energy or intellect; rather they find their source in something about the nature of the values themselves. But it may also be that the surrounding context is playing a role. In a great many cases, the tension between values (and thus incompatibility between the corresponding forms of expertise) is in part due to inessential features of the current social and political system. It seems likely that within our current way of life, the distinctive value-expertise involved in (for instance) bringing up children stands in some tension with the value-expertise involved in creating great art, but this needn’t necessarily be so under all possible social structures.

Still, the case of opera suggests that not all such tensions could be resolved in such a way. For opera is itself a culturally and historically specific phenomenon, one that could not survive as the sort of thing that it is without its particular cultural and historic context. As such, it is the bearer of a form of value that is itself culturally and historically specific. Perhaps something similar to the value of opera might be realizable under a quite different social order, and perhaps in this case the similar value would be fully compatible with socialist ideals. But such a thing would not, it seems to me, quite be *opera*, and those who love opera now would not necessarily be wrong in thinking that something of value had been lost in this new world order—although they may be wrong in the significance they accord to this loss, considered in light of other gains.

4.2 MUST ONE OF US BE WRONG?

For these reasons, then, I am inclined to think that values *can* stand in tension with one another, and that possessing certain forms of expertise can of necessity preclude possessing others. (It should be noted, however, that this particular claim is independent of the broader argument of the dissertation; one can reject it while still accepting the rest of my account.) In such cases, the realist can say something like what I said about opera: there *is* value here, just not value with which I am directly acquainted. And I could not (or at least it seems to me likely that I could not) come to such direct acquaintance without losing the direct acquaintance I currently have with other kinds of value. The value of opera remains out there in the world, and I am genuinely missing out by not developing my expertise in this area. But I would be missing out on something *else* if I did develop this expertise, and the importance of this something-else is vividly luminous to me right now, and is thus not something I can give up on. Acknowledging the reality of values that stand in tension with one another thus does not require one to assent to two outright contradictory statements. I can speak

coherently about the reality of values that lie beyond, or even in conflict with, my own areas of expertise.

This mode of talking is useful in illuminating the question that I raised at the end of section 3.3. If I prefer peanut butter cup ice cream, and you prefer vanilla, and some third party prefers rum and raisin, does my realism about the value of ice cream flavors mean that some of us must be *wrong*, must be making some sort of evaluative mistake? Let's leave aside the possibility that the rum and raisin chap is just totally confused, and suppose for the sake of argument that there is actually some value in that ice cream flavor. Now, as we know from the history of metaethical thought, a difference in ice cream preferences and the like is often presented as a clear example of something that isn't properly understood as a *disagreement* at all; this sort of example is then used to consider whether we ought to think of other seeming evaluative 'disagreements' (such as moral disagreements) on a similar model, say in terms of a (mere) expression of (mere) preferences. However, since I have argued that even ice cream preferences should be understood in terms of a perception of real value, the question arises whether we then *should* see you and I and our third companion as in disagreement here—disagreement, for example, about which ice cream flavor is 'best'. And if this is how we should interpret our differing preferences, then it seems we should also be able to say that at least some of us are wrong.

My suggestion here is that whether or not a particular party is 'wrong' will depend on how she conceptualizes her preference. When I described how I think things stand, I said that *I* prefer peanut butter cup, but that if *you* prefer vanilla that may be because your expertise lies in the realm of vanilla appreciation. As far as I am concerned, our preferences do not involve any contradiction, and so neither of us needs to be wrong, and likewise with our mysterious rum-and-raisin-loving friend. If you share my basic metaethical picture, then you can agree with my account of how things

stand: we all have different preferences, reflecting our different areas of expertise, and though each of us may be missing something in our perception of the world (you and I are missing out on the value of rum-and-raisin ice cream, for instance), none of us need be wrong or mistaken in our beliefs, for we can acknowledge that we are missing something at the same time as asserting our preferences.

Now suppose that rum-and-raisin doesn't accept this pluralist metaethical picture, but is fully persuaded of value realism regarding ice cream. (This is, I think, kind of an odd combination of positions to hold. But he likes rum-and-raisin ice cream, so maybe it shouldn't be surprising.) *He* is adamant that rum-and-raisin is genuinely the best flavor, and furthermore that vanilla ice cream is just boring, and that peanut butter cup is actively bad. This person's conception of his preference means, I claim, that he *is* making a mistake. Not only is he missing out on something—the value of both vanilla and peanut butter cup ice cream—but he also holds false beliefs about the things that he is missing. Thus, while a disagreement in preferences need not involve anyone being wrong, certain ways of conceptualizing and articulating those preferences may be wrong. (This is just one way in which a mistaken metaethical picture can lead to mistaken ground-level evaluative judgments.)

