Normative Reality
heavy-duty realism and the expressivist challenge

Vladimir Vlaović
This dissertation by Vladimir Vlaović is accepted in its present form by the Department of Philosophy as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Jamie Dreier
Date: 22 April, 2013
James Dreier (Advisor)

Charles Larmore
Date: 22 April, 2013
Charles Larmore (Reader)

Nomy Arpaly
Date: 22 April, 2013
Nomy Arpaly (Reader)

Approved by the Graduate School

Date
Peter M. Weber
(Dean of the Graduate School)
Curriculum Vitae
Vladimir Vlaović

117 Hudson St.
Providence
RI 02909
U.S.A.
401-390-6527
vlaovic.philosophy@gmail.com

Education
Ph.D., Brown University (2013)
M.A., University of Manitoba (2006)
B.A. Honours, University of Manitoba (2004)

Doctoral Dissertation committee
James Dreier (adviser), Charles Larmore, and Nomy Arpaly

Doctoral Dissertation

Normative Reality
heavy-duty realism and the expressivist challenge

ABSTRACT

Practical facts about what we have reason to do are no less real and no less factual than the truths of natural science, mathematics, and our everyday observations about the nature of our surroundings. Practical facts differ from these other truths by being, as is sometimes said in philosophy, irreducibly normative. Normative facts fall on the ought-side of the is-ought divide. They are facts about what it makes sense to think or do, or what we have reason to think or do. These facts are here assumed to be fundamentally unalike any other species of fact, including those species of fact that are unquestionably naturalistic in virtue of referring to the physical universe of space, time, matter, energy, and relations of cause and effect, or else more abstract in virtue of referring to the entities of, say, mathematics. Meta-ethics, or we might call it the philosophy of normativity more inclusively, investigates, among other things, the nature of normative facts by asking whether these facts are as special and distinct as the remark I just made suggests, whether they are facts at all and if so in what sense, and other related questions. Two views in this branch of philosophy that are consistent with the sharp divide my remark insists upon are heavy-duty realism and expressivism. Here I offer a limited and modest defense of heavy-duty realism, construed as a kind of default view, against the expressivist challenge. This brand of realism is metaphysically and epistemologically costly. I do my best to make the costs tolerable, though I have no and express no delusions about doing this in a way that will satisfy critics of realism. However, defending expressivism against my objections, which come from the philosophy of language and logic, requires desperate measures—desperate enough to keep me out of the expressivist camp, realism's cost notwithstanding. I conclude that if my assumptions that heavy-duty realism is a kind of default position, and that normative facts are fundamentally unalike any other species of fact, are reasonable, then the challenge presented by expressivism is met.
Area of Specialization
   Ethics

Areas of Competence
   Metaphysics, Epistemology, Philosophy of Language, Early Modern Philosophy

Peer-reviewed Talks
   2011 (April) Ohio Philosophical Association
      “Correctness and Belief”
   2012 (April) American Philosophical Association (Pacific division)
      “Negation and Normative Belief”

External Post-Graduate Awards
   2004 Social Sciences and Humanities Research Council of Canada ($17,500Cdn.)
   2005 Manitoba Graduate Fellowship ($15,000Cdn.)
Acknowledgments

I owe many thanks to Sean Aas, Nomy Arpaly, Derek Bowman, Michael Bratman, David Christensen, Jamie Dreier, David Estlund, Allan Gibbard, Richard Heck, Chris Hill, Paul Klumpe, Charles Larmore, Bernard Reginster, Andrew Rotondo, Josh Schechter, Mark Schroeder, Tim Schroeder, Neil Sinclair, Vasilis Tsompanidis, and Michael Young.
Preface

One good way to understand this project is as a limited and modest effort on my part to
defend a thoroughgoing, non-reductive realism about practical truth, and normative truth
more generally. It makes sense to eat in moderation. There's no reason to raise one's voice
on many occasions that provoke anger. These are facts that hold true in many cases.
Being a reason, making sense, goodness, foolishness, and many other normative
properties are real, and their reality goes farther than the realms of thought and talk.
Because the normative facts of which these properties are constituents are not reducible
to the sorts of facts that comprise those aspects of reality that instantiate these properties,
nor to facts about the opinions, feelings, intuitions, plans, etc. of agents, the very
existence of these facts gives rises to interesting puzzles in metaphysics and epistemology
which are not altogether easy to solve. But, diluting the reality of these facts in a way
amenable to the expressivistic style of anti-realism invites worse problems—those
stemming from the philosophy of language. Or so I hope to convince you below.
# Table of Contents

Curriculum Vitae \hspace{1cm} iv

Acknowledgments \hspace{1cm} v

Preface \hspace{1cm} vi

Chapter 1: Some Preliminary Controversy \hspace{1cm} 1

Chapter 2: Existence \hspace{1cm} 9

Chapter 3: Epistemology \hspace{1cm} 31

Chapter 4: Supervenience \hspace{1cm} 42

Chapter 5: The Inconsistency Problem \hspace{1cm} 65

Chapter 6: Conclusion \hspace{1cm} 97

References \hspace{1cm} 100
Chapter 1: Some Preliminary Controversy

Normativity is a real feature of the world in which we live. There are better and worse ways to act and think. Spending one's money on lottery scratch tickets when one needs to buy food is a bad way to act. Believing that one won't get wet whilst walking without an umbrella, hat, raincoat, or any other kind of outerwear, is a bad way to think. The sense in which these things are bad is normative. Goodness and badness are prime examples of normative properties. These properties are every bit as real as any real thing. If there are such things as moral wrongness and moral excellence, then these things, too, are normative properties. Normativity is all about how things should be. The normative way things are—and there is, I'm going to argue here, such a thing—is identical to the way the non-normative, or naturalistically describeable, should be.

My primary aim here is to offer a limited defense of what I and other philosophers call realism about practical truths—in other words, realism about the way things normatively are with respect to practical rationality understood as a kind of subdomain of the larger or broader domain of normativity; realism, that is, about how we should act. There are facts about what it makes sense to do. There are facts about what's rational. I don't know whether there are facts about what's morally okay. But there are (again, my defense is limited), I say, facts about what to do under which conditions.

One such truth or fact (I don't distinguish the two here) is the fact, F*, that the fact, F, that it can save your life with very little cost to you is a pretty good (at least) reason to wear a seat belt whilst driving. All practical truths—truths or facts about what to do in which circumstances, what's rational to do, what it makes sense to do, and so on
—are, I suspect, facts about what we have reason to do. The domain of normative epistemology—what we might call the study of theoretical rationality, or the study of what to think given various bodies of evidence, what it's rational to think, what it makes sense to think, and so on—is a domain similar to the practical in the sense of being a domain of mind-independent, irreducibly normative facts, that is, facts about which features of reality are genuine reasons for agents like us to respond to those features in particular ways, where the responses are psychological rather than behavioral. Even though my explicit focus remains on the practical domain, what I say about this domain I also take to be true, with the necessary modifications, about the theoretical domain on analogous grounds, though the specifics are best left for another day.

The reader who expects an indubitable chain of reasoning ending with practical realism, construed as the thesis that there are irreducibly normative truths about what beings like us have reason to do, may be disappointed, for I offer no refutation of the competitor to realism that I consider and argue against in the chapters to follow (chapters 4 and 5, mostly)—a view which, these days, goes by the name of 'metanormative expressivism'. I'll explain what expressivism is, or at least do my best to give you a sense of what it suggests about normativity, below. If you've heard the term 'expressivism' and have some familiarity with the theory to which it refers, then, for now, just assume that I'll be talking about that theory. Probably the best way to understand this essay, then, is as but a modest, and limited, cost-benefit analysis—one which tells in favor of realism over the competing, anti-realist view, viz. expressivism.

Expressivism as I understand it (the characterization of expressivism, I should probably point out, evolves throughout the chapters to follow) is anti-realist in the sense
that, on a particular way of conducting metaphysical discourse, expressivists deny that there are irreducibly normative truths, instead claiming that the best we philosophers can do is systematically investigate what it is to say and think that there are such truths. For example, the state of affairs consisting of a hungry person sitting in front of his favorite dinner involves an irreducibly normative component, viz. the fact that this person's hunger and the availability of his favorite food provide him with at least some reason, though perhaps a defeasible one, to start eating this food. Let' call the person Al. The claim that Al has the above reason, expressivists would like to say, is true, though its truth consists of nothing over and above the fact that Al has a reason to start eating his dinner. The claim that his having this reason is an irreducibly normative feature of reality adds nothing, except—maybe—a semantic or, perhaps, merely pragmatic marker signaling that its speaker or thinker really means it (in the colloquial sense of 'really means it').

Let me try to clarify a little. Al has a reason to start eating because his hunger and the availability of his favorite food count in favor of his doing so. I, the practical realist, believe that this and all other instances of the relation of fact F's counting-in-favor-of agent X's performing action\(^1\) is a relation which really exists—in the same way and in the same sense of 'exist' as its relata, whereas the expressivist, in virtue of having\(^2\) a conception of

\(^1\) Like I said earlier, the abstract schematics of the counting-in-favor-of relation on which I focus are such that psychological states can take the role assigned to actions in this essay with minimal modifications to the overall schematics.

\(^2\) More informatively, I would say that the expressivist is forced into accepting, as a background theory (whether to call it a part of expressivism is something I'm not going to focus on here), a particular approach to metaphysical inquiry—a theoretical recipe I call the domain-specific approach to doing metaphysics, which I'll go on to explain, the basic idea being that the schema 'There is an X such that X is F' is a kind of shorthand for something along the lines of: 'According to the standards of domain D, there is an X such that X is F'. The implication—rather an implication I'm going to use as a recurring example—is that questions such as 'Is there an even prime number?' cannot be answered by philosophical analyses of the sort Hartry Field (1980, 1989, 2001) offers as he negatively answers the aforementioned question, but rather can only be settled using mathematical standards (number theory, more precisely). Since mathematicians and other people can straightforwardly prove that there is indeed
metaphysics that differs sharply from my own, believes that no sense can be made of this manner of speaking, other than the sense that can be made of the claim that our agent and other relevantly similar agents have the reasons they do in the circumstances they are in. There is no further question about whether the counting-in-favor-of relation, as I put it, really exists, that is, in the way that its relata really exist (whatever that is).

The conception of metaphysics working in the background of my metanormative theory suggests a kind of polysemous reading of the existential quantifier. There is a sense in which the question of whether or not Al has a reason to start eating is a question within the domain of practical reasoning. In that sense, the answer is yes. Al has a reason to start eating, because anyone in his circumstances would have such a reason. As far as practical reasoning goes, there is, on my view, nothing further to be said about why Al has this reason. But, there is an existential question within metaphysics about the counting-in-favor-of relation the presence of which makes it true that certain facts constitute reasons for Al and people in Al's circumstances, viz. 'Does this relation really exist?', to which I am going to offer up a limited defense of an affirmative answer.

Now, by contrast, the conception of metaphysics working in the background of the expressivist's view—and, interestingly, also Thomas Scanlon's (2009), despite Scanlon's outspoken self-image as a metanormative realist—implies that existential claims are fundamentally domain-specific, in the sense that the standards of correctness operative in the domain within which some existential claim is made are the soul arbiters of the truth.

---

3 Please do not read too much into the 'because'. For example, I ask that you let it be true that p because p for any substitution of p, even if this usage departs from common sense a little.
of this claim. Our example is a claim about what Al has reason to do. Because this is a claim within the practical domain, its truth depends on the standards of correctness operative within this domain. I think that the truth of this claim (assuming it’s true) adds something to metaphysics and not just ethics and practical life more generally—something, well, metaphysical.

Metaphysics, as I think of it, has a special place within intellectual inquiry. This branch of philosophy is, among many other fascinating things, home to an unrestricted notion of existence, about which I am going to say a number of things below, but not nearly enough. The metaphysician is in a position to survey other domains of discourse and intellectual inquiry, and assign, by way of the aforementioned unrestricted notion of existence, a certain status—one of two statuses, more informatively—to the entities, things, facts, or whatever, to which these domains are existentially committed. The domain of practical reasoning says (we're supposing) that a certain fact (or set of facts), F, counts in favor of a certain agent, X, performing a certain action, A, in circumstance C. The fact that Al is hungry, and has his favorite dinner in front of him, which he can freely eat, together (at least⁴) constitute a reason for him to start eating, which is to say that the facts or states of affairs denoted by the first three clauses of this sentence count in favor of Al's performing a certain action, to wit eating. Is this relation a real feature of the world, understood as the sum total of what there is, in the most unrestricted sense of 'is'? Is this relation not real, but rather a mere phantom projected onto a world devoid of value—projected by the thoughts, feelings, and interests of creatures like us? Is the normative domain like the domain of magic and sorcery? Or is it more like what we today by and

---

⁴ Perhaps each is a reason on its own. I don't know, and won't attempt to argue for claims about who has what reason in which circumstances here. It suffices that someone has a reason in some circumstance.
large think of as domains of intellectual integrity, such as the natural sciences, or mathematics (more on mathematics below)?

It's not important what we call either of the binary values metaphysics assigns to the things other domains are committed to, as long as it doesn't use the same name for both statuses. It could be real vs. not-real; exist vs. not-exist; chocolate vs. vanilla; and so on. I'm using the idioms of 'reality' and 'real existence' for the positive value, and the same thing prefixed by a negation for the negative value. On this usage, one could, in metaphysics, hold the view that tigers really exist, whereas keyboards are nothing over and above a collection of molecules arranged keyboard-wise. Common sense is committed to keyboards. But, when we enter the metaphysics room with the notion of a keyboard, and proceed to distill this notion, we may or may not be left with a concept that actually denotes something that exists in the most unrestricted sense of 'exist'. How to assign a more literal meaning to this metaphor is not something I am in a position to undertake here, though I will say that folk intuitions are in all likelihood the best place to find the materials one aims to distill. In many cases, we hope these materials will survive the distillation process. Counting-in-favor-of relations are materials that we, unless our opinions are already philosophically informed, almost universally hope will survive the distillation process. Witches and goblins are materials that most of us expect won't survive the process, though there was a time when this was not so.

Expressivism in its most sophisticated incarnations, when it is used as a catalyst for the distillation process, does not result in the destruction of irreducibly normative properties, relations, and facts. Expressivists make some good points. But these points are not good enough for it to make sense for us to start believing that the property of making
sense to which this very sentence refers is not a real property, but rather a phantom with which beings like us try to and color a reality in which, strictly speaking, it never makes sense to do something—in other words, that nothing is a good or else bad idea, or constitutes a reason, or is good or bad, or anything in that neighborhood.

Many philosophers do not want to say these things, and have devised clever ways to avoid saying them while also avoiding, or trying to avoid, the epistemological and metaphysical burdens of denying these things explicitly. Some of these philosophers happen to be expressivists, whereas others—strangely—identify themselves with members of the heavy-duty realist camp (where heavy-duty realism is the view that there are irreducibly normative facts, whose status of facts does not depend on anyone’s psychological stance toward them5), adhering to a kind of low-calorie heavy-duty realism which is designed to avoid vexing metaphysical and epistemological questions about the nature of our epistemological access to, and the nature of, irreducibly normative facts.

Scanlon is a prominent figure among these philosophers. Before getting to expressivism, then, I will focus on his view. My goal is not to refute this view, but rather to show that in the end it is difficult to construe it as a variety of realism. Instead, it is best seen as a version of quasi-realism, which is popular among today's leading expressivists. The only difference is that arguably the most sophisticated version of quasi-realism, viz. Gibbard's (forthcoming), has more things to say about the philosophy of language than Scanlon's theory. It is, moreover, unclear what alternative to those things a proponent of Scanlon's view has available to him, or would have available to him, were he to consider the surrounding issues explicitly.