4.3 A NEW APPROACH TO THE PROBLEM OF MORAL DISAGREEMENT

It is often noted that there is widespread and seemingly intractable disagreement about moral and evaluative questions. This is both a philosophical problem for the realist, and a practical problem for us as people who need to live together. These are not entirely separable problems, since any steps successfully taken towards the resolution of the practical problem will at the same time lessen the philosophical burden faced by the realist. I shall end this section by highlighting the resources that

my account offers for new approaches to both forms of the problem of moral and evaluative disagreement.

First, we now have a new kind of model for understanding what is going on in cases of evaluative disagreement: on my account, at least some such disagreement is likely to be a result of differences in value expertise. Of course, knowing that this is so will not often be enough to resolve or dissolve the disagreement, for the question of (say) whether person x should leave her family given circumstances y doesn't go away just because we come to realize that the conflicting answers offered to this question are a result of people's differing forms of value expertise. However, notice that nothing in my account guarantees that there must *be* a determinate answer to this question: if values can be deeply in tension with one another, realism need not entail that there is a determinate answer to every well-formed practical question. Further, whether there is or is not such an answer, my model allows us to give a realist explanation of what is going on that gives us much more insight than merely having to insist that one party is wrong. For even if one party is wrong in the final conclusion they draw about the case under consideration, still we can allow that the considerations that lead people to their differing conclusions are real. We can allow that people may be, in their evaluative disagreements, each responding to genuine sources of value in the world.

Second, we now have the possibility of a much richer account of the particular intractability that has seemed to many to be distinctive of evaluative disagreement (according to some philosophers, it is the fact of the intractability that is the real problem for the realist). For if differing opinions are a result of differing forms of value expertise, it is not surprising that the disagreement cannot be resolved merely through *argument*. This is because expertise is much harder-won than this—one cannot, typically, be argued into expertise on some topic. Instead, expertise is a result of practice and engagement; it is the slow process of developing sets of skills and modes of seeing, of

gaining mastery of structures of concepts, of training oneself to respond to the fine detail of the world.¹⁷

Corresponding to these two philosophical points, my account offers two useful practical outcomes. First, we have a mode of thinking and talking about disagreement that can help to clarify actual disagreements. Of course, it won't necessarily resolve or dissolve the disagreement, but it can help to make one's opponent seem less baffling. Understanding, for instance, that she may be missing the significance of the value that you see in devoting one's life to great art not merely because of a blindness on her part, but instead because she *can* see with clarity something else—the value of nurturing one's family relationships, say—opens up possibilities for understanding more deeply what is going on in your interaction. And further, this sort of clarification can allow opponents to turn their attention beyond the immediate situation of disagreement and towards larger questions about the way the social structures of the world they live in may force certain forms of value into tension with one another in ways that could be overcome within a different and better system. Perhaps we can, for instance, stop arguing about whether someone is being too selfish a mother, and start working to create a world in which motherhood is no longer so incompatible with creating great art.

Alongside this approach, we also have the possibility of another sort of approach to moving forwards in cases of disagreement. Once we have accepted that certain disagreements are a result of a divergence in value-expertise, we can start to think about how we might help the other person to develop some measure of expertise in relation to the value which is moving us. Thinking in terms of

¹⁷ “Education isn't what some people declare it to be, namely, putting knowledge into souls that lack it, like putting sight into blind eyes. [...] [T]he instrument with which each learns is like an eye that cannot be turned around from darkness to light without turning the whole body.” (Plato 1997a: 518b-c)

expertise can allow us to move beyond the mode of (mere) argument, and towards a more holistic approach. Even if our opponent cannot be brought around by argument, perhaps she may be re-oriented in her thinking over a longer period of time through exposure to certain forms of experience, or through some other form of guided practice. And if this is a reciprocal process, we may likewise change and develop in our own forms of expertise. Again, of course, nothing guarantees that this will be a smooth and happy process of coming to share the same evaluative outlook, but it does open up possibilities for moving forwards in cases that might otherwise have seemed thoroughly intractable.

The purpose of the preceding discussion was to point towards ways in which my account here may be developed and put to use in thinking about the philosophical and practical problems of moral and evaluative disagreement. The above is not presented as itself an intervention into any such debates, but rather merely to signal directions that such interventions might take.

5. CONCLUSION

Let me sum up what I believe I have achieved in this chapter. Most importantly, I have set up a picture of value and our relationship to it that will allow me to resolve a particular problem that arose in the course of my overall argument here. In Chapter 2, I raised the question structuring this argument: what are we up to when we're deciding what to do? My goal was to show that the answer to this question is that we look outwards towards a value-laden world, and the arguments of Chapters 3 and 4 were intended to establish this point. However, this left me with a puzzle, for these arguments on their own left it quite unclear how it could be that I am able to make decisions about what to do that are guided by things that are idiosyncratic to me, such as my particular tastes and preferences, or my personal relationships or evaluative commitments. In this chapter, I've laid out

an account of pluralist value and value-expertise that will allow me to make sense of these things within a fully realist account. The way that this account enables me to complete my answer to the question about what we do when deliberating is hopefully already reasonably clear. In the dissertation's Conclusion, I'll spell out explicitly how all these pieces fit together, and introduce some further questions that arise for my account of deliberation.

In addition to solving the specific difficulty that arises for my argument here, my account in this chapter also has the more general benefit of showing how realism can be made fully compatible with the idiosyncrasy of the practical. This is a bonus for the realist quite independent of the argument I'm developing here, since the realm of the practical does, in general, seem to involve many aspects of idiosyncrasy. One way that this more general issue emerges is, for example, in the remark from Williams that I used as one of this chapter's epigraphs. Williams is interested in the idea that the reality of ethical life does not seem capable of being neatly captured by any one set of rules of demands:

[T]here are many and various forms of human excellence which will not all fit together into one harmonious whole, so any determinate ethical outlook is going to represent some kind of specialization of human possibilities[...].(Williams 1985: 153)

As we saw in Chapter 1, Williams is typically interpreted within metaethical circles as a proponent of neo-Humeanism. And we can see how his commitment to the ineradicably personal and idiosyncratic nature of our ethical lives might seem to point towards such an account. However, the picture I've developed here shows that this commitment and others like it need not require a neo-Humean account of value, nor indeed any sort of anti-realism. We can make sense of idiosyncrasy and the ineradicably personal without having to suppose that the source of value lies in a person's idiosyncratic and personal attitudes, rather than out in the world.

In a similar way, my account has allowed me to show how certain metaethical questions can be separated out from the question of whether or not value is real, and in this way it removes some of the burdens that the realist has sometimes been assumed to face. For instance, we saw in section 3.3 that we can understand a person's evaluative attitude to be a response to realist value without this entailing that the person must suppose that everyone ought to agree with her evaluative attitude. The reality of some particular value is thus a separate matter from its universal applicability.

As well as offering support to realism in these various ways, my value-expertise account has rich resources for developing novel arguments within several different debates in ethics and metaethics. I mentioned at the end of part 4 the way that my account might be utilized in the context of discussions of moral disagreement. And my discussion of loving one's children in section 2.5 suggests interesting prospects for developing a new philosophical account of love, a topic that has seen an increase in popularity over the last few years.

In addition to these two suggestions, my approach raises the possibility of distinguishing between importantly different ways that one may possess evaluative knowledge. Samuel Scheffler makes a distinction between, on the one hand, believing or judging something to be valuable, and actually valuing it on the other (Scheffler 2010). Within the framework of my realist account, we have the resources to make a related distinction in a more objective register: the value-expertise model allows for a distinction between ('merely') knowing something to be valuable, on the one hand, and possessing the expertise that allows one to have something more like the knowledge of first-hand experience or acquaintance, on the other.¹⁸ I may come to believe that something is valuable on the basis of the testimony of someone else, even when I cannot see the value myself,

¹⁸ Compare Iris Murdoch's reading of the post-*Phaedo* Plato, who, she says, understands the Forms as "objects of spiritual vision known by direct acquaintance rather than through the use of language (propositions)." (Murdoch 1977: 26)

and under suitable circumstances my belief may count as knowledge. But this is not the same as possessing the relevant expertise myself.

This distinction offers interesting prospects for thinking about akrasia. Recall, for instance, my claim that my own knowledge of the value of art-house cinema comes “as though by testimony from my past self”. Suppose, then, that we attempt to model the akratic individual’s (purported) knowledge that she ought not to do so-and-so as a sort of self-testimony: while she may once have known this via direct acquaintance with the relevant value, at the time of acting she does not possess the live acquaintance of value-expertise, but only a sort of memo to herself—something which can provide propositional knowledge but not the sort of richer experience that may be able to hook one motivationally.¹⁹

The final thing that this chapter aims to do is to explore a wide enough range of cases of value that the reader may, in one of these examples, find something familiar in a way that will lend a distinctive sort of support to my account. My hope is that at some point some particular example of value and its phenomenology will have resonated in a certain way for you; that I will have hit upon some particular area of value expertise that you possess, so that your own first-hand experience of value will stand as evidence in support of the account I’m presenting. Many of my examples of value have been drawn from the realm of the aesthetic—music, film, and so on. This is because there seems to me to be something compelling in the thought, which we get from Plato via Iris Murdoch,