5 I'll say more about the terminology in the chapter to follow.
1 Realism: anti- vs. robust

I'll frame the discussion of this chapter by starting with a bit of apparatus from the domain-neutral approach to doing metaphysics. Recall that this view implies that metaphysicians are in the business of, among other things, assigning one of two statuses to the things other domains of discourse talk about or at least try to talk about—things like microwave ovens, house cats, prime numbers, superheroes, and countless others. It doesn't matter what we call these binary statuses. We could, for example, call them real versus not-real; non-deflated versus deflated; ontological versus non-ontological; blue versus red; or chocolate versus vanilla. We could say that the things to which one of these statuses is assigned exist, or else that they don't exist. We could say just about anything, really, as long as we countenance two statuses, and assign every property, relation, individual, fact, state of affairs, event, or whatever, exactly one of these statuses.

The defining features of realism about a domain of discourse are the claims that at least parts of this domain are apt for assignment of a truth value, or that these parts express propositions; that the truth values of these propositions do not depend on any of our psychological stances toward the facts these propositions are about; and that some of these propositions are true (or have whatever truth value closest plays the role of truth in classical logic). To accept the negation of this claim (by accepting the negation of one or

---

6 Just as a matter of idiosyncratic convention, I'm going to insist that we don't use 'fundamental' and 'derivative' as our names for the two statuses under present consideration. That's because I'm going to, following Enoch, use the prefix 'robust' to describe my realism (as he describes his, which is a similar view), and 'fundamental' will name the sense in which normative facts have the "really exists" or whatever status, where my claiming that the normative facts have this status and is, in addition, fundamental, makes the view I'm defending a robust, as opposed to non-robust, realism. Cornell realisms, and so-called reductive realisms, a newer (2007) version of which is Mark Schroeder's theory, are, according to my usage non-heavy-duty realisms, for these views acknowledge the real existence of the normative facts, but take these facts to be, in the final analysis, somehow constituted by facts that a true and complete natural science could describe, at least in principle. Such scientific (in a broad sense which includes idealized sciences) descriptions are not only fanciful, according to heavy-duty realism, they are off the menu, so to speak, or, in other words, utterly hopeless.

7 Why not more? … think about this...
more of its conjuncts) is to be an anti-realist about the domain in question. To be a quasi-realist is to stake out an area of anti-realism that does not force you to deny the common sense platitudes realists tend to have at their disposal, but also to avoid taking on the metaphysical and epistemological burdens of realism proper. Finally, to be a heavy-duty realist about a domain is to be a realist who claims that the facts about which she is a realist cannot be reduced to any other kinds of facts. Thus, heavy-duty realism about normativity is the view that there are normative facts whose truth does not depend on our psychological stance toward the things these facts are about, viz. normative properties, relations, etc. (the normative facts, or some normative facts, I'll say), and that these facts cannot be reduced to non-normative facts of any kind(s) or degree of complexity. When I, as an anti-realist about certain aspects of Christian theology, such as the condemnation of male homosexuality as an immoral practice, say that certain claims this body of discourse treats as moral facts are not facts at all—that these would-be facts do not really exist—I am giving voice to an instance of an error-theoretic version of anti-realism, by saying that a certain discourse is of the right kind to be assigned a truth value, but also that the propositions it affirms are uniformly false. John Mackie (1977) defended an error theory about moral discourse. It is, however, somewhat more difficult to defend an error theory about all of normative discourse, not just its moral components. But this is not the occasion for properly addressing error theories. Rather, I'll start with Scanlon's attempt at formulating a kind of low-calorie heavy-duty realism, that is, realism with the metaphysical and epistemological import of anti-realism.

2 Scanlon's low-calorie heavy-duty realism
I'll begin with some clarification about the nature of the irreducibly normative truths
Scanlon thinks he is a realist about. In the Locke Lectures, he invokes a distinction between “pure” and “mixed” normative claims. Pure normative claims have only normative implications, whereas mixed claims have normative and non-normative implications. An example of a mixed claim is this. The fact that it is raining outside provides Tom with a reason to take an umbrella in the circumstance of preferring to stay dry. Call this claim Mixed. This claim implies that it is raining, and that Tom prefers to stay dry, so it has non-normative implications. If it weren't raining, or if Tom didn't prefer to stay dry, he wouldn't have the reason expressed by this claim. Pure normative claims, on the other hand, are expressed in the subjunctive mood, and their truth or falsity does not depend on non-normative matters. Here is an example of such a claim: if it were raining, and if so-and-so were in the circumstance of preferring to stay dry, (s)he would have a reason to take an umbrella. Call this claim Pure. The truth of Pure, assuming it is true, does not depend on its actually raining or on any individual’s actually preferring to stay dry. If Pure is true, it is not true in virtue of anything naturalistically describable, says Scanlon:

As a first step, I will take the thesis of the autonomy of the normative from the naturalistic to be the thesis that no pure normative claim is entailed by any combination of claims about physical and psychological facts. Given any combination of non-normative claims, it is a further claim that some pure normative sentence is true. (2009, Lecture 2: ms. 27)

The claims that it is raining, and that Tom is in the circumstance of preferring to stay dry, do not jointly imply the claim that he has a reason to take an umbrella, or that it would be good for him to take an umbrella. To bridge the logical gap between the former two claims, which natural science is well equipped to clarify, comment on, etc., and either of

9 Sometimes I'll focus on reasons, other times on goodness, …
the latter claims, which attribute some normative properties to some naturalistically describable events or states of affairs, we need, according to Scanlon, something like Pure.

Do pure normative claims have metaphysical implications? If 'metaphysical' refers to what there is in the most fundamental sense of 'is', then yes, because truths expressed in the subjunctive mood of English are still truths, and Pure is such a truth. Since it is something, viz. a truth, it exists. But if the aforementioned sense of 'is' turns out to be domain-specific in the sense envisaged by the domain-specific approach to doing—rather, doing away with—metaphysics, then the answer is no. Under the latter supposition, neither pure normative claims, or anything else, have metaphysical implications. The claim that there is a cat in the room does not have metaphysical implications, but rather physical, chemical, biological, and (I would suggest—even insist) sublime implications, among others, but each of these sets of implications are within some particular domain of discourse, and each of these domains has its own standards of correctness that are the final arbiters for which truth value a claim falling under those standards has. There is no such thing as existing full stop, no domain-neutral, fundamental sense of 'is'. There is only existing-according-to-standard-S, where S governs the assignment of truth values for its respective domain of discourse.

In his own words, Scanlon describes the domain-specific view as the:

permissive first-order view, according to which we should decide what sentences to accept by applying the criteria appropriate to the relevant first-order disciplines—empirical science, mathematics, and so on perhaps including … our best thinking about reasons—and simply accept the set of ontological commitments that these sentences have (2009, Lecture 2: ms. 3)

There are normative standards for accepting the truth of a claim such as Mixed. But
Mixed is not a claim belonging only to the first-order domain of moral and practical reasoning. It is also a claim about the weather, and a claim about a person. So we should accept some meteorological commitments, and some psychological commitments as well, if we accept Mixed. As far as the normative commitments involved in accepting Mixed go, this claim seems to ontologically commit us to a counting-in-favor-of relation, if having a reason is featuring among the relata of one of these relations. So Mixed commits us to the existence of a reason-relation, which Scanlon accepts.

Of course, he accepts it in domain-specific fashion:

The claims we make about moral right and wrong generally presuppose that there are moral standards that everyone has good reason to take seriously as guides to conduct and as standards for objecting to what others do. But the ordinary ways of understanding morality, and ordinary ways of arguing for moral conclusions, do not make clear what these reasons are. There is therefore a question, external to morality, whether there are such reasons, and whether the usual ways of establishing that a form of conduct is wrong also guarantee that there are good reasons not to engage in it. This question is not scientific or metaphysical (unless metaphysics is understood to include normative claims), but rather normative, hence “practical” in one sense of the term—a question about what we have reason to do. (2009, Lecture 2: ms. 10)

So the question of whether there are true claims referring to irreducibly normative facts is not a metaphysical question “unless metaphysics is understood to include normative claims”, where by ‘normative’ Scanlon presumably means irreducibly normative.

Questions about the existence of reason-relations are settled exclusively on practical grounds. These just are questions about whether people have reasons. The are not questions about whether the relational property of having a reason is on some grand list of what there is in the most unrestricted sense of ‘is’, because there are multiple such lists, one for every domain of discourse.

You might suspect that another sense in which the existence of irreducibly
normative truths—pure ones, that is—does not, on Scanlon's view, have metaphysical implications, is that these truths don't conflict with claims made within other domains, nor do they even have implications in these other domains. To say that an irreducibly normative truth exists, on this view, is to say that someone either does or could have a reason to do or to think something. That's all. To say that a reason-relation makes an ontological contribution, or that it is real, or that it is true that there are such relations, are ways of saying the same thing. Because irreducibly normative facts does not causally interact with anything, for if it did, it would be capable of standing in relations that are not normative but rather something else, in this case causal, this facts never encroaches on territory that doesn't belong to it. This may seem difficult to disagree with, but what would, or at least could, lead to disagreement is if it turned out that our accepting some true normative claim was caused by the irreducibly normative referent(s) of this claim. I will return to one facet of this issue, viz. the way in which the domain-specific theorist handles inter-domain conflict, in the next section, and another, viz. the causal efficacy of irreducibly normative facts, in the next chapter within the context of the epistemology of the normative.

2.1 The regress problem introduced
Some metaphysicians have used the phrase 'hard to stomach' as a predicate for some class of brute facts posited by their colleagues. The name or description of any one of the irreducibly normative truths Scanlon posits may look like a prime candidate for being the subject of this predicate. In this subsection, I'm going to flesh out the, or an, objection to that effect, inspired by a paper from Tristram McPherson (2011). In the following

---

10 This presupposes that causal truths are not normative truths, which is something I don't take a stand on here, one way or another.
subsection, I'll answer the objection on Scanlon's behalf, and then, in the remainder of the chapter, make my own attempt to cut to the heart of the issue, which is not the question of whether Scanlon's is a bad view, but rather the question of what kind of view it really is.

To get the regress to which the title of this subsection alludes up and running, I'm going to return to the claim I called Mixed, and I'll abbreviate its name further, to just 'M'. M says that if it were raining and Tom preferred to stay dry, he'd have a reason to take his umbrella. If M is true, it is not true in virtue of any natural fact(s) or natural facts, according to Scanlon. In virtue of what, then, is it true? Scanlon thinks this is a legitimate question, since he himself poses it in the Locke Lectures, calling it the “Ground of Truth” question (though I call it the grounding question):

In virtue of what are claims about reasons true, when they are true? Does the idea that claims about reasons are true or false, independent of our opinions about them, and that truths about reasons are irreducibly normative, have unacceptable metaphysical implications? (2009, Lecture 1: ms. 4)

While he says that “It would be desirable to have an account of what makes claims about reasons true, when they are true”, “such an account, if it could be given, would be itself a normative thesis, not a metaphysical one” (Scanlon 2009, Lecture 1: ms. 5). From this we are, I take it, to think that there exist norms or standards for correctly deciding whether a particular pure normative claim is true, and that these standards, rather than some property or relation external to the normative domain, are what's relevant to the acceptability of any given pure normative claim. The grounding question is a question about what there is and what it's like to the extent that this is a matter of what it makes sense for creatures like us to do and to think.

Now, it is pretty clear that this leads to a kind of normative regress. The question is whether this regress is vicious. First the regress. According to Scanlon, M, if it is true
and if there is a non-trivial answer to the question of its ground of truth, then M is true in virtue of the fact that there are substantive, normative standards of correctness for arriving at M which we have no reason to regard as defective (pretty close to his exact words). What might a standard of this sort look like? It could be the further claim that if one were in any circumstance whatsoever, one would have a reason to satisfy one's preferences. Call this claim MG (for M's ground). Is it intelligible to ask in virtue of what MG is true? The question seems intelligible, but it may not have a good answer. This is because the answer to the grounding question posed in relation to MG, if a non-trivial answer could be given, would presumably consist of the fact that there are further standards of correctness for arriving at MG which we have no reason to regard as defective. Suppose for the moment that there are such standards. Call the set of these standards S. What is contained in S? Presumably the contents of S are further normative truths, say T₁, … , Tₙ. After all, what else could S possibly contain besides further normative truths? MG is a pure normative claim, after all, so it seems incoherent on Scanlon's view to suppose that the standards of correctness for arriving at MG involve non-normative truths. But notice that we have no reason to exempt each Tᵢ-truth from an intelligible grounding question about it. According to Scanlon, though, these questions have substantive normative answers appealing to standards of correctness, which we have no reason to regard as defective. But these standards can be expressed by further normative claims, about which we may, in turn, ask still further grounding questions. If the grounding questions posed in relation to each Tᵢ-truth themselves have normative answers, which appeal to substantive substantive standards of correctness expressed by further normative truths in relation to which we may, in turn, pose still further grounding
questions, then we are indeed faced with a normative grounding regress.

A moment ago I said that the real question is whether the regress is vicious. A related question is: what is the structure of the regress? Does the regress terminate in a brute normative fact? If so, then this fact grounds itself. Does the regress instead go on ad infinitum? If so, then we can never know the whole story, but rather only a tiny portion, of why Tom has a reason to take his umbrella on a rainy day in his more or less hydrophobic frame of mind. Or perhaps the grounding question-answer pairs are eating their own tail, thus forming a circle, in which case normative reality grounds itself in a somewhat more complicated way than that envisaged by the proponent of brute normative facts. No one of these three options seems appealing. This trilemma is, then, what I'm calling the regress problem. Interestingly, the solution I favor won't take us much farther than a kind of restatement of one of the essential components of heavy-duty realism itself.

2.2 The regress problem solved
The threat of viciousness with regard to this regress is, according to the solution I favor, more apparent than real, because the question of in virtue of what MG is true may not have, or need, a non-trivial answer. In other words, S is empty—there aren't any T₁-truths. There is just MG (if that), and it (or maybe M) just is true, and that's all there is to it. heavy-duty realism—part (not necessarily proper part) of its essence—is the claim that there are irreducibly normative, brute facts (or truths).

2.3 MacPherson's objection: why it doesn't work
Doesn't appealing to brute normative facts invite the question of why they, rather than
some other putative normative facts, are true, and have genuine normative implications, while these other putative normative facts are false, and have merely apparent normative implications?\textsuperscript{11} McPherson (2011) thinks there is a deep problem with this. But the problem is not what he thinks it is. Here's what he thinks (or thought at one point—I'm not sure which is more accurate).

If MG is true as a matter of brute normative fact, then a standard of correctness expressed by a another claim, MG*, isn't, for MG* says that regardless of what other circumstances one might be in, one has a reason to refrain from doing what would satisfy one's preferences. MG* grounds claim M*, which says that if it were raining and Tom preferred to stay dry, he'd have a reason to leave his umbrella at home.

Or you can imagine that M** (and likewise for MG**, with the necessary modifications) says that if it were raining and Tom preferred to stay dry, he'd have, as MacPherson puts it, a schmeason to leave his umbrella at home. Why, then, are M and MG true, but M* and M1* false? Why isn't Tom obliged to act in accordance with the standards grounding M** and MG**, but is obliged to act in accordance with the standards grounding M and MG, which, we are assuming in light of the result of the previous subsection, just is either, or both, of the latter claims?