¹⁹ I also think that these ideas about differing forms of knowledge offer promising resources for a philosophical account of the nature of depression, something that is sometimes mentioned in moral psychology only to be cast immediately aside (I did just this myself in this dissertation: see Chapter 2, section 1). The phenomenology of depression frequently involves an experience as of being shut off of from a realm of value that one recalls having been previously known more intimately; of seeing only as through a glass darkly the good that one once saw face to face. The distinction between propositional and acquaintance knowledge (expertise) of value offers the possibility of giving a satisfying account of this phenomenology.

that while “virtue in general may not attract us”, we often find that “beauty presents spiritual values in a more accessible and attractive form” ((Murdoch 1977: 36); cf. the *Symposium* (Plato 1997b)). But I have also discussed the values involved in one’s personal relationships with other people, and in political causes. So my hope is that one of these examples will have made contact with your own experience of relating to the value-laden world, and in particular that it will, in the context of my arguments here, have been revealed to you *as* a form of relation that you bear to a value-laden world.

CONCLUSION

1. WHAT DO YOU DO WHEN YOU'RE DECIDING WHAT TO DO? (REDUX)

We are now in a position to answer the question I raised in Chapter 2: what are you up to when you are trying, on some occasion, to decide what to do? Putting together my arguments from the intervening chapters, the answer is as follows. When you are trying to decide what to do, you look outwards towards a value-laden world that you hope will guide you in making your decision. What you 'see' in that world is in part a function of the particular forms of value-expertise that you possess, and this is why it makes sense that you will pursue the forms of value with which you are able to be acquainted.

Your own practical attitudes thus do play an important role in your deliberations, but it is not the role that the anti-realist had hoped they might play. These attitudes cannot, all on their own, make it the case that certain courses of action count as 'better' than others. Instead, they play this role only because and in so far as they are already, themselves, your having found some certain part of the world to be value-laden.

There is more that can be said here. At the beginning of this project, I mentioned that it is common practice in contemporary metaethics to treat a wide range of normative terms as more or less interchangeable, at least for the purposes of the realist/anti-realist debate. My arguments here give us some cause to resist this tendency. There are already points at which certain distinctions along these lines have proved to be important. For instance, in Chapter 3 I argued, against Finkelstein's critique of a transparency-based account by suggesting that we need not interpret the notion of transparency in terms of strict *obligation*, but can instead usefully move to talk of what is

‘called for’ or ‘warranted’ by the world. And now, my arguments in Chapter 5, and the conclusion I have drawn above, suggests another set of distinctions, although sadly not one that I will be able to fill out with any adequacy here. I’ll offer a brief overview.

I have argued that we deliberate about what to do in the face of a world that we find to be value-laden. But our experience of that value is of necessity limited and patchy. Further, it seems to me that there is reason to think that there’s really an *awful lot* of value out there, of a wide variety of different kinds. If I am right, this means that the world’s ‘calling for’ responses on our part is really quite a cacophony—in the face of which we’re each only able to catch bits and pieces. (“If we had a keen vision and feeling of all ordinary human life, it would be like hearing the grass grow and the squirrel’s heart beat, and we should die of that roar which lies on the other side of silence. As it is, the quickest of us walk about well wadded with stupidity” (Eliot 2000: 185).) And while I have suggested that it makes sense that one will respond to the bits and pieces of callings that one is able to catch, there is nonetheless a further question about the significance, to a particular person, of those areas of value with which she is *not* herself directly acquainted.

In particular, in developing this account, I have deliberately tried to make room for the possibility that there may be sources of value in the world with which you are not directly acquainted, but which you may nonetheless have knowledge about (say, on the basis of testimony). It seems to me that here it might make sense to ask a question of the form: what sorts *reasons* apply to a person on the basis of a source of value about which she possesses knowledge—knowledge, that is, of the fact that there is value in this part of the world, and perhaps even some things about *why* and *how* there is such value—but where this knowledge does not amount to knowledge-by-acquaintance?

One very interesting set of questions here concerns how we should understand the idea that she might, in some such cases, have reasons to pursue expertise in this arena.¹ Another set of questions concerns what reasons she might have to take her knowledge of this value into consideration in her deliberations even if she does not expect that she will ever end up with knowledge-by-acquaintance in this particular arena. For it is important to notice that knowledge of this kind does *not* (not necessarily) amount to the person's suffering from rational alienation of any kind: I know opera is valuable, though I cannot 'see' it myself—and this does not mean that I must experience either this piece of knowledge about opera, nor the lack of my ability to experience the value in opera, as a part of myself from which I am alienated. So my knowledge that there is value here can potentially enter into my deliberations without having to go via the sort of defective circuitous route that I described in cases of alienation.

How might this happen? Well, perhaps the value of something might give me reason to preserve the thing, even if I myself cannot experience its value.² As Hercule Poirot tells us, "One destroys not a thing of beauty wantonly", and this may be so even if we ourselves cannot directly see its beauty.³ If I find myself in possession of free tickets to the opera, but feel I have no opportunity (or no inclination) to adequately develop my abilities so as to appreciate this event, I will not merely throw the tickets away but instead try to find someone who will appreciate them. And I will do this

¹ Agnes Callard tackles related topics in her recent *Aspiration* (2018).

² Thanks to Agnes Callard for pressing me on a related but more difficult question to which I am not yet sure of the answer, so that I have been forced to answer this easier formulation of the question instead.

³ Poirot is here talking about his magnificent moustache, which someone has foolishly suggested he might wish to remove for the purposes of disguising himself whilst on a case: "he even suggested that I—I should—I hesitate to say it—should shave off my moustache! Yes, *rien que ça!* I pointed out to him that it was an idea ridiculous and absurd." (Christie 1974: 110).

in part because I believe that opera *is* valuable, and *not* merely because I know other people happen, perhaps foolishly, to think so.

All of this is to indicate that the realist account that I have offered suggests that there might be room to usefully separate, for instance, such terms as ‘reason’ and ‘value’, and to open up the possibility that the relationships between them might be complex:

If there is such a thing as the truth about the subject matter of ethics—the truth, we might say, about the ethical—why is there any expectation that it should be simple? In particular, why should it be conceptually simple, using only one or two ethical concepts, such as *duty* or *good state of affairs*, rather than many? Perhaps we need as many concepts to describe it as we find we need, and no fewer. (Williams 1985: 17)

2. ON WHAT I HAVE SHOWN, OR NOT SHOWN

I’ll return now to an issue I raised in Chapter 1, concerning what, exactly, the form of argument I have developed here can show. For a certain sort of interlocutor might argue that all I have *really* shown is that things *seem* a certain way from within a certain perspective.⁴ I will not be able to deal adequately with this complaint here, but let me make the following brief remarks.

First, as I stated in Chapter 1, even the more limited conclusion about ‘how things seem from within a particular perspective’ already undermines a wide range of anti-realist accounts, in particular those which claim to be giving an account that is not only compatible with but indeed derived from within that perspective (see especially Korsgaard and Street). Second, I have along the way established a set of claims which independently offer either support for realism, or increase the costs of anti-realism. For instance, an important corollary of my philosophy of mind conclusions

⁴ And that perspective, the interlocutor may wish to add, is known to be highly unreliable. Consider, for instance, the range of empirical evidence and argumentation marshaled in support of such a conclusion in (Arpaly 2003), (Doris 2015), and (Kornblith 2012). Notice, however, that I have rejected the conception of ‘the first-person perspective’, ‘the deliberating self’, and so on, that each of these philosophers work with (see Chapters 2 and 3).

from Chapters 3 and 4 is that the way of thinking about attitudes that made them especially attractive as a starting point for the anti-realist's metaethical account does not survive the transition to the first-person perspective. Attitudes had seemed unworrisome to the anti-realist largely to the extent that they were viewed third-personally, as a purely psychological fact that could be treated quite separately from the question of whether the attitude was *right* (or warranted, correct, appropriate, etc). As we saw in Chapters 3 and 4, viewing one's own attitudes in such a way is the mark not of an ordinary case of self-relation, but rather of alienation. This point, though not decisive on its own, does go some way towards reducing the appeal of an anti-realist approach to metaethics.

In addition, the model of value-expertise that I offer in Chapter 5 is supposed to render realism more attractive than it might otherwise be thought, allowing the realist to harness some of the intuitive appeal of anti-realist account. In that chapter, I separated out the question of whether the realm of value makes the *same* demands of each of us from the question of whether value is really 'out there', thus showing how there is space for personal and idiosyncratic relationships to value within a thoroughly realist framework.

Finally, let me repeat something I said in Chapter 1. In arguing that deciding what to do involves looking outwards towards a value-laden world, my hope has been to draw your attention to the ways that you yourself already seek—and, if you're lucky, find—value in the world; value, that is, as the realist understands it. This point has a double significance. First, in so far as you yourself are already seeking value as the realist understands it, you are already committed to (at least) the possibility of such value. You are in some sense, that is, already a realist about value. And second, through identifying the activity of deciding what to do as a process of seeking value, my hope is also that you might on some actual occasion of deciding what to do notice an instance of real value itself—in the wild, as it were—and be in a position to see it as such.