MacPherson tries out a couple of unhelpful answers. First, he points out, we can conceive of a whimsical demon who decides he prefers M* and MG* to M and MG.\textsuperscript{12} Then we can try to say that the true or real or obligatory or whatever standard of correctness is the one that is ranked below the alternatives on the demon's preference scale. But being ranked below the alternatives on a demon's preference scale is too

\textsuperscript{11} Compare McPherson (2011).

\textsuperscript{12} I'm going to henceforth stay away from MacPherson's schmeasons-talk since I think it only obscures the central issues in a way that will become apparent, if it isn't already, in the later portions of this chapter.
arbitrary a feature to ground a claim of genuine, rather than merely apparent, normativity.

MacPherson's error is in supposing that the demarcating feature can't be just that acting according to the standard of correctness expressed by MG* constitutes making a mistake according to the standard of correctness expressed by MG. He supposes that it can't because it is also true, MacPherson confusedly suggests, that acting according to the standard of correctness expressed by MG constitutes making a mistake according to the standard of correctness expressed by MG*. We need an account, he thinks, of why one of these mistakes has genuine normative implications while the other does not.

The confusion is that acting in accordance with the standards grounding M and MG, viz. the claims themselves, is not a mistake according to the standards grounding M* and MG*. Rather, it is a mistake*! The notion of making a mistake about what to do is itself normative—irreducibly so on the heavy-duty realist's view. According to Scanlon, normativity is all about reasons for thought or action. So it's just not true that Tom has a reason to leave his umbrella at home. This is why I warned you that the response to MacPherson's pseudo-problem doesn't take us much farther than a restatement of heavy-duty realism. And if we want to, with MacPherson, talk about schmeasons, then fine, Tom has a schmeason to leave his umbrella at home. But he has a reason—a reason—to take the umbrella with him. To say that Tom makes a mistake according to the standards governing schmeasons if he acts on his reasons is nonsense. The closest thing is the fact that Tom does something out of sync with his schmeasons, but this doesn't matter—it's not a practical error—because mattering is a normative property, and normative properties are all about reasons, not schmeasons. Practical mistakes, according to any heavy-duty realist who thinks normativity is all about reasons,
are actions that the people who make mistakes have reasons to avoid. That's just what a mistake is: something you have conclusive reason to avoid. It's not something you have schmeason to avoid. And that's all there is to it.\footnote{Some readers may end up predicting where the rest of this chapter's discussion is going, viz. toward the nature of the relation between having sufficient or perhaps decisive reason to believe P on the one hand, and the truth of P on the other.}

Instead of seeing this, MacPherson makes the rhetorical suggestion that we care about the standard expressed by MG whereas we are unmoved by the standard expressed by MG*, and then goes on to dismiss this is an implausible would-be explanation of why MG is true whereas MG* is not. It is indeed implausible from a realist's point of view, because saying something like it runs counter to the basic motivating force behind realism, as MacPherson is aware. This is because saying it amounts to suggesting that normative truths such as MG are true in virtue of the fact that we care about them, which, in turn, suggests that such truths are not independent of our psychological stances toward the facts they are about, or to these truths themselves. To say that MG is true and MG* false in virtue of the fact that we happen to be moved by MG and left cold by MG* is therefore straightforwardly anti-realist, at least according to the usage of 'realism' I've stipulated here.

However, MacPherson's toy suggestion brushes up against the central problem with Scanlon's low-calorie heavy-duty realism, viz. the fact that this view really is, in the end, anti-realist.

\section*{3 Metaphysics: two perspectives on how to do it}
Recall from the previous chapter that the domain-neutral theorist sees metaphysics as
being, among other things, in the business of assigning one of two statuses to the facts
other domains—that is, as Scanlon would call them, first-order domains, including
physics, mathematics, practical reasoning, and many others—are existentially committed
to. So, for example, either photons, which are referred to by many of the truths or at least
what we take to be the truths within the discourse of physics, really exist, as this
discourse implies, or else they don't really exist but rather are best (or better) viewed as a
kind of useful fiction.

Thus, the domain-neutral view sees metaphysics as a truly “meta” enterprise, a
central task of which is to argue in a systematic way whether or not the existential claims
of first-order, that is, non-meta, domains, such as physics, psychology, art, morality, and
practical reasoning more generally, to name but a few, are true—if the facts they talk
about really exists. Assuming the domain-neutral view for the moment, it is natural to
think that if our best physics entails that there are photons, then our best metaphysics
should entail whether or not photons really exist. The claim that photons really exist,
then, is a claim within metaphysics, about the facts existentially quantified over by the
truths, or at least what we take to be the truths, of physics. Similarly, the claim that the
normative facts (or some normative facts, more cautiously) really exists is a metaphysical
claim, one which is an essential component of what I've been calling metanormative
realism. The claim that some of the normative facts that really exists is fundamental, that
is, not reducible to any naturalistically describable facts, or, indeed, any non-normative
facts period, is an essential component of what I've been calling heavy-duty realism.

There is, however, no room for these or any other metaphysical components
within the study of normativity, according to the domain-specific view. Existential claims
are not metaphysical claims, it turns out—not according to domain-specific theorists. The claim that photons really exist is a philosopher's way of saying something that properly belongs within the body of discourse properly belonging to physics. It is not a claim about some subdivision of our best physics-talk, but a funny way of giving voice to something with this talk. Similarly, the claim that some normative facts exists is a normative claim about normative facts, viz. it makes sense to think that there is facts that it makes sense to do or to think.

### 4 The mumbo-jumbo problem

#### 4.1 Heck's proliferation problem

On its face, the domain-specific view invites Richard Heck's (2000) proliferation problem. The problem as I'll construe it here—which is not exactly the way Heck construes it but will do for our purposes—for domain-specific theorists is that these philosophers seem to be under no restriction to refrain from (over)populating the world with as many entities as they like, as long these entities are referents of some domain(s) of discourse with its own standards of correctness. But, even with this restriction, the standards of correctness such a philosopher can appeal to may very well be just as much creatures of his own stipulation as are the entities (or, as I've been saying, the facts\(^{14}\)) referred to by the claims that come out correct according to these standards.

For example, let's say it's frood to blink twice upon seeing a cat and under no other circumstances—and that anyone in those circumstances has a freason to blink twice. Why has no one explicitly acted on the basis of his or her beliefs about froodness or freasons? More to the point, why do so few (oh, so very few) people even have beliefs about these properties? Let me add that froodness is one of a number of lormative

\(^{14}\) In case 'entities' doesn't pick out things like relations.
properties, and also that there are, in addition to these properties, lormative relations, facts, and the like—in short, that there is lormative facts, and that this facts is irreducibly lormative. According to my usage, I am now a heavy-duty realist about lormative facts.

Okay, no I'm not. But I'll pretend that Trisa is. In fact, I happen to be an error theorist about lormative facts, for I think this fact consists entirely of things Trisa made up one day. But Trisa is very imaginative—overly so, in fact. Her powers of imagination sometimes come in handy, as when she writes successful works of fiction, but the downside to her having these powers is that she sometimes just goes and makes up a bunch of facts that isn't real, and starts believing that it is real. As a domain-neutral theorist, I get to say this in a fairly straightforward manner. But the domain-specific theorist is going to run into a problem. For Trisa has not only told us what lormative facts there is, she has, in the process of doing this, given a detailed outline of a rich set of standards of correctness for lormative claims. If I see a dog and blink twice in response, and then tell you that I did this because it was the frood thing to do in those circumstances, thereby giving voice to my belief that blinking twice upon seeing a dog is frood, then, should Trisa happen to overhear our conversation, she may point out that this belief is false according to the standards of correctness governing lormative discourse.

Jeff, a fourth party, may, however, express puzzlement, and ask Trisa whether lormative discourse is an artifact of one of her ingenious fictional works. Trisa will tell Jeff that the answer is no, for there is nothing fictional about lormative facts or the parts of lormative discourse referring to it. Jeff asks: “How come I never heard of this facts, then?” Trisa retorts (somewhat impatiently, with a bit of sarcasm): “What, Jeff, you've heard of everything, have you?” Jeff points out that he wasn't intimating having heard of
everything there is to hear about, but that (almost) nobody has heard of any formative facts existing out there, as it were, in the real, that is, non-fictional, world.

If it is acceptable for Trisa to conjure formative facts into existence, Jeff wonders, then why can't he posit a goblin war raging among non-physical and non-natural but nevertheless very real combatants, a war which is responsible for all the negative emotions in the world.

Trisa has an answer, to wit she tells Jeff that his goblin war, if it were real, would involve causal interaction with the naturalistically describable world. Because there is no scientific evidence of this, Trisa argues, we ought not to accept the reality of such a war. The irreducibly formative facts, however, is causally inert, much like the irreducibly normative facts, which, too, is causally inert according to many defenders of the reality of the latter.

The principle working in the background of Trisa's response to Jeff is something Scanlon draws upon in his Locke Lectures, as we saw above. In response to an anticipated objection inspired by Heck's proliferation problem, a version of which I've been sketching via the voice of Jeff, Scanlon may well want to say that there is nothing overly worrisome about populating the ontology with whatever one cares to stipulate into existence, as long as there is no resulting conflict with plausible claims made within other domains—especially, I think but am not sure, he wants to say, domains as reputable or useful as the natural sciences. Since Jeff's goblin problem asks for a principled reason to disavow the existence of the above mentioned goblin war, it can be solved by making the observation that the purported reality of such a war contradicts plausible claims made within various natural sciences. Scanlon's own case is that of the now-mostly\textsuperscript{15}-archaic

\textsuperscript{15} People still call one another witches, and the like, but as far as I am aware, no widely recognized
(thankfully) discourse, which was once taken seriously (to say the least), of witchcraft. Since witch-claims are in part causal claims that contradict physics, chemistry, and other natural sciences, it won't do to stipulate the existence of witches.

This response leaves more questions in its wake than it answers. I'll discuss only a few of these questions—the ones that strike me as especially pressing.

Assuming, then, that some domains are sacred\(^\text{16}\) and that the natural sciences are among the sacred domains, and also assuming that the best\(^\text{17}\) natural sciences have the last word on causal claims (which is an assumption I'll revisit in the next chapter), there is a question about how domain-specific theorists are supposed to adjudicate inter-domain conflicts among various purportedly non-causal domains. Suppose Trisa, on top of stipulating into existence lormative facts, stipulates into existence some zormative facts as well. One zormative property is goodness, and, it turns out, it's good to avoid at all costs blinking twice upon seeing a cat. I see a cat, and I'm not sure what to do: the frood thing, which is blink twice, or else the good thing, which is anything but blink twice.

Ah, but now, even if Trisa doesn't have an answer, Scanlon does. We've already seen it. It's the answer I suggested on Scanlon's behalf when arguing against MacPherson above: in the envisaged cat-viewing scenario, I should do the thing I have most reason to do, or, if you like, the good or best thing to do. It doesn't really matter whether an action is frood or grood. What does matter is whether the action in question is good, which, on Scanlon's view, is determined by whether or not I, or anyone else who happens to find

\(^{16}\)The third question, I'll reveal, is what could make a domain sacred in the relevant sense (and what that sense is to begin with).

\(^{17}\)The question of whether, say, the best physics must be defined in terms of truth (the best physics is the true completed physics, one might think), where the operative notion of truth is, as Dummett says, evidence-transcendent in the sense that the best evidence could turn out to be misleading, is a question I'll return to shortly. Unfortunately, it is not one I can do justice to in the present discussion.
him or herself in my circumstances, has reason to prefer it over the alternatives (or something like that). After all, goodness and having a reason are normative things. These things, not the lormative, or the zormative, facts, are guides to action. To be a heavy-duty realist about normative facts just is to dig one's heels in and say that part (maybe improper) of the essence of all normative things is the propensity to guide the thoughts and actions of sophisticated enough agents—that is, agents like us.

(The related topic of what makes it possible for us to, on occasion, be motivated by our moral beliefs (assuming we have some, and that the ones we have occasionally motivate our intentional actions) is treated in like fashion by the heavy-duty realist. Sometimes—maybe every time conditions are ideal in some suitable sense—moral beliefs motivate precisely because they are about normative facts. The very existence of this facts as conceived by heavy-duty realists (in other words, the existence of this facts plus its irreducible character) makes moral motivation possible, to the extent that the moral facts is a kind of normative facts. Unfortunately moral motivation and its conditions of possibility are topics I must leave for another day.)

Whether to act in accord with the lormative, or else zormative, analogs to having a normative reason, the heavy-duty realist's answer is simply that the nature of the thing to do depends on which option—the one in line with the lormative lay of the land, or else the one in line with the zormative lay of the land—is the option we have most normative reason to take. What to do is a normative question. Neither lormative nor zormative questions are questions about what to do, for, if they were, they would be normative.

---

18 Psychological, not just environpsychological, circumstances count.
19 Something to do with reasons, in other words, for the relation of F's being a reason for any person, x, in circumstance C to perform action A (or take up mode of thought T) is, among the normative properties, fundamental in the sense that all the other normative properties are reducible to, or derivative from, it.
20 Interestingly, when this is conjoined with the sharp, Hume-inspired, division of belief and desire, it implies, or comes close to implying, non-cognitivism about moral belief/judgment.
That, in a slogan, just is heavy-duty realism (or part of its essence at any rate), not that any of this is news given some of the preceding material of this chapter.

It is, however, contrary to the spirit of any kind of realism about the normative facts to suppose that our role in relation to this facts is more like that of creators than explorers. Realists generally don't identify with views that imply, on any level of explanation, that we have the reasons we do in fact have because we believe we have them. Rather, when things are going aright, we believe we have certain reasons because we do in fact have them, at least according to metanormative realism, though the same things can be said, with the necessary modifications, of realism about the material world. Opponents of Bishop Berkeley do not in general suppose that the material world exists in pretty much the way they envisage because they envisage it in this way—because, in other words, they disagree with Berkeley. Rather, they, or we, think that our holding beliefs about the material world that cannot be squared with the good Bishop's is a result of the fact that there really is a material world out there in much the way that we, but not he, think this world exists. Why, then, do realists about the discourse of natural science get to be right-minded error theorists about the goblin war, as it were, free of charge?

Answering this question is of course a general problem for the domain-specific view, rather than a problem specific to the conjunction of this view and heavy-duty realism about normative facts. Indeed, when conjoined with heavy-duty realism about the norms governing rational belief, the domain-specific view seems to almost solve, or come close to solving, the problem. What should we believe, the claims made by goblin war realists, or else the claims made by realists about the bodies of discourse belonging to the natural sciences? Exactly the claims we have most reason to believe, for the reasons we
have to believe those claims, is an answer open to heavy-duty realists about norms of evidence.

Assuming the balance of reasons weighed in favor of the body of claims made by natural scientists (against goblin war theorists), we would have a solution to a—but not the—problem. The solution is not a solution to the right problem, in other words. What we ought to believe, natural science or else goblin war theory, is one question. What exists, the facts of natural science or else the facts of goblin war theory, is another question. Appealing to the reality and fundamentality of norms of evidence can, given some auxiliary premises, get the domain-specific realist about the facts of natural science, one who also wants to be an error theorist about the facts of goblin war theory, just about all the way home to a solution to an epistemological version of the proliferation-goblin war problem. But, as I presented my version, this is a problem about the facts of existence, not about what to believe in relation to those facts.