3. AFTERWORD: BOTH SEEING AND DOING

The impetus for this project came first and foremost from my reading Richard Moran's *Authority and Estrangement*, and the influence that Moran's work has had on my thinking here should be obvious. There is, however, another author whose role within my philosophical development has not made its way to the surface of my arguments very explicitly, but who deserves a final mention here. That author is Iris Murdoch. I am indebted in particular to her insistence on the significance, for one's philosophy, of the metaphors that one works and thinks within, and to her own allegiance to the importance to ethics of the metaphor of vision (this is, of course, an idea that she takes from Plato). Murdoch develops the importance of this metaphor in contrast with what she takes to be an opposed constellation of images concerning the 'free agent' (which becomes, in certain hands, the 'blind leaping' of existentialism). Her foils are, at various times, and among others, Stuart Hampshire and Jean-Paul Sartre—and both of these are cited by Moran as important influences for his own work.⁵ In some sense, then, my aim here has been to offer some sort of reconciliation of these two traditions, to show how the agent must be understood as acting *in relation to*, indeed in submission to, the reality within which she acts, and the demands that this reality makes of her; that her 'leaping', such as it is, need not be blind:

In the knowable realm, the form of the good is the last thing to be seen, and it is reached only with difficulty. Once one has seen it, however, one must conclude that it is the cause of all that is correct and beautiful in anything, that it produces both light and its source in the visible realm, and that in the intelligible realm it controls and provides truth and understanding, so that anyone who is to act sensibly in private or public must see it. (Plato 1997a: 517b-c).

⁵ (Hampshire 1975), (Sartre 1956).

BIBLIOGRAPHY

- Altham, J. E. J. 1986. 'The Legacy of Emotivism.' in Graham Macdonald and Crispin Wright (eds.), *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth and Logic* (Blackwell: Oxford and New York).
- Anscombe, G. E. M. 2000. *Intention* (Harvard University Press: Cambridge, MA).
- Arpaly, Nomy. 2003. *Unprincipled Virtue* (Oxford University Press: Oxford).
- Ayer, A. J. 1957. *Language, Truth, and Logic* (Dover Publications: New York).
- Blackburn, Simon. 1984. *Spreading the Word: Groundings in the Philosophy of Language* (Clarendon Press: Oxford).
- . 1993a. *Essays in Quasi-Realism* (Oxford University Press: New York and Oxford).
- . 1993b. 'Moral Realism.' in, *Essays in Quasi-Realism* (Oxford University Press: New York, Oxford).
- . 1993c. 'Supervenience Revisited.' in, *Essays in Quasi-Realism* (Oxford University Press: New York, Oxford).
- . 1998. *Ruling Passions* (Clarendon Press: Oxford).
- Boyle, Matthew. 2009. 'Two Kinds of Self-Knowledge', *Philosophy and Phenomenological Research*, 78: 133–64.
- . 2011. 'Making up Your Mind' and the Activity of Reason', *Philosophers' Imprint*, 11: 1–24.
- Brewer, Talbot. 2002. 'The Real Problem with Internalism about Reasons', *Canadian Journal of Philosophy*, 32: 443–73.
- Callard, Agnes. 2018. *Aspiration: The Agency of Becoming* (Oxford University Press: New York).
- Christie, Agatha. 1974. 'The Lost Mine.' in, *Poirot's Early Cases* (William Collins Sons & Co Ltd: London & Glasgow).
- . 2011. *Miss Marple: The Complete Short Stories* (HarperCollins: New York).
- Churchland, Paul M. 1989. *A Neurocomputational Perspective: the nature of mind and the structure of science* (MIT Press: Cambridge MA).

- Conan Doyle, Arthur. 2001. *A Study in Scarlet* (Penguin Books: London).
- Davis, Lennard J. (ed.). 2013. *The Disability Studies Reader* (Routledge: New York).
- Doris, John M. 2015. *Talking to Our Selves: Reflection, Ignorance, and Agency* (Oxford University Press: Oxford; New York).
- Drier, James. 2004. 'Meta-Ethics and the Problem of Creeping Minimalism', *Philosophical Perspectives*, 18: 23–44.
- Dworkin, Ronald. 1996. 'Objectivity and Truth: You'd Better Believe It', *Philosophy and Public Affairs*, 25: 87–139.
- Edgley, Roy. 1969. *Reason in Theory and Practice* (Hutchinson: London).
- Eliot, George. 2000. *Middlemarch* (Random House: New York).
- Enoch, David. 2006. 'Agency, Schmagency: Why Normativity Won't Come from What Is Constitutive of Action', *Philosophical Review*, 115: 169–98.
- . 2009. 'Can there be a global, interesting, coherent constructivism about practical reason?', *Philosophical Explorations*, 12: 319–39.
- . 2011. *Taking Morality Seriously: a defense of robust realism* (Oxford University Press: Oxford).
- Evans, Gareth. 1982. *The Varieties of Reference* (Clarendon Press Oxford).
- Finkelstein, David. 1999. 'On Self-Blindness and Inner Sense', *Philosophical Topics*, 26: 105–19.
- . 2003. *Expression and the Inner* (Harvard University Press: Cambridge, MA).
- . 2012. 'From Transparency to Expressivism.' in Günter Abel and James Conant (eds.), *Rethinking Epistemology* (De Gruyter: Berlin).
- FitzPatrick, William J. 2005. 'The Practical Turn in Ethical Theory: Korsgaard's Constructivism, Realism, and the Nature of Normativity', *Ethics*, 115: 651–91.
- Frankfurt, Harry. 1988a. 'Freedom of the Will and the Concept of a Person.' in, *The Importance of What We Care About* (Cambridge University Press: Cambridge).
- . 1988b. 'Identification and Externality.' in, *The Importance of What We Care About* (Cambridge University Press: Cambridge).
- . 1988c. 'Identification and Wholeheartedness.' in, *The Importance of What We Care About* (Cambridge University Press: Cambridge, UK).

- . 1988d. 'The Importance of What We Care About.' in, *The Importance of What We Care About* (Cambridge University Press: Cambridge, UK).
- . 1999a. 'Autonomy, Necessity, and Love.' in, *Necessity, Volition, and Love* (Cambridge University Press: Cambridge).
- . 1999b. 'On Caring.' in, *Necessity, Volition, and Love* (Cambridge University Press: Cambridge).
- . 2004. *The Reasons of Love* (Princeton University Press: Princeton and Oxford).
- Gauthier, David P. 1986. *Morals by agreement* (Clarendon Press: Oxford).
- Gertler, Brie. 2011. *Self-Knowledge* (Routledge: London and New York).
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings* (Harvard University Press: Cambridge MA).
- . 2003. *Thinking How to Live* (Harvard University Press: Cambridge MA, London UK).
- Hampshire, Stuart. 1975. *Freedom of the Individual* (Princeton University Press: Princeton).
- Hare, R. M. 1952. *The Language of Morals* (Oxford University Press: Oxford).
- Harman, Gilbert. 1977. *The Nature of Morality: An Introduction to Ethics* (Oxford University Press: New York).
- Harman, Gilbert, and Judith Jarvis Thomson. 1996. *Moral Relativism and Moral Objectivity* (Blackwell: Oxford).
- Hinshelwood, Alec. 2013. 'WINNER OF THE PHILOSOPHICAL EXPLORATIONS ESSAY PRIZE 2013: The relations between agency, identification, and alienation', *Philosophical Explorations*, 16: 243-58.
- Johnston, Mark. 2001. 'The Authority of Affect', *Philosophy and Phenomenological Research*, 6: 181–214.
- Joyce, Richard. 2001. *The Myth of Morality* (Cambridge University Press: Cambridge).
- Katsafanas, Paul. 2013. *Agency and the Foundations of Ethics: Nietzschean Constitutivism* (Oxford University Press: Oxford).
- Kornblith, Hilary. 2012. *On Reflection* (Oxford University Press: Oxford).
- Korsgaard, Christine. 1996. *The Sources of Normativity* (Cambridge University Press: Cambridge, UK).
- . 1997. 'The Normativity of Instrumental Reason.' in Garrett Cullity and Berys Gaut (eds.), *Ethics and Practical Reason* (Clarendon Press: Oxford).

- . 2009. *Self-Constitution* (Oxford University Press: Oxford).
- Lear, Jonathan. 2011. *A Case for Irony* (Harvard University Press: Cambridge MA, London).
- Loeb, Don. 2003. 'Gastronomic Realism - A Cautionary Tale', *Journal of Theoretical and Philosophical Psychology*, 23: 30–49.
- Mackie, J. L. 1990. *Ethics: Inventing Right and Wrong* (Penguin Books: London).
- Moran, Richard. 2001. *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton University Press: Princeton).
- . 2003. 'Responses to O'Brien and Shoemaker', *European Journal of Philosophy*, 11: 402-19.
- Murdoch, Iris. 1977. *The Fire & The Sun: Why Plato Banished the Artists* (Oxford University Press: Oxford).
- . 1999. 'Vision and Choice in Morality.' in, *Existentialists and Mystics* (Penguin Books: New York).
- Nagel, Thomas. 1970. *The Possibility of Altruism* (Princeton University Press: Princeton).
- . 1986. *The View From Nowhere* (Oxford University Press: New York, Oxford).
- . 1997. *The Last Word* (Oxford University Press: New York, Oxford).
- O'Brien, Lucy. 2003. 'Moran on Agency and Self-Knowledge', *European Journal of Philosophy*, 11: 375–90.
- Pauer-Studer, Herlinde (ed.). 2002. *Constructions of Practical Reason: Interviews on Moral and Political Philosophy* (Stanford University Press: Palo Alto CA).
- Plato. 1997a. 'Republic.' in John M. Cooper (ed.), *Plato: Complete Works* (Hackett Publishing Company: Indianapolis/Cambridge).
- . 1997b. 'Symposium.' in John M. Cooper (ed.), *Plato: Complete Works* (Hackett Publishing Company: Indianapolis).
- Platts, Mark de Bretton. 1979. *Ways of Meaning: an introduction to a philosophy of language* (Routledge & Kegan Paul: London; Boston).
- Quine, Willard Van Orman. 1960. *Word and Object* (MIT Press: Cambridge MA).
- Quinn, Warren. 1993. 'Putting rationality in its place.' in, *Morality and Action* (Cambridge University Press: Cambridge, UK).