4.2 Victory by defection?
Before moving on, I’ll close this section by pointing out that there is a way for domain-specific theorists to get around the worries I’ve been sketching. It is, however, a defection into the anti-realist camp—more specifically, a broader kind of anti-realism as conceived by Dummett. Simply put, Dummettian anti-realism is the denial of an evidence-transcendent notion of truth. If the truth about what there is cannot outrun the limits of good, or the best, evidence for what there is, then, because the best evidence favors the existential import of scientific claims over that of goblin war-theoretical claims, there is no further question about what facts—science-facts or else goblin-war-facts—exists. I’m not about to say that the resulting theory—that is, the theory that results from disavowing
the existence of goblin-facts, and accepting the existence of science-facts as well as at least some, viz. evidential, irreducibly normative facts—is bad. But heavy-duty realism about normative facts and a broader Dummettian anti-realism seem to me to be strange bedfellows. I for one would rather take on the metaphysical and epistemological burdens of original heavy-duty realism than Scanlon's low-calorie alternative and the commitments it demands.

Having made peace with the caloric content of full-blooded heavy-duty realism, the next two chapters are aspects of the exercise program intended to burn some of those calories off. A little less metaphorically: the next chapter defends heavy-duty realism against epistemological objections. For example, heavy-duty realists seem to need to explain how it is that our minds, which are constituted or realized by neurological and thus physical facts, and maybe some irreducibly psychological facts, depending on the nature of the relationship between mind and brain, manage to correctly represent irreducibly normative facts. Whenever a domain-neutral metaphysician (henceforth metaphysician) posits two or more kinds of facts, and wants to say that these kinds of facts are related in systematic—indeed logical or conceptual, thus necessary—ways, he or she (who is in this case me) seems to need to explain how it is that these, as Hume and his followers might put it, distinct existences stand in such relations to one another. I'm going to try to take some of the bite off of any objections leveled against heavy-duty realism that follow this pattern of thought. This, then, is the business of the next chapter, to which I now turn.
Chapter 3: Epistemology

1 Preliminary remark
Normative properties, relations, and facts are, according to heavy-duty realism, not like
any other sorts of properties, relations, and facts. (I'll mostly speak of facts, but if at times it makes more sense to you to substitute in the language of properties, please do so.) Since psychological and physical facts count as other kinds of facts in this context, it is a mystery, at least on its face, how these facts manage to interact in systematic and predictable ways with normative facts. How does Paul, a creature of flesh and blood, and perhaps a soul or some soul-like qualities, manage to know or at least rationally believe that his having the plan or intention to write a really good book makes it a good idea for him to also form a plan or intention to do some research for this book? The question is pressing for heavy-duty realists, because we think the value of Paul's putative intention to do some research given that he has set out to write a good book is not a feature of this intention that can, even in principle, be assimilated to the realm of the natural world; or numbers or other abstracta; or souls; or whatever. One answer I will do my best to sell on behalf of what I'll conjecture is a rather large majority within the heavy-duty realist camp, though I for one don't count myself among this majority, is that Paul and the rest of us who have some rational beliefs about what to do are very lucky to have come by these beliefs. I don't think these beliefs are products of dumb luck. Rather, I think they are the products of intelligent design, but this is not something I'm prepared to defend in the present essay.21

I'll start with Gil Harman's infamous case:22 you round the corner and see kids

---

21 A few years ago, I described heavy-duty realism to a highly intelligent, now-former member of an English department. My friend got the basics on the first pass, and his first and most enthusiastic question was whether the view I described to him springs forth from what he called a religious impulse. Another, equally intelligent person, but this time a professional philosopher working outside of metaethics, on another occasion voiced a concern—and I do mean concern, for this person finds the notion of intelligent design not only metaphysically and epistemologically, but morally, repugnant—that heavy-duty realism treads in an area she takes to be dangerously close to non-secularism. I agree with my two friends to the extent that I do think that heavy-duty realism works best when conceived as a non-secular view.

22 Ref
pouring gasoline on a cat, intending to light her on fire. You don't need to reflect on the moral lay of the land, so to speak. You can just see that what's about to happen is wrong—to say the least! (If you're in my shoes, moreover, you immediately do something to cause the kids to scatter, and tend to the poor cat in as loving a way as possible.) This quasi-perceptual model of moral knowledge, or justified belief, probably won't work if taken too literally, however. At least I'm not comfortable taking it literally under the supposition of heavy-duty realism (a supposition which I'll retain for the rest of the chapter). This is because the wrongness of the action the kids are about to undertake if you or I don't stop them, which, I'm going to assume for the sake of my own psychological comfort, is an action that I, you, or whoever the protagonist of the story happens to be, prevent, is not a property capable of physically impacting the human retina. It may be that secular heavy-duty realists have to or ought to, in the end, deny this, which, if done, would entail that non-physical facts can and does causally interact with physical facts. Because a proper discussion of this issue would take us too far away from the metaethical questions I'm trying to answer, or at least sharpen, in this discussion, I set it aside for another day. Henceforth, I'll assume that all causal relations are physical relations.

Given this assumption, in exploring the prospects of a quasi-perceptual model of normative knowledge, or justified belief, it seems that we have to take Harman's example of just seeing that a particular course of action is wrong more or less metaphorically. I propose to give literal sense to the contents of the metaphor by employing the notion of inference. If normative knowledge is a matter of seeing the normative facts heavy-duty realism says is out there in the world, then episodes of seeing this facts are not, strictly
speaking, results of a perceptual faculty, though they are, strictly speaking, the products of a type (or types) of non-inferential psychological process. I'll call this process intuiting, and name its products, which are psychological representations of how things are (how things are normatively in the case of normative intuition, how things are naturally in the case of perception proper, how things are mathematically in the case of mathematical intuition, and so on, if there are any others) intuitions. Normative intuitions, then, are what thinkers like us end up with when we see that some spatial and temporal area or, if you like, this area's occupants, are are right, wrong, good, bad, permissible, required, or what have you. According to this non-inferential model, Lisa can intuit her way to certain beliefs about what to do, such as the belief that it's a bad idea to embark on her usual inter-state commute when there is an emergency weather warning in effect, from which she will go on to infer further normative beliefs, such as the belief that it's an even worse idea to undertake the commute and exceed the posted speed limit on the way there and on the way back. If there are innate normative beliefs that people like you, Lisa, and I have, say the belief that sugar and fat are good, then these these beliefs qualify as intuitions, too. And so do any beliefs that are not the products of inferential processes that start with other beliefs.

2 Basic psychological representation
Suppose normative facts are pretty much the way heavy-duty realism portrays them to be. To wit, these facts are fundamentally unalike any other kind of fact, including psychological facts and physical facts, whatever exactly the relation between the latter two kinds of facts turns out to be. Wouldn't there, then, have to be some kind of coincidence between the structure of the normative facts that are out there in the world,
waiting to be responded to by beings like us on the one hand, and the psychological representations of these facts that exist in our heads on the other? One component of realism, robust or otherwise, is the claim that we have a bunch of true beliefs about what to do. Strictly speaking, it could, as far as heavy-duty realism is concerned, turn out that none of these beliefs are rational or justified, or that we have no normative knowledge. This is consistent with heavy-duty realism, as long as some of the beliefs in question are true. But given the statistical regularity with which we end up with true or correct normative beliefs, it is difficult to plausibly deny that we are fairly reliable when it comes to thinking about what to do (remember Paul's case, and think about how many times you've been in a similar situation, and came up with a reasonable belief about which means to take in order to achieve your given end or ends). Indeed, for an agent to have reliable beliefs about what to do just is for her to have beliefs about what to do that tend to be true. Thus, if all we know about this agent is that she's pretty much like you and me—epistemically—then we know there's a fairly decent chance that her next belief about what to do is going to turn out to be true, provided that the conditions in which this belief is formed are epistemically normal or better than normal.

Maybe reliability so understood is neither necessary or sufficient for rationality or justification. Still, the fact that we are reliable thinkers (for the most part, or at least to a noteworthy extent) when it comes to practical matters, is interesting enough in its own right. It's interesting enough to get the coincidence problem up and running. Simply put, we'd like to know how it is that beings made of one or more kind(s) of facts—psychological properties, physical properties, matter, energy, etc.—that is or are utterly unalike normative properties, relations, etc. manage to correctly represent to ourselves the
There is a family of theories acknowledging a logical connection between our causal interactions with the material world, and our having, on occasion, justified beliefs or items of knowledge about this world. Medium sized dry goods impact our sensory organs in scientifically predictable ways. These impacts, via various biological mechanisms, result in beliefs about the natural world. And, some people go on to say, our propensity toward reliably true beliefs about the natural world is, in some way or other, logically tied to our having justified or rational beliefs in this area of inquiry. Assuming that this causation-based model is roughly on the right track, it becomes an alarming piece of news that this model cannot be applied to a causally inert normative reality.

### 3 Further details

I may have mentioned the concept of a perceptual, or sensory, intuition. Such intuitions I think of as non-inferential representations that, with the aid of other psychological faculties, give rise to (further\(^{24}\)) beliefs about the natural world as it impresses itself upon our sensory organs. These sensory intuitions cause our belief-forming, and knowledge-forming, representational capacities to be exercised in the ways that they are exercised when we see, hear, etc. aspects of the natural world, such as cars, cats, trees, and the like. The justification or knowledge that consists of, or is derivative from, sensory intuitions seems to some philosophers to be a rather secure kind of knowledge. If these philosophers were right—and they may well be right—then it would be good news for heavy-duty realism if it also turned out that normative knowledge is made possible by a

---

\(^{23}\) The canonical example is Goldman (1979).

\(^{24}\) If they themselves count as beliefs, they cause (at least in part) further beliefs. Otherwise they cause beliefs, sans ‘further’.
human, though perhaps not uniquely human, capacity to form intuitions that are analogous to sensory intuitions, which, however, are about a completely different subject matter, viz. normative reality consisting of normative facts and their constituents, including normative properties and relations.

It is, however, easy to misconstrue the existence of such a capacity as an empirical possibility, and then proceed to offer a misguided neurobiological argument purporting to show that brain science is inconsistent with this possibility. Instead of being an empirical hypothesis, however, the claim that agents like us are equipped with a capacity for forming normative intuitions is more of a conceptual claim than it is an empirical one. Assuming that the concept of normative knowledge and that of justified normative belief aren't empty but rather denote something real, it may be that, from the point of view of heavy-duty realism, these concepts are best seen as logically related to the concept of a normative intuition. I think this is so.

The fact that there are a bunch of neurobiological processes going when and only when we acquire new justified normative beliefs, or new items of normative knowledge, in no way bears on the question of whether the concept of a justified normative belief and that of normative knowledge are logically tied to the concept of a normative intuition. Maybe the same cannot be said when we change the topic of conversation to sensory intuitions. Maybe the concept of a sensory intuition is in part an empirical concept, the best understanding of which requires not only the philosophy of perception but also empirical science. This would not be very surprising, for sensory intuitions stand in causal relations to the natural world, and empirical science seems to be the most reliable mode of inquiry into the causal structure of the natural world, which is, after all, the
subject matter of sensory intuitions.

Of course, any given normative intuition, or at any rate its biological basis, is causally imbedded no less than any given sensory intuition. Nor do the biological substrata of these two kinds of intuition differ in their metaphysical makeup. Some branches of so called experipsychological philosophy, with help from the natural sciences, try to map the causal relations of these biological substrata. Whatever the nature of the logical connection of these substrata and their causal relations on the one hand, and the psychological states these structures and these relations underpin on the other (whether this connection is best understood as identity or else something weaker), it is, by the definition of the brand of heavy-duty realism I am here expounding and defending, plain to see that neither these psychological states or their biological substrata consist of the same type of facts as their subject matter, for this subject matter consists of a mind-independent realm of normative facts (which, in turn, consist, in part or in whole, of normative properties and relations).

To sum up with some question-answer pairs: How do the psychological facts, properties, etc. and their biological underpinnings deliver reliable information about the natural world? Well, by being causally related to it. How do these things deliver reliable information about the mind-independent realm of normative facts? Not in the same way. How, then? It is to this question I now turn.

4 A problem and a solution
One answer is that every time you or I or whoever has a normative intuition which is or causes a justified normative belief or item of knowledge we are on the fortunate side of a metaphysical coincidence between how things are in normative reality, and how things
are psychologically and biologically in here in the mind and brain. I prefer to think of this so-called coincidence as no coincidence at all, but rather a marvelous product of intelligent design. As I said, I can't defend intelligent design here, though. Nor do I wish to. Instead, I'll do the best I can without invoking this sometimes heated controversy.

If the coincidence between normative knowledge and normative fact is truly a coincidence, and not a product of intelligent design, then it seems to me that the best way for secular heavy-duty realists to proceed is to downplay the role played by this coincidence as much as possible. This is what I'm going to try to do on these philosophers' behalf in this, penultimate, section of the present, epistemology, chapter.

One way to start is to consider where the epicenter of this coincidence is. Normative facts are, according to heavy-duty realism, out there in the world. Normative intuitions, however, are in here—in the mind and brain. The respective constituents of the two kinds facts belong to distinct ontological categories. When we talk about these constituents in the two cases, we're talking about different kinds of properties and relations. Descartes famously argued that psychological properties and material, including biological, properties are as different from one another as I think normative properties are from either of the former. He thus faced the infamous problem of explaining the nature of the relation between the two kinds of properties he postulated (ignoring God, who consists of a third kind of facts, and ignoring abstracta, of which Descartes has no account). I face a similar kind of problem here.  

Suppose that every time a person correctly intuits a normative fact, the correctness

---

25 I mentioned intelligent design as a possible solution, which is a move that is, in a sense, analogous to Malebranche's solution to Descrates's problem. At least that's one way to read it: there is no interaction between normative intuition and normative fact, if by 'interaction' is meant causal interaction. Instead, what looks like interaction is a continuous set of syncing processes among two kinds of properties—a set of processes I and, in the aforementioned context, Malebranche, take to result from intelligent design.
of his or her intuition is a product of dumb luck. The problem with this is that, in general, beliefs—one might think—don't get to be justified or count as knowledge when they are the products of dumb luck. Heavy-duty realists are therefore going to need to say something about why the chancy sync of normative intuitions and the subject matter of these intuitions doesn't undermine the justificatory status of the beliefs and, hopefully, knowledge we get by correctly intuiting the normative lay of the land.

One way to start is to argue that the chances of this sync occurring on any given occasion of an agent’s correctly intuiting a normative fact are fairly good. Return to sensory or perceptual intuition. It's not as though perceptual intuitions always sync with their naturalistically describable subject matter. But, setting aside deep skeptical worries, these intuitions and their subject matter do indeed sync up nicely on a frequent enough basis for us to think, not unreasonably, that they are under many conditions reliable. Heavy-duty realists can, then, try to say that the same thing is true of normative intuitions and the facts these intuitions are in the business of psychologically representing: statistically, there is a good enough chance that, under normal normative intuition-forming conditions, any given normative intuition will constitute or result in a true normative belief or item of knowledge. If true, this would make normative intuition a reliable belief-forming process, and, even if reliability isn't good enough for justification or knowledge, it's a decent start.

Okay. But why is normative intuition reliable?

The troublesome disanalogy between perceptual intuitions and normative intuitions follows from the fact that there is, apparently, a plausible account of why perceptual intuitions are in sync with their subject matter on a regular enough basis, to
wit causation, whereas the same thing cannot be said, or at any rate isn't obvious, when it comes to normative intuitions. In the good case, the naturalistic facts cause the perceptual intuitions that represent these facts. In the good case, the normative facts are in sync with the normative intuitions that represent these facts because the two are in some sense intended to be in sync, according to the sort of non-secular epistemology for heavy-duty realism I alluded to earlier. But standard heavy-duty realism doesn't incorporate intelligent design. What, then, are the majority of heavy-duty realists to say here?