- Raz, Joseph. 1999. 'The Truth in Particularism.' in, *Engaging Reason* (Oxford University Press: Oxford).
- . 2005. 'The Myth of Instrumental Rationality', *Journal of Ethics and Social Philosophy*, 1: 2–28.
- Robinson, Daniel N. 2002. *Praise and Blame: Moral Realism and its Applications* (Princeton University Press: Princeton and Oxford).
- Russell, Francey. 2017. *Self-Opacity, Human Agency, and Ethics*, PhD Dissertation, University of Chicago.
- Sartre, Jean-Paul. 1956. *Being and Nothingness* (Washington Square Press: New York).
- Scanlon, T. M. 2014. *Being Realistic About Reasons* (Oxford University Press: Oxford).
- Scheffler, Samuel. 2010. 'Valuing.' in, *Equality and Tradition: Questions of Value in Moral and Political Theory* (Oxford University Press: New York).
- Schroeder, Mark. 2007. *Slaves of the Passions* (Oxford University Press: Oxford).
- Searle, John R. 1979. *Expression and Meaning* (Cambridge University Press: Cambridge).
- . 1983. *Intentionality: An essay in the philosophy of mind* (Cambridge University Press: Cambridge, UK).
- Shoemaker, Sydney. 1994. 'Self-Knowledge and Inner Sense, Lecture 2: The Broad Perceptual Model', *Philosophy and Phenomenological Research*, 54: 271-90.
- Smart, J. J. C., and Bernard Williams. 1973. *Utilitarianism For and Against* (Cambridge University Press: Cambridge).
- Smith, Michael. 1994. *The Moral Problem* (Blackwell: Oxford).
- Stevenson, Charles L. 1963. *Facts and Values* (Yale University Press: New Haven).
- Street, Sharon. 2008. 'Constructivism About Reasons.' in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics* (Clarendon Press: Oxford).
- . 2010. 'What is Constructivism in Ethics and Metaethics?', *Philosophy Compass*, 5: 363–84.
- . 2012. 'Coming to Terms with Contingency: Humean Constructivism about Practical Reason.' in James Lenman and Yonatan Shemmer (eds.), *Constructivism in Practical Philosophy* (Oxford University Press: Oxford).

- Svavarsdóttir, Sigrún. 2008. 'The Virtue of Practical Rationality', *Philosophy and Phenomenological Research*, 77: 1–33.
- Tolstoy, Leo. 1979. *What is Art?* (International Translations and Publishing Company: Detroit).
- Velleman, J. David. 2000. *The Possibility of Practical Reason* (Oxford University Press: Oxford).
- Vogler, Candace. 2001. 'Anscombe on Practical Inference.' in Elijah Millgram (ed.), *Varieties of Practical Reasoning*.
- . 2002. *Reasonably Vicious* (Harvard University Press: Cambridge MA, London).
- Watson, Gary. 1975. 'Free Agency', *Journal of Philosophy*, 72: 205–20.
- Wedgwood, Ralph. 2007. *The Nature of Normativity* (Clarendon Press: Oxford).
- Weil, Simone. 2002. 'Attention and Will.' in, *Gravity and Grace* (Routledge Classics: London and New York).
- Williams, Bernard. 1981a. 'Conflicts of values.' in, *Moral Luck* (Cambridge University Press: Cambridge).
- . 1981b. 'Persons, character and morality.' in, *Moral Luck* (Cambridge University Press: Cambridge).
- . 1985. *Ethics and the Limits of Philosophy* (Harvard University Press: Cambridge MA).
- . 1995. 'Internal Reasons and the Obscurity of Blame.' in, *Making Sense of Humanity* (Cambridge University Press: Cambridge).
- Wright, Crispin. 1992. *Truth and Objectivity* (Harvard University Press: Cambridge MA).
- Zangwill, Nick. 2008. 'Besires and the Motivation Debate', *Theoria*, 74: 50–59.