They should, I think, press further on the analogy with perceptual intuition. Descartes called it a metaphysical doubt that our beliefs about the external world and perhaps even our mathematical beliefs might be false, where their falsity, if it were actual, would be a product of a systematic deception orchestrated by a powerful though malevolent entity—an evil genius. It's hard to conclusively demonstrate the error of this pessimistic hypothesis. But most of us don't worry about it. I mean, even those of us who think about this sort of thing on a regular basis for the most part don't worry about it. Secular-minded philosophers who fall into this group of non-worriers, as far as I can tell, are for the most part implicitly or explicitly comfortable with believing that human beings are in a sense lucky to not be the victims of the sort of hoax envisaged by Descartes on behalf of his skeptic. Sure, there are plenty of ways in which a powerful deceiver of the sort Descartes worried about could have systematically tricked us. It just so happens that we're pretty darn lucky that none of those ways corresponds to how things actually are, since, luckily, there is no super-powerful evil genius to deceive us. The analogous move for secular heavy-duty realists, then, could be to make the claim that the fact that normative intuitions and their subject matter sync up with one another on a
regular enough basis for us to rightly think of the faculty of intuition as a reliable belief-forming process is a product of dumb luck in the same way that the fact that we are not the victims of a systematic hoax orchestrated by a malevolent and extremely powerful deceiver is a product of dumb luck.

I am not certain how fruitful this strategy will prove to be. My hunch is that, with some ingenuity, it will be sufficient to insure that the bullet heavy-duty realists are in all likelihood going to end up biting will be of a relatively low caliber. Conclusive arguments for big conclusions, such as the falsity of heavy-duty realism, are rare in metaethics. Arguments that exploit the problem heavy-duty realism faces in light of heavy-duty realists’ apparent need to explain the remarkable sync of normative intuition and normative reality are unlikely to prove exceptional in this regard. These arguments will, I predict, eventually be filed alongside general skeptical worries, which may be very hard to soothe but are relatively easy to not think about.

Chapter 4: Supervenience

1 The basic idea
Normative facts necessarily depend on, and co-vary with, natural facts. In other words, normative facts supervene on natural facts. It is also true that normative properties
supervene on natural properties. The way(s) in which they do, as we'll see shortly, invites the question of whether normative properties could even be a kind of non-natural, as opposed to natural, thing, which, in turn, calls into question the sharp distinction between natural fact and normative fact entailed by heavy-duty realism. The short answer to this set of questions is that construing normative properties as non-natural presupposes a controversial view of properties, according to which necessarily coextensive properties could nevertheless be distinct. If such a view of properties is to be avoided, non-naturalism about normative properties must be abandoned along with it. Since I think there are antecedent reasons to accept heavy-duty realism, normative property non-naturalism, which is an essential component of heavy-duty realism, needs to stay. Views of properties that are inconsistent with normative property non-naturalism therefore need to be discarded. Unfortunately I can't argue for this result here—at least not to anyone's satisfaction. The best I'm going to be able to do on behalf of heavy-duty realism is argue that there is enough of a stalemate between views according to which logically equivalent properties are identical, and views according to which such properties could nevertheless be distinct. Given such a stalemate, heavy-duty realists should simply accept the view of properties that doesn't contradict normative property non-naturalism and with it, heavy-duty realism as a whole.

Even then, heavy-duty realism faces what some philosophers would describe as a dilemma. Either these realists learn to live with unexplained yet metaphysically necessary connections between what Hume might have called distinct existences (in this case normative properties and natural properties); or else they construe these connections as the products of intelligent design. As before, I'll leave the hypothesis of intelligent design
for another day, and instead argue that unexplained necessary connections between normative facts and properties on the one hand, and natural facts and properties on the other, are not especially worrisome in this context.

2 Kim's notion of weak supervenience
We think that the layout of the normative properties in some sense depends on, and co-
varies with, the layout of the natural properties. But what is this sense? I hope to arrive at a precise enough answer to this question in this section. I will begin by considering three interesting supervenience claims. All three of these claims seem to be true of normative and natural properties. That is, the normative properties supervene on the natural properties in all three senses of 'supervene'. I will, however, appeal to one of the weaker supervenience claims in the next section, even though I think it is too weak to capture the intuitive sense in which we think the normative is dependent on, and co-variant with, the natural.

Let us, then, consider the three supervenience claims themselves. An easy way to distinguish among these claims is to consider which “possible worlds”, or maximally specific ways things are or might have been, these claims rule out. I take no position on the status of possible worlds, except that I take them to be fully specified ways things might be. The complete specification of what the actual world is like, what it contains, and the relations that hold among its constituents, constitutes one possible world. This specification, altered only with respect to the location of a single atom on Mars, constitutes another possible world. Since these worlds are very similar, we might say that they are in close proximity to one another in the space of possible worlds. The language of proximity is only a metaphor, though, since each possible world is not spatially or
temporally related to any other possible world. Even if David Lewis (1986) is correct to
grant each possible world the same ontological status that you and I grant the actual
world, we ought not to confuse the metaphor of proximity with the real thing. These
worlds, even if on an ontological par, are causally, spatially, and temporally isolated from
one another.

Consider, then, a possible world, w, similar to our world, except that in w there is
a “Twin Earth”—a planet qualitatively indistinguishable from w's Earth, but spatially
distant from it. Every individual on w's Earth has a twin on w's Twin Earth whose
thoughts, feelings, actions, and causal histories mirror his or her twin's thoughts, feelings,
actions, and causal histories on w's Earth. Suppose, moreover, that there is, on w's Earth,
an individual, Jones. Suppose she is a good person by most commonsensical standards.
Now suppose her twin, Twin-Jones, who is spatially and temporally located on w's Twin
Earth, and who is exactly to Jones when it comes to thoughts, feelings, and actions, and
even has an almost exactly similar causal history (I'll explain below why I need the
qualifier 'almost'), is not a good person. The supposition seems implausible.

One supervenience claim that could but might not (I'll say why in a moment,
when I explain the above qualifier, 'almost') rule out worlds such as w is weak
supervenience. If the goodness of Jones weakly supervenes on certain natural properties,
then Jones and Twin-Jones could not differ with respect to being good, or not good, as
long as the two individuals are exactly similar with respect to these natural properties. Of
course, this invites the difficult question of which natural goodness supervenes on. After
all, Jones and Twin-Jones do differ in having some and lacking other natural properties,
notably relational properties. Such minor, seemingly insignificant (morally insignificant,
that is) differences in relational properties are why I needed the 'almost' qualifier above. Jones has the property of being a citizen of w's w-Earth. Twin-Jones doesn't have this property—she has the property of being a citizen of w's w-Twin Earth. But surely this difference among the two individuals cannot explain why Jones, but not Twin-Jones, is a good person, if such a profound—or any—moral difference existed between the two. For no moral difference could exist, given the ways these twin-like people are alike, and the ways in which they are unalike. If we could narrow down the set of natural properties in virtue of which a person is good or as the case may be, not good, to at least a level of specificity more strict than that of all the properties a person has—and I don't know how much more strict such a set would have to be—then weak supervenience would provide us with an elegant restriction on the space of possibilities, on which rules out worlds relevantly similar to w. This is of course a difficult project in normative ethics and metaphysics, which I postpone for another occasion.

Even though weak supervenience can rule out worlds such as w, with the just-mentioned and rather large proviso, it cannot rule out certain pairs of worlds that we would intuitively like to rule out. I'll say a little more about weak supervenience before moving on to the more interesting supervenience theses which can rule out the pairs of worlds I have in mind. Rather than proposing my own definition of weak supervenience or the other two supervenience claims to be discussed here, I'll borrow two of Jaegwon Kim's (1987), and one of Frank Jackson's (1998), definitions, and work with those.

Here, then, is Kim's definition of weak supervenience, where a set of properties, A, is said to weakly supervene on a set of properties, B (A, we may say, is the supervenient family of properties and B the base family of properties):
Necessarily, for any x and y, if x and y share all properties in B, then x and y share all properties in A. In other words: “indiscernibility in B entails indiscernibility in A” (Kim 1987: 315). An equivalent formulation of weak supervenience is this:

\[(WS2) \text{ Necessarily, for any object } x \text{ and any property } F \text{ in } A, \text{ if } x \text{ has } F, \text{ then there exists a property } G \text{ in } B \text{ such that } x \text{ has } G, \text{ and if any } y \text{ has } G, \text{ it has } F\]

A and B are taken by Kim to be closed under the standard Boolean operations of conjunction and negation. In other words, if G is a property in B and H is a property in B, then (G and H) is a property in B, too; and so is (G or H); and so is not-G, and so on. It is, as I claimed a moment ago, an elegant means by which to dismiss as impossible putative possibilities along the lines of w.

Unfortunately, weak supervenience isn’t enough for us meta-ethicists—not for the job we want a supervenience claim to do. And the reason it isn’t is that weak supervenience doesn’t rule out certain pairs of worlds that we would intuitively like to rule out. And the reason it doesn't is that weak supervenience is not a cross-world relation. It is a relation between the A-properties and the B-properties within a single possible world. In other words, the conditional in the second formulation, (WS2), which says that if A has G it has F, too, holds at individual worlds, and doesn't hold across worlds. This suffices to rule out w, because w's w-Earth and its population, as well as w's w-Twin Earth and its population, are constituents of w—a single world. The natural properties with respect to which Jones and Twin-Jones are exactly similar, for example,

---

28 G or F is of course equivalent to not-(not-G and not-F).
being generous, honest, and adept at overcoming fear—and we can call these the B-properties—are the base properties, and the property of being good is a supervenient property, which we can call an A-property. Weak supervenience tells us that necessarily, Jones's being a good person implies that there is some set of properties, say being generous, honest, and adept at overcoming fear, which Jones has, and that if Twin-Jones has the latter set of properties, then she, too, is good. So, weak supervenience rules out the possibility that Twin-Jones might not be good even though Jones is, despite the fact that both are generous, honest, and adept at overcoming fear.

3 Kim's notion of strong supervenience
But it seems to me to be no less absurd than a world that is just like this one in all natural properties, except my cat Eliot's counterpart in this other world is evil. Let me assure you that even if non-human animals are capable of evil, this one is just about as far from evil as a sentient being can get! (Well, toward people at any rate.) Strong supervenience rules out pairs of worlds such as the pair I've just described (to the extent that an impossibility can be described).

To rule out this and the like, we need a relation, R, between the B-properties—the natural properties—and the A-properties—the normative properties—such that R holds across worlds: a so-called cross-world relation. This requires that we put the narrow-scope conditional in the above formulation (WS2) within the scope of its own necessity operator. Doing so yields strong supervenience, which, I think, is the relation we intuitively would like to predicate of normative properties and natural properties.

29 I happen to think that 'Jones is a good person' says something about the existence of reasons to prefer individuals with the natural properties that wholly or partially constitute Jones's personal identity—in some sense of 'prefer', whether it be preferences to spend time with people like Jones, or encourage one's children to exemplify Jones-like properties, or whatever.
Here's Kim's formulation:

(SS) Necessarily, for any object x and any property F in A, if x has F, then there exists a property G in B such that x has G, and necessarily if y has G, it has F.\(^{30}\)

As before, we take the subvening B-properties to correspond to some subset\(^{31}\) of all natural properties. The supervenining A-properties then correspond to normative properties. And both sets of properties we take to be closed under the Boolean operations of conjunction and negation. Understood in this way, (SS) rules out not only worlds like w. It also rules out world-pairs such as the one that, were it to exist, would consist of the actual world and a world exactly similar to the actual world, except that the Professor Allan Gibbard in that other world is or would be mean.\(^{32}\)

The trouble with appealing to strong supervenience resembles one of the problems I mentioned above in connection to weak supervenience. For all Kim's (SS) says, the subset of all natural properties we single out as the subvenience base for some normative property or properties could, formally, be an improper subset of the set of all natural properties in a world—in other words, the B-properties could be the set of all natural properties. This special case of (SS) is often called global supervenience. Intuitively, we metaethicists would like to do better. I would, at any rate. I would like to be able to tell you which natural properties make a person good. Not the least because I would like to work on acquiring these properties myself. Maybe these properties are the ones God commands us to exemplify. But many philosophers don't believe in God. And, of the ones

\(^{30}\) See Kim (1987: 316); original emphasis.

\(^{31}\) Formally, this could still be an improper subset, though if it is merely an improper subset, we end up with a special case of strong supervenience called global supervenience, which is the topic of the next section.

\(^{32}\) This is close to being as obviously absurd as the putative world in which my cat is evil.
that do, at least one, viz. me, has serious doubts about the existence of God's commands. Instead, these could, for all I know, be the properties exemplifying which maximizes aggregate utility, or stops me from treating you in ways to which no rational person would consent. And even if we knew some of these answers, it would be even nicer to know which practical possibilities fall under the criteria delivered by these answers, and which don't. Our knowledge regarding these applied- and even more so normative-ethical areas is not nearly at the level I wish it were, which leaves me wondering how much practical mileage I can expect from insisting that we metaethicists insist on formulating the supervenience of the normative on the natural as (SS) minus its special case of global supervenience. Unfortunately this is not a venue in which I can do justice to these problems. I will mention them again in this chapter but without suggesting concrete solutions that tell us which natural properties make a person good, or which thoughts and actions a person needs to have and take in order to count as being good. I turn to global supervenience.

4 Global supervenience
If normative properties globally supervene on natural properties, then any two worlds that are exactly similar in all natural respects are also exactly similar in all normative respects. This of course rules out pairs of worlds where the arrangements of the members' natural properties are exactly similar, while the layout of the normative properties varies from one member to the other. This is great, as far as it goes. To rule out the absurd putative world z, however, requires of that we go farther, all the way to (SS) rather than the weaker claim presently under discussion. For all that, Jackson has threatened the very heart of heavy-duty realism, to wit normative property non-naturalism, with even this
weaker claim.

4.1 Jackson's problem
Here's something very close to Jackson's formulation of global supervenience:

\[(GS) \text{ For all } w \text{ and } w^*, \text{ if } w \text{ and } w^* \text{ are exactly similar in all natural}
\]
\[
\text{respects, then they are exactly similar in all normative respects.}^{33}
\]

Jackson argues from (GS) and some auxiliary premises to (1998: 122-23) to the falsity of normative property non-naturalism (henceforth non-naturalism):

Take a normative sentence, say

\[(A) \text{ Lying is wrong.}\]

Suppose (A) is true at world w1 but false at w*. Then, w1 and w* differ normatively, since lying is wrong at w1, but not wrong at w*. Moreover, the natural properties at w1 are distributed thus and so. They must be distributed in some way or other, because without natural properties there would be no normative properties. This is simply a conceptual truth about normative properties. Now consider all the worlds at which (A) is true, and call these worlds w1, w2, w3, … For each of the wi worlds, there will be a large conjunctive sentence that describes the natural properties of that world in complete detail. Call the descriptive sentence that completely describes the layout of the natural properties in w1 (D1). And call the descriptive sentence that completely describes the layout of the natural properties at the other wi worlds, including w2, (D2). Then there will be an analogous {w3, (D3)} pair, a {w4, (D4)}, and so on for the rest of the wi's and their respective (D_i)s. Now take the disjunction of all the (D_i)s and call it (D). Jackson says that (A) entails, and is entailed by, (D). Wherever—in whichever possible world—(A) is true,

33 Not only does (GS) not entail (SS), it doesn't even entail (WS1)/(WS2). See Kim(1987: 319).
there is a corresponding \((D_i)\) sentence that is also in that world. And each \((D_i)\) sentence entails \((D)\), because a disjunct always entails its disjunction (\('p' entails 'p or not-p'\)). So, \((A)\) entails \((D)\). In addition, wherever—in whichever possible world—a \((D_i)\) sentence is true is a world where \((A)\) is true, too. And each \((D_i)\) sentence entails \((D)\). So \((D)\) entails \((A)\). So \((A)\) entails, and is entailed by, \((D)\). For any normative sentence, there is a corresponding, and extensionally equivalent, descriptive sentence, says Jackson. All normative possibilities can be denoted in purely naturalistic terms, though maybe only in principle, but that's enough to generate the problem.

It seems, then, that correct descriptions of how things are normatively are necessarily coextensive with correct descriptions of how things are naturalistically, though the latter may be practically, though not logically, impossible for creatures like us to state. Is this sufficient to show that normative properties just are natural properties? Does global supervenience rule out property non-naturalism? No.

4.2 A response to Jackson's problem
According to one view, if a particular \(x\) is \(G\), where \(G\)-ness is a property, just in case \(x\) is \(F\), then \(G=F\). On another view, this is not so. Elliot Sober (1981) defends this second view, and Jackson (1998: 125-28) argues against it since he accepts the first. Jackson's argument is far from decisive. Because discussing this complex controversy is beyond my scope here, the structure of my response to Jackson's argument is only this: Jackson hasn't shown that pairs of properties that are like \(F\) and \(G\) in the sense that, if a thing has the one, it has the other, too, are one and the same property. There is a reason I'll go over, from Sober, to think that logical equivalence in this sense is insufficient for identity. It, too, isn't exactly decisive. The move heavy-duty realists should make is to ally
themselves with Sober's preferred view of property identity in this complicated issue within metaphysics, if for no reason but because it saves the heart of our view. I for one happen to be inclined to make this move as a metaphysician, not just a heavy-duty realist in metaethics.

Sober's argument for the distinctness of logically equivalent properties relies on the following principle

(P) If property F makes a causal contribution not made by property G, then F and G are distinct properties

This is, of course, simply an application of Leibniz's law, according to which a difference in properties entails a difference in the entities to which the properties belong. And we may assume that the causal contribution of a property is a property of a property, so that if properties F and G make distinct causal contributions, they have distinct properties, and are therefore distinct properties themselves.

Now consider the property of being triangular, and the property of being trilateral. These properties are unquestionably coextensive. If something is triangular, it is trilateral, and if it is trilateral, it is triangular, too. Might these properties nevertheless be distinct, given principle (P)? Sober has a thought experiment meant to show that the answer is yes. I'll change the case in a few superficial ways, but leave the thought experiment pretty much as Sober presents it.34

The idea, then, is something like this. Imagine a machine that takes line figures, such as the ones displayed below, as inputs, then scans them and produces printouts of the line figures as outputs. The machine has two components. One component is a mode of output sensitive to whether the line figure that is fed into the machine as an input is

closed and straight-sided. If it is, the machine will output a printout of that line figure, and if not, it won't. The second component is a mode of output sensitive to whether the line figure that is fed into the machine as an input has three angles. If it does, the machine will output a printout of the line figure, and if not, it won't. To illustrate more vividly, here are Figures 1 through 4:

![Figures 1 to 4](image)

Inputting Figure 1 will not result in an output from either component of the machine, because Figure 1 isn't either closed or triangular. Inputting Figure 2 will result in an output from the first component only, because Figure 2 is closed and straight-sided. But inputting Figure 2 will not result in an output from the second component, because Figure 2 isn't triangular. Inputting Figure 3 will not result in an output from the first component, because Figure 3 isn't closed, though it has three angles, which will result in an output from the second component of the machine. Finally, inputting Figure 4 into the machine will result in an output from both components of the machine, because Figure 4 is both closed and has three. Sober asks which property of Figure 4 causes the output from both components, and answers that it is the property of being a triangle. It is not, he says, the property of being a trilateral, even though being a triangle and being a trilateral are logically equivalent properties.

The idea, then, is that it is possible, at least in principle, for necessarily
coextensive properties to have distinct causal profiles. The triangular nature of Figure 4 causes the machine to produce outputs from both components, whereas the trilateral nature of Figure 4 does not. Thus, by principle (P), the property of being triangular is not identical to the property of being trilateral.

However, Jackson is unconvinced by this argument. He insists that being triangular and being trilateral are one and the same property (Jackson 1998: 125-28). He thinks we might distinguish being triangular from being trilateral in thought and speech, as when we cognitively and linguistically represent closed-sided and three-angled figures under the one description rather than the other, but he does not think there are two features of reality here. Jackson's view is that it seems to us correct and intuitive to parse Sober's case in the way that Sober parses it, because angles play a causal role in the production of the output from the second component of the machine, whereas sides do not. “But”, Jackson insists, “this only bears on the common ground doctrine that sides are distinct from angles.” (1998: 127). On Jackson's view, it is irrelevant to the supposed distinctness of being triangular and being trilateral that the machine in our thought experiment is sensitive to angles and not to sides in the way that it is, since this only shows that angles and sides are different things, and doesn't show that being triangular is distinct from being trilateral.

I am not sure exactly what to say about our thought experiment, or Jackson's reply. The issue is whether it is credible to claim that our inputting Figure 4 into the machine produces an output from both components in virtue of the fact that Figure 4 is a trilateral. If one has an antecedent conviction to the effect that being triangular is the same thing as being trilateral, then one will be inclined to say that it is not wrong to think
that the machine produces outputs from both components in virtue of Figure 4’s being trilateral. But if one has the opposite antecedent conviction, one will say otherwise. In short, if being trilateral is the same thing as being triangular, then the fact that both components of the machine produce outputs from our inputting Figure 4 does happen in virtue of the fact that Figure 4 is a trilateral. But if being triangular is distinct from being trilateral, this is not so. Unfortunately I am not sure that there is a non-question begging means to weigh in on this point one way or another—not without taking us chin-deep into a muddy collection of metaphysical controversies.

Does any of this help answer the question of whether the fact that normative properties are logically equivalent to natural properties implies that normative properties just are natural properties? Since the Sober thought experiment is not decisive, it may seem tempting to appeal to principle (P) directly. Principle (P) says that if two properties have distinct (second order) properties, then they are themselves distinct properties. In chapter 1 we saw that non-naturalists construe normative properties as being causally inert. Now, if normative properties are causally inert, and if the natural world, including all natural properties—including the supervenience base of the normative properties—is causally closed, then it would seem that the base properties upon which normative properties supervene have a feature that the supervening properties don’t: the base (natural) properties feature in causal relations, whereas the supervening (normative) properties do not. If this is so, then principle (P) implies that normative properties are distinct from natural properties, according to this account of normative properties—an account according to which normative properties are causally inert.35

35 Non-naturalists who affirm the causal efficacy of normative properties may therefore have some reason to appeal to something along the lines of the Sober thought experiment, since they cannot claim that normative properties differ from natural properties by being causally inert.
However, one might think that the strategy of distinguishing between normative and natural properties by appealing to the causal of efficacy of the natural, and the causal inertness of the normative, begs the question against Jackson's view that necessarily coextensive properties are identical. We know that, given global supervenience, normative properties are necessarily coextensive with natural properties. According to Jackson, this implies that normative properties just are natural properties. But on my view, normative properties don't stand in causal relations. According to the causal closure of the natural world, all natural properties do stand in causal relations with one another. Conjoining my view with the thesis that the natural world is causally closed, then, has the effect of contradicting Jackson's view of property identity. The question is which combination of commitments is more plausible: the causal closure of the natural world plus the causal inertness of normative properties and the negation of Jackson's views on property identity; or else the causal closure of the natural world, Jackson's views on property identity, and the negation of my view.

Since I'm not going to delve any deeper into issues surrounding property identity, I'll end this section by officially adding to my version of heavy-duty realism Sober's view of property identity, which contradicts Jackson's. Sober's view is coherent, and not clearly worse than Jackson's, and that will have to be enough for today.

In the remainder of the chapter, I'll discuss what I take to be The Big Worry for my view in relation to the supervenience of the normative on the natural.

5 Explaining supervenience
5.1 A scanlonean strategy
The fact that normative properties supervene on natural properties is itself a normative
fact. I at any rate think that it is, and so does Scanlon, despite the many disagreements among his views and my own. In the present dialectical situation, however, I'll rely on a distinction of his.

An, as I'll call this species of truthbearing structure, impure normative sentence, claim, proposition, or thought, if true or factual, has naturalistically describable implications. An example of an impure proposition (or whatever—I'll mostly speak of propositions, though little turns on this) is this:

(B) The fact it's raining and Jones wants to stay dry gives her a reason (not necessarily a decisive one) to take her umbrella.

(B) is an impure proposition, because if it weren't raining, or Jones didn't want to stay dry, or didn't have an umbrella, then the naturalistic facts affirmed by (B) wouldn't give her any sort of reason because they wouldn't exist. In other words, (B) implies that it's raining and that Jones wants to stay dry and that she has an umbrella, and these are naturalistic facts if they are facts. So the truth or falsity of (B) depends in part on naturalistic facts, so (B) is partly normative and partly naturalistic, thus normative but impure or, henceforth, just impure.

Pure normative propositions, such as

(C) If it were raining, and if Jones preferred to stay dry, she would have a reason to take an umbrella

do not entail, nor are they entailed by, naturalistic facts. Scanlon appeals to what I'm calling the pure vs impure distinction—he uses 'mixed' instead of 'impure'—when (among other times) he describes his version of normative property (relation, more

---

36 I'm not going to distinguish these truth-bearers in any systematic way. Philosophers of language and mind, please feel free to arrange the concepts denoted therein as you see fit.
accurately) non-naturalism

the thesis of the autonomy of the normative from the naturalistic [is] the thesis that no pure normative claim is entailed by any combination of claims about physical and psychological facts. Given any combination of non-normative claims, it is a further claim that some pure normative sentence is true. (2009: ms. 27)

He adds that “the truth of pure normative claims, by contrast [to impure claims, propositions, or whatever], does not depend on, or co-vary with, non-normative [naturalistic] facts” (2009: ms. 33). In other words, the fact, if it is one, affirmed by (C) does not supervene on the natural facts. Arrange the natural facts however you like, (C) will be true (or false) irrespective of these facts, for (C) is necessarily true, thinks Scanlon.

But (C) only ascribes properties in a possibly-counterfactual situation. It says that if things were naturalistically thus and so, then certain normative properties would be arranged in a particular way. By contrast, (B) directly ascribes a normative property to a naturalistically describable situation. Since (B) ascribes a normative property, and since we know that normative properties supervene on natural properties, we might wonder why, or how, given Jackson's result from the previous section, (B) is not entailed by some set of naturalistic propositions. I'll return to these questions shortly.

We might also wonder which natural properties the normative property ascribed by (B) supervenes on; and we might wonder what the correct explanation of this supervenience relation is. Well, if the fact that it's raining and Jones wants to stay dry whilst walking gives her a reason to take her umbrella, then Jones's having this reason supervenes on this, or rather these facts, to wit the facts that (i) it's raining, (ii) Jones wants to stay dry, (iii) she is about to start walking, and (iv) has an umbrella at her disposal. Scanlon agrees, then, that the normative aspects of impure propositions do
supervene on natural facts, and he clarifies that “which non-normative [naturalistic] facts they depend on is a normative matter, determined by the truth of pure normative claims” (2009: ms. 32). In other words, what explains the supervenience of the normative on the natural are sets of pure normative propositions—ones that are true. If (C) is true, it explains why the reason Jones has, which is correctly ascribed to her by (B), supervenes on the naturalistic facts entailed by (B), viz. (i) - (iv).

A virtue of this normative explanation of supervenience is that it specifies the base properties on which a given normative property, albeit an essentially relational property (of F's being a reason for x to A), supervenes by way of the naturalistic contents of mixed claims. So, this strategy allows us to appeal to strong supervenience. For example, we can know that in every possible situation in which it's raining, Jones wants to stay dry whilst walking, and has an umbrella at her disposal, she also has a reason to take this umbrella on her walk. Thus, we can rule out pairs of worlds that are exactly similar with respect to the distribution of natural properties, except that a single particle on a distant planet in one world is in a slightly different spatiotemporal location than it is in this planet's counterpart in the other world, but where the Jones in the first world has a reason to take an umbrella though her counterpart in the second world does not. It doesn't bear on the question of whether Jones has a reason if she stands in one or else another, similar relation to the distant particle, because her having the reason she has supervenes on facts that are insensitive to which of these similar relations causally connect Jones and the

37 Original emphasis.
38 I'm using 'counterpart' naively in the sense that I take no stand on issues surrounding cross-world identity. In other words, I'm not committed to Lewis's (1986) counterpart theory, according to which it's being possible that I have an eye color other than blue is equivalent to my having a modal counterpart in another possible world who has eyes that aren't blue. Nor do officially reject cross-world identity, according to which I and that counterpart are numerically identical.
Next up are a couple of issues, the first of which I've already mentioned.

5.2 A complication
Jackson tried to show that global (and therefore also strong) supervenience of normative properties on natural properties entails that, for any true normative sentence, there is a true naturalistic sentence that entails, and is entailed by, this normative sentence. Does the pure-impure distinction rule this out? After all, Scanlon says that pure sentences, when true, are necessarily true. To that extent, Jackson's result is a bit misleading.

This is not to say that there aren't any necessarily true subjunctive English sentences with naturalistically describable contents. 'If there is any water, it consists of two hydrogen atoms bonded to an oxygen atom' may be such a sentence. The way I introduce people to the idea of supervenience is by saying that the A-properties supervene on the B-properties just in case there's no A-difference without a corresponding B-difference. Since the referents of necessarily true sentences are fixed in the sense that there's never an A or B or whatever difference, the supervenience of the A-properties, assuming that such a sentence correctly describes some formation of these properties, on the B-properties, if true, is trivially true. But the triviality or else informativeness of such sentences does not bear on Jackson's result as he presents it. More precisely, we'd need to know more about what Scanlon means by 'entailment' when he says that pure sentences neither entail, nor are they entailed by, naturalistic sentences in order to proceed further in evaluating this complex matter. Does the entailment of 'Q' by 'P' consist of nothing other

39 I know this may be inconsistent with certain views of property identity. If you subscribe to one of these views, then ignore my preference for strong over global supervenience, since, to you, it probably makes no difference which of these kinds of supervenience we attribute to normative and natural properties.
than the fact, assuming it is one, that all the worlds at which 'P' is true, 'Q' is true, too? In that sense, every necessarily true sentence entails every necessary true sentence. So, presumably Scanlon has something else in mind with his notion of entailment. My guess is that he has, at least, intensional contexts in mind. Jones may self-referentially believe (C) without believing any necessarily true naturalistic subjunctives, though (C) and any such naturalistic subjunctive is true in all the worlds in which (C) is, to wit all worlds. So, if the logical equivalence of (C) and one of these subjunctives requires that the two be substitutable salva veritate in intensional contexts, then clearly they are not logically equivalent.

As far as the putative co-entailment of (B) and some, possibly gigantic, naturalistic sentence goes, the same is true: (B) and this sentence, whatever it is, are not going to be substitutable salva veritate in intensional contexts, though they may well (by Jackson's result) be true at the same set of possible worlds.

Similar things can be said about deducibility. Is someone who knows that (B) is true (assuming it is) thereby in a position to know that some gigantic naturalistic sentence that is true at the same worlds as (B) is in fact true, too? Not if that someone is you or me. And, even if the naturalistic sentence in question turns out not to be gigantic, I still won't be in a position to know that it's true just by knowing that (B) is true. Not unless I'm a philosopher who has correctly reduced normative discourse to naturalistic discourse and remembers enough of the reduction.

In short, I'm pretty sure that we don't have to worry about Jackson's result invalidating Scanlon's point about pure claims, viz. that they are not deducible from naturalistic claims, nor are these claims substitutable salva veritate in intensional contexts
with pure claims. If Scanlon has more in mind, I'm not sure what to say about it. And if he doesn't, then I agree with him completely.

5.3 An apparent problem
5.3.1 What the problem is supposed to be
Why are the three supervenience claims we've considered true? What has the heavy-duty realist to say by way of explanation of the normative on the natural?

Pure facts are necessarily true. And impure facts supervene on the natural facts they incorporate, and they do so because of the pure facts that logically connect the normative components of impure facts with these facts' normative components. Why does (B) supervene on the naturalistic facts it entails? Because (C) is true. Any given pure claim, we can now see, is true in virtue of some further pure claim, or else it's just true and that's all there is to it. At least this is so on Scanlon's view. The pure claims Scanlon is pressed into saying are just true and that's all there is to it are, on my heavy-duty realist view, true because they correspond to normative reality. These claims are irreducibly normative, yes. But a claim's being irreducibly normative, according to my domain-neutral method of metaphysical inquiry (cross-ref), doesn't exempt it from falling within the scope of this type of inquiry. Such claims, when true, describe normative reality, thus reality. And the fundamental structure of reality understood as the totality—not just the physical or naturalistic parts—of what there is and what it's like is precisely what metaphysicians are in the business of studying. Keeping in mind the distinction between Scanlon's domain-specific and my domain-neutral approaches to metaphysical inquiry,\[40\]

\[40\] To oversimplify a little: I, but not Scanlon, think there is a place—in a sense a privileged place—for metaphysics in our quest for philosophical knowledge. Questions of fundamental existence are metaphysical questions. If it's numbers whose fundamental existence or lack thereof we're wondering about, number theory, unlike the philosophy of mathematics, won't help. If it's reasons we're interested in, we have to ask the hard metaphysical questions, such as: Is the property of F's being-a-reason-for-X to-A real, or else something we color a normatively neutral world with through our thought and talk?
all of this should be intelligible enough by now.

Here's the problem. Don't my view and Scanlon's alike posit inexplicable logical relations between completely unalike properties, to wit natural properties and normative properties? I'm not going to suggest anything on Scanlon's behalf, instead focusing on the problem as it relates to heavy-duty realism.

5.3.2 Why the problem is only apparent
These two sets of properties are properties of the same thing, viz. fundamental reality.

The truth of a pure claim is a property of fundamental reality. The normative property this claim denotes also belongs to fundamental reality. Now suppose this pure claim, though expressed as a conditional in the subjunctive mood of natural English, has a true antecedent. More concretely, (C) and (B) are both true. (C) attributes a property to fundamental reality, and so does (B) including its naturalistic components. Perhaps the question I'm supposed to answer is why the irreducibly normative property denoted by (B), viz. that of being-a-reason, is a property of the same fundamental reality as the natural properties denoted by (B). This is not a straightforward question (to say the least), nor do I quite understand what a right or wrong answer would even look like. The two-ness of my couch cushions is a mathematical property of a thing that is reducible to the physical aspect of fundamental reality. Yet two-ness is a mathematical property, which, I think, makes it a property of fundamental reality. How can, or why does, something physical have an irreducibly non-physical property like two-ness? Regarding how: the physical properties, and the two-ness, are in this case properties of the same two things, viz. my couch cushions. If asked how a single thing can have totally unalike properties, I reply: the same way a thing can have properties in the first place. Regarding why: insert
the true causal story of the design and manufacturing of the couch to which the cushions belong, and the transfer of ownership of this couch to me. Okay, the objector might think, but why do the cushions have two-ness as a property? Because, I reply, when we count up the cushions there are exactly two.

If I am supposed to add something—if the problem is not, as the title of this subsection implicates, merely apparent but also real—then I say that I don't know what this something is and it's not my fault.

6 Conclusion
In this chapter we looked at the other Big Problem a lot of heavy-duty realists worry about: supervenience. The first is, of course, epistemological (see chapter 3). Next I'm going to set the stage the Big Problem I show threatens the anti-realist competitor to heavy-duty realism with which I'm concerned here. This is expressivism. And the stage-setting is the negation problem. Not only the stage-setting but also the main performance are within the philosophy of language and logic. For reasons I won't go into here, defending expressivism has become in large part an exercise of applied philosophy of language and logic, and it is into this territory that I now venture.

Chapter 5: The Inconsistency Problem

1 Stage-setting
1.1 Expressivism
Expressivism is these days often enough conceived as (i) a theory of moral judgment. To wit moral “judgments” are non-belief-like, and more like decisions, plans, intentions,\footnote{There is a great deal of controversy surrounding the question of whether planning or intending or even preferring X involves having a belief of some kind. I think that in the first two cases it does, and that in the third case it may, but postpone a defense of these views for another day.}
preferences, or the feelings, if any, that accompany these states of mind. And expressivism is conceived as having (ii) an essential metasemantic component. 'What does 'P' mean?' is a semantic question. 'Given that it means P, how and why did it end up with this meaning?' is a metasemantic question. Expressivists hold a view which I call psychologicalism (to be further discussed below) according to which facts about the meaning of 'P' are true in virtue of the psychological states 'P' expresses. And, since normative thought is in the business of non-belief-like, that is, non-cognitive states of mind, normative talk gets its meaning from some of these states, whether they are plans, preferences, feelings, or whatever.

Construed as a semantic thesis, expressivism is obviously false, since it is plain to see that the meanings of simple normative sentences of the form

x is N (where N is a normative property)

that are embedded into intensional contexts such as

S thinks x is N

do not, or at any rate need not, refer to S's non-cognitive state of mind. Any competent linguist can tell you that 'Jack believes jaywalking is wrong' does not mean that Jack hates, disapproves of, or plans to avoid jaywalking. This is why expressivists posit a deep, metasemantic structure to our normative talk—a structure that is to be read off of psychological states.

But there's more. The logic of our normative talk is given by the psychological states we express by saying and writing normative sentences. I'll focus on one of the simplest logical operations—negation—and the essentially related subject of inconsistency.
1.2 The inconsistency problem introduced
Explaining why sentences of the form 'p' and '~p' are inconsistent is a problem for everybody. Mentalists are no exception. Unlike traditional attempts to solve this problem, however, mentalists start with the claim that sentences of this form are inconsistent because the psychological states they express are inconsistent. The inconsistency problem for mentalism consists of the need to explain what makes inconsistent psychological states inconsistent.

The phrase 'inconsistency in psychological states' is far more ambiguous than 'sentential inconsistency'. When people use 'inconsistent' as a predicate describing a set of meaningful indicative sentences (henceforth sentences42), we're not left wondering about things like whether the inconsistency in question is bad or undesirable, or something to be avoided. It's a formal property. That's all. This is not so when the subject to which this predicate is applied stands for a person's state of mind.

In ordinary thought and speech, it makes perfectly good sense to say, for example, that Bill's wish to have a drink now is inconsistent with his wish to, as of last week, never drink again. But some metaethicists43 use 'inconsistent' in a sense in which Bill's wishes aren't inconsistent. When discussing psychological states, these people reserve this term for a property of mental states that, as Blackburn put in an influential paper from the late 1980s,44 matters to the agent. I take Blackburn's use of 'matters' to signal that the kind of inconsistency in question involves an element of rational pressure to which the owner of the inconsistent state(s) is subject. Bill may be under overwhelming rational pressure to

42 I don't address other forms of sentence here. 
43 See, for example, Schroeder (2008). 
44 The paper, called “Attitudes and Contents” (1988, reprinted in 1993), is a response to G.F. Schueler (1988). My critique of expressivism is in some ways a development of Schueler's.
refrain from having a drink now. But the source of this pressure isn't the inconsistency of his having this and the contrary wish to never drink again after last week's social, legal, rational, and moral disaster. The source of this pressure is, rather, a large and historically informed set of considerations pertaining to Bill's drinking habits. Similarly, when I have made up my mind to order just a single iced coffee and wish to get a Venti-sized one, but also wish to get a Trenta-sized one with three shots of espresso mixed in, I, like Bill, need to weigh reasons for, and reasons against, acting on each of my two wishes. Neither, or just one, or perhaps both options are irrational. Maybe I need to stay away from that much caffeine. But there is no irrationality simply in my having inconsistent wishes.

The kind of inconsistency mentalists need to explain is a kind that matters in the sense of involving an element of rational pressure on the person whose psychological states are inconsistent—pressure that wouldn't be there without the inconsistency of those states (or that state\textsuperscript{45}). Everyone who has inconsistent beliefs is irrational to some, in many cases minor, extent. Take Adam, who believes that all dogs bark, but also believes that some dogs don't bark.\textsuperscript{46} (Pardon the artificiality of the example, which is there only to make the exposition easier.) The main topic of the sections following this one is the question of whether non-cognitive attitudes or states of mind can be inconsistent in the way that Adam's beliefs are inconsistent (henceforth in the belief-like way).

Why insist on the same, rather than just a close-enough, way? For one thing it would be unclear at best, especially to philosophers who don't work in metaethics, whether any weaker sense of inconsistency could form the basis of a logic of normative language and thought.\textsuperscript{47} I'll return to this deep problem toward the end of the chapter.

\textsuperscript{45} Say a single belief to the effect that a particular shape is a quadrilateral circle.
\textsuperscript{46} It turns out that there's an African species of dog which never bark.
\textsuperscript{47} Indeed, it's not clear to many philosophers (myself included) how a property of psychological states could do this at all.
Also, it seems difficult to specify what is, and what isn't, close-enough and explain why. This I'll return to as well.

1.2.1 A component problem
Assume for the moment expressivism and a close-enough strategy. Simple normative sentences of the form x is N, expressivism says, get their meanings from the non-cognitive attitudes they express. Sentences about chemical compositions, everyone says, express beliefs if they express anything. 'Water = H\textsubscript{2}O' expresses the belief that water is H\textsubscript{2}O. 'Water ≠ H\textsubscript{2}O' expresses the belief that water isn't H\textsubscript{2}O. These sentences are inconsistent because the beliefs they express are inconsistent, or so say expressivists qua mentalists.\textsuperscript{48} Suppose expressivists say that 'Jaywalking is wrong' expresses the speaker's acceptance of a plan of action that forbids jaywalking. 'Not-jaywalking is wrong' then expresses the speaker's acceptance of a plan that requires jaywalking. These sentences are inconsistent, then, because the states of mind expressed therein are inconsistent in a close-enough sense, according to our supposition.

But notice that we now (that is, under this supposition) have a disjunctive explanation of sentential inconsistency. Some sentences (such as the former pair) are inconsistent in virtue of expressing states of mind that are inconsistent in the belief-like way. For a set of psychological states to be inconsistent in this way is for this set to have one property: the belief-like inconsistency property. Other sentences (such as the latter pair) are inconsistent in virtue of expressing states of mind that are inconsistent in the, or a, close-enough way. For a set of psychological states to be inconsistent in this way is for

\textsuperscript{48} They have to say this because, if they didn't, they'd face a version of the problem that motivates the move from semantics to metasemantics. What evidence is there of a sharp divide between normative and non-normative semantics? None. Is there evidence of such a divide at the metasemantic level? Well, no. I'll have more to say about this, and how I think expressivists should respond to it, below.
this set to have another property: the close-enough inconsistency property. Now the expressivist account of sentential inconsistency starts to look like this: 'p' and '~p' are inconsistent for any substitution of p just in case (and because) the states these sentences express are inconsistent, either in the belief-like way, or else in the close-enough way. The property that explains sentential inconsistency thus turns out to be disjunctive, which is striking enough to stand in need of explanation.

It could, however, turn out that belief-like inconsistency and close-enough inconsistency are features that share some key property, in which case the disjunctive nature of the above explanation of sentential inconsistency would be no objection to expressivism. It serves little purpose, then, to quibble about the viability of a close-enough approach. It seems better to look for such a key property, if there is one. And if there is, it's not going to be too important whether we, in the course of evaluating expressivism, call the way in which normative thoughts as expressivists construe them are inconsistency-apt in the belief-like, or else some close-enough, way. For convenience, I'll collapse this distinction for the time being, and say that close-enough inconsistency is basically the same thing as belief-like inconsistency—that the two really are close enough.

But: if we don't find a key property common to the cognitive and non-cognitive incarnations of belief-like inconsistency, expressivists will have to say something about the disjunctive explanation problem.

2 An idea from Mark Schroeder
Mark Schroeder, though he is a cognitivist (2007), says or at any rate implicates (2008: 43) that expressivists are entitled to assume that inconsistency in intention is belief-like.
Expressivists, he says, are also entitled to assume that the belief-like inconsistency aptitude of intention is not a bi-product of the belief-like inconsistency aptitude of beliefs that necessarily accompany, or partially constitute, intentions, since (expressivists are entitled to assume) there are no such beliefs, which (the thought continues) has been shown by Michael Bratman (2009a, 2009b). Tempting as it is, I won't discuss the arguments by Bratman to which Schroeder refers. Instead I'll assume the conclusion of those arguments for the time being.

It should also be noted that if intentions were inconsistency apt in the belief-like sense, their having this property would not imply or even suggest that expressivists must or ought to construe all normative sentences as expressing states of intending. Instead, all expressivists would have to do is stipulate that the attitudes that feature essentially in normative thinking are capable of belief-like inconsistency, much like intentions. If one non-cognitive attitude, viz. intention, has the belief-like inconsistency property (whatever this property may be), then why not another attitude? There's no conceptual barrier to stipulating that whatever attitude plays the role of normative “judgment” as construed by expressivists has a property that we know is already had by a non-cognitive attitude we are antecedently familiar with. This, as far as I can tell, is roughly what Schroeder had in mind when he briefly addressed the inconsistency problem on behalf of expressivism in Being For.

But things aren't as smooth as Schroeder's strategy may suggest, in part because only belief may be uncritically assumed to have the belief-like inconsistency property. So it's no surprise that a philosopher of intention as careful as Bratman doesn't say or even suggest otherwise in his work. (Of course, since Bratman isn't a non-cognitivist in
metaethics—at least not openly—he doesn't stand to gain much from discussing this property in our sense of it.) If expressivists just stipulated that intention, or some other non-cognitive attitude, has the belief-like inconsistency property, they would be stipulating something that needs to be argued for, because the presence of this property is obvious only in the case of belief. Cognitivist mentalists, then, are in a position to base the logic of normative thought not on the stipulation that there is such a thing as a belief-like inconsistency apt non-cognitive attitude, but rather on the trivial truth that inconsistent beliefs are inconsistent in the belief-like way. In other words, expressivists can't just help themselves to a normative logic. To borrow another phrase from Blackburn, they need to earn the right to it.

One way for expressivists to do this is to give a plausible explanation of how an attitude, belief-like or otherwise, ends up having the belief-like inconsistency property. If expressivists can show that the belief-like inconsistency property isn't a biproduct of the essentially representational role that belief plays in our patterns of thought, they can do all that needs to be done in order to solve, or dissolve, the inconsistency problem. The next section evaluates two approaches to doing this.

3 The common sense approach
This approach I'll call constitutive functional. The reason for calling it this will become obvious shortly. Backing up a little, what this is an approach to is the discovery, rather than stipulation, of a property shared by inconsistent normative judgments, attitudes, beliefs, and so on, as these psychological states are construed by expressivists, and

---

49 Not to mention trivial.
50 We have beliefs in part because we need a psychological mechanism by which we may gain insight into how things are.
inconsistent non-normative, or, as Gibbard sometimes says, prosaically factual, beliefs (henceforth ordinary beliefs), such as Adam's belief that all dogs bark. A property that looks like it might fit the bill is joint frustration of a constitutive function, which is something like a teleological purpose behind one's having a certain attitude. Attitudes are for something, the thought goes, and what they are for tells us something, or everything, about what they are. When they are necessarily incapable of doing what they're for, they are inconsistent in the belief-like way. That, at any rate, is the idea in a nutshell.

3.1 How it's supposed to work
If asked why beliefs such as Adam's are inconsistent, most philosophers who haven't thought about this question for very long would probably say that these beliefs are inconsistent because it is impossible for both to be true at the same time. The natural and perhaps even obvious, though ill fated, expressivist move, which Blackburn makes, is to say that inconsistent moral attitudes (as these attitudes are construed by expressivists) are inconsistent because they can't be jointly realized. Moral attitudes necessarily motivate their owners to adopt policies of reaction to certain aspects of the natural world and to act in particular ways in relation to those aspects (or so many or all expressivists say, think, or assume). If (and probably only if) an agent's having a particular set of moral attitudes makes it impossible for him to do what each attitude motivates him to do, the members of the set are inconsistent. Presumably, the inconsistency in question is belief-like in the sense outlined in the previous section. When, for example, Jack disapproves of jaywalking, but also tolerates it, he is to some extent motivated to express anger at jaywalkers and avoid joining their ranks in virtue of being in the relevant state of

\[51\] I extend the discussion to normative attitudes more generally.
disapproval, but also to ignore these people's behavior (perhaps 'ignore' is too strong, and
'be indifferent' is more accurate, but you get the idea) in virtue of being in the tolerant
state. To disapprove of X is to adopt one policy of psychological and physical reaction
toward X, whereas tolerating X amounts to adopting some alternative policy. Because
these policies can't both be adopted, since the reaction mandated by the one cannot be
implemented alongside the reaction mandated by the other, disapproval of X along with
tolerance of X are an inconsistent pair of attitudes toward X.

3.2 Why it doesn't work
3.2.1 First attempt
Dana wants to live in a non-urban environment, but also wants to live in Manhattan. Her
desires have the very property that supposedly explains why Jack's attitudes toward
jaywalking are inconsistent, viz. both pairs of attitudes consist of members that cannot be
jointly realized. But Jack's attitudes need to be inconsistent in the belief-like way or
something near enough, whereas Dana's desires are inconsistent in some other way, one
which carries with it no element of rational pressure. Therefore an attitude set's being
jointly unrealizable can't be the whole story about why these attitudes are inconsistent in
the belief-like sense. Being jointly unrealizable might be part of the story, but it's
certainly not the whole story, because, if it were, Dana's desires would be inconsistent in
the belief-like sense, which they're not.

Why not? According to Blackburn, inconsistency in, say, mere wishing doesn't
matter because when we wish for things we are “spinning fictions”, as he says, and it's
okay for the fictions we spin to be inconsistent, whereas things may be different in cases
of inconsistent desiring. Inconsistency in, as he puts it, “real” desire “may matter” (1988:
The move of distinguishing wishes and desires in this way, and then rescuing the constitutive-functional approach by saying that inconsistency in desire-proper is just as bad as inconsistency in ordinary belief, makes for a dubious strategy. This is true for several reasons.

First, it is unclear what it is to “spin a fiction” other than to make-believe some set of propositions.

Second, the distinction between wishing for some state of affairs to occur, and wanting it to occur, doesn't seem deep enough, nor does it have the right shape, to justify the strategy I described a moment ago. There are some superficial distinctions between how we ordinarily use the term 'wishing' and that of 'desiring'. In some contexts, it sounds a little odd to say that one wants some event in the past to take place. For example, 'I want to have gone to the store' sounds more strange than 'I wish I had gone to the store'. But there are contexts in which it sounds perfectly fine to talk about past-directed desires. Person A asks: “What do you most desire?” Person B responds: “I most desire to have never acted as I did.”

Third, it sounds awkward to say that you wish for something to be the case, but don't want it to be the case. The two attitudes seem to go together in all contexts.

In short, it seems doubtful that there is a distinction as logically weighty as the one Blackburn is hoping for lurking in this particular neck of the woods. Explaining belief-like inconsistency in the way Blackburn has in mind isn't up to the task.

The problem, moreover, can't be quickly resolved by refining this putative solution to the inconsistency problem by switching to the language of constitutive
functions, or as I first conceived the idea, success conditions. The notion of a constitutive function is to some extent metaphorical, much like a great many philosophical notions. I’ll try to elucidate the structure of the metaphor by starting with a similar, also metaphorical, concept, viz. that of a success condition. Some attitudes are “successful”, we might say, under certain conditions. The belief that p is successful just in case p. That’s the only circumstance in which the belief is in some sense as it should be. The intention to X is successful just in case, in the near future, X.

A success-conditional solution to the inconsistency problem, one which generalizes to non-cognitive attitudes, could, then, take the following form: when, and only when, a set of beliefs are necessarily incapable of jointly fulfilling their success conditions are they inconsistent.

This is very similar to a constitutive-functional account. (I think they basically come to the same thing.) The constitutive function of belief is truth/correctness. Truth/correctness is what beliefs are for, and what they are for explains, in part or in whole, what they are, according to the constitutive-functional account. Belief is not unique in having such a function. Non-cognitive attitudes, too, have functions which tell us something about the nature of these attitudes. The belief-like inconsistency property, then, is the property had by a set of psychological states when, and only when, these states, taken together, necessarily frustrate their constitutive function, where the property of being a set of attitudes that do (or does) this is what I’ll abbreviate as the NFCF property. Solutions to the inconsistency problem that take this form seem appealing to some expressivists because these solutions generalize enough to pay the expressivist’s bet that belief-like inconsistency isn’t a bi-product of the essentially representational, truth
apt, correctness apt nature of belief. Performing the explanatory work in question is the
general notion of a constitutive function. The parallels between this and the success-
conditional approach should be fairly clear at this point. Take a property that ordinary
beliefs and at least some non-cognitive attitudes share, viz. NFCF-aptitude/necessary
inability to meet success conditions, and explain belief-like inconsistency by appeal to
this property.

The problem with this species of strategy, though, is that these putative solutions
_over_-generalize to attitudes we don't think are ever inconsistent in the belief-like way,
including desires and wishes. Mike's state of wanting to go and train at noon, but also
wanting to take a nap at noon, has the very property that's supposed to explain what
makes my intending to go to the store at noon, and also intending to stay at home and
work at noon a belief-like inconsistent state. But there's nothing wrong with being in the
state Mike is in.

3.2.2 Second attempt
Here's a way to try to refine this species of would-be solution to the inconsistency
problem. It starts by moving away from an explanation that says, roughly, that
psychological state M1 is inconsistent in the belief-like way with state M2 just in case
(and because) M1 and M2, taken together, necessarily frustrate their constitutive
function, to an explanation that says that M1 and M2 are inconsistent just in case (and
because), taken together, these states necessarily frustrate a specific constitutive function,
F. The expressivist then assigns a value to F that prevents his explanation of belief-like
inconsistency from over-generalizing in the way that the previous attempt did.

Not just any value of F will suit the expressivist's purpose. All constitutive
functions are by no means created equally in this regard. If F is simply realization, then expressivists are going to run into the very same problem confronting the previous, Blackburn-inspired attempt. In other words, when F is realization, the over-generalization problem looms large, and expressivists get the wrong result: clashing desires are on a par with clashing beliefs. When truth/correctness is assigned to F, the constitutive functional strategy doesn't look so bad, but then this strategy won't generalize in the way expressivists hope. This is because non-cognitive attitudes aren't truth/correctness apt in the same way as ordinary beliefs, for truth/correctness aptitude is a bi-product of the essentially representational nature of belief.

The expressivist can and should, then, acknowledge that realization is an ill conceived value to be assigned to F, but go to suggest that some non-cognitive attitudes—the normative attitudes, whether they are—“aim at” more than just realization. In a recent (2009b) paper, Bratman briefly suggests that inconsistency in intention can be explained by the necessary frustration of a “coordination” function. And Neil Sinclair—an expressivist who tackles the inconsistency problem head on in his (2011)—makes a similar suggestion about moral attitudes. I'll focus on Sinclair, because, as an expressivist, he is directly concerned with the inconsistency problem.

He suggests that the function of normative attitudes is to, in addition to motivating us to act, coordinate thought with action, both inter- and intra-personally. I'll focus on the intrapersonal case. Normative attitudes, on Sinclair's view, are functions (in the

52 I will extend Sinclair's remarks about moral attitudes so as to include normative attitudes more generally.
53 Sinclair's views on these topics are well developed, as I had the pleasure of learning during the course of an exchange prompted by his commentary on a paper I presented at the 2012 Pacific Division of the American Philosophical Association titled “Negation and Normative Belief” where I argue that the first refinement of Blackburn's explanation of belief-like inconsistency doesn't work.
54 Sinclair focus on moral attitudes, though I'm going to extrapolate what he says about them so as to cover all normative attitudes.
mathematical sense) that take features of the world as inputs and yield psychological reactions to those features as outputs—reactions that motivate action and, to that extent, “aim at” realization. In this way, normative attitudes logically, and not just causally, our best thinking about what to do with the internal motivating forces that move us to act. Normative attitudes are thus more complex than desires, and have the dual functions of this sort of coordination, and also realization. The presence of the dual functions is very important, because expressivists need normative attitudes to differ in a relevant way from wishes and desires, and having a coordination function looks, on the surface, like it could be a relevant difference.

Here, then, is my second constitutive-functional pass at solving the inconsistency problem on behalf of expressivism—an attempt heavily inspired by Sinclair's work: when, and only when (and because), a set of normative attitudes together necessarily frustrate their coordination function, or their truth/correctness function, are these attitudes inconsistent.

But now expressivists are going to need a solution to what I'll call the disjunctive explanation problem.

3.3 The disjunctive explanation problem
The disjunctive explanation problem arises for explanations of belief-like inconsistency, or, for that matter, sentential inconsistency, that don't have this feature, which I'll call the unity property. The absence of the unity property is the sort of thing that seems to stand in

---

55 I'll return to the issue of logical vs causal relations of motivating attitudes with normative thoughts in the penultimate chapter.
56 How exactly to construe the relationship between the states that motivate us and the functions Sinclair posits would need to be settled. The relationship might be identity, or else it might be something weaker, say logically necessary accompaniment of these functions with the psychological states that motivate us to act.
need of explanation: why would it matter what kinds of attitudes make up an inconsistent\textsuperscript{57} psychological state (for example, which functional role(s) the constituent attitudes play; or what these attitudes' constitutive function(s) are), or, for that matter, what the contents of sentence are (normative or not) to the nature of their inconsistency? To say that these things do matter is, if I may say, a rather striking claim. And striking claims tend to require explanation. The only way to solve the disjunctive explanation problem, then, is to explain the striking claim that functional roles and contents do, somehow, matter to the nature of psychological and sentential (respectively) inconsistency.

It may be asked why the claim that the correct account of sentential inconsistency lacks the unity property should be found to be so striking. Here's what I take to be the main reason. Why a complex\textsuperscript{58} psychological state is inconsistent, and why inconsistent sentences are inconsistent, and whether these things happen to have one functional role as opposed to another, or, in the sentential case, whether the sentences are normative or otherwise, seem to be two separate questions. 'What are these attitudes for—representing the world or changing it?' seems unrelated to 'Why are these attitudes inconsistent?'. This may be even more clear in the sentential case: 'Are these sentences about what I'm supposed to do or not?' and 'Why are these sentences inconsistent?' have no conceptual or logical connection to one another. Positing such a connection is a move that requires justification.

Here's an argument from analogy. When an automotive insurance company claims that people driving red cars are more likely to get into fatal accidents than people who

\textsuperscript{57} It is belief-like inconsistency which is at issue here.

\textsuperscript{58} One that includes more than a single attitude or simple psychological state.
drive beige cars, it needs to produce statistical data to demonstrate a correlation. Otherwise, people will, or at any rate ought to, find this claim rather hard to believe, because it is far from obvious how vehicle color could bear on safety in the way that braking distance, crumple zones, and the number of airbags in a car, do. Philosophical theories are no more exempt from having to explain striking claims that people have no antecedent reason to believe than automotive insurance companies are.

One way for expressivists to proceed at this juncture is to try to show that there is no disjunctive explanation problem, because the correct explanation of sentential inconsistency is only superficially disjunctive.

3.4 Solving the disjunctive explanation problem
According to my second attempt to arrive at a suitable constitutive-functional solution to the inconsistency problem on behalf of expressivism, the reason inconsistent psychological states are inconsistent (in the belief-like way) is one the presence of one property in the case of ordinary belief, but a different property in the case of normative attitudes as these attitudes are construed by expressivists, to wit non-cognitive. Disjunctive explanations of sentential inconsistency are undesirable unless they are, I'll say, “tolerably” disjunctive, meaning that each disjunct refers to a property, P, had by all of the other disjuncts, where P does the majority of the explanatory work. I'll introduce a convention whereby the relational property of an attitude set's necessarily frustrating its or their truth/correctness function is property α. The property of an attitude set's necessarily frustrating its or their coordination function I'll call β. And the realization function necessary frustration property I'll call γ. Assuming β really does explain belief-like inconsistency in non-cognitive attitude (which assumes that there is such a thing as
belief-like inconsistency in non-cognitive attitude), α and β have a couple of common features. Does either one (or both) of these features meet the conditions I stipulated for a property that can credibly be said to play the role of property P?

First, an agent whose psychological states suffer from either of these properties is going to have difficulty acting. But γ has this feature, too, so it can't be the property expressivists are after. If I want to write philosophy at home for the next three hours, and also want to take a break at the coffee shop in an hour and a half, I'm going to have a difficult time of acting on both of these states of mind. Expressivists therefore need some other property besides the hard-to-act property.

Here is a candidate. When an agent's states suffer from either α or β, this agent is, as a result, irrational. When an agent's states suffer merely from γ, this is not so. States that make it true that the agent to whom they belong is irrational (to some extent, perhaps very minor) are indeed the states expressivists want to, or need to, construe as belief-like inconsistent states.

This, then, is my third pass at a constitutive-functional expressivist solution to the inconsistency problem on behalf of expressivism, and this one is comprehensive in the sense that it tries to explain sentential as well as psychological inconsistency, and links the two together in expressivist fashion: 'p' and '¬p' are inconsistent sentences for any value of p just in case (and because) the psychological states these sentences express are inconsistent in the belief-like sense, where being inconsistent in this sense is constituted by these states' necessarily frustrating at least one constitutive function in a way that makes the owner of these states irrational.

The problem I'm about to turn to is one of circularity. Calling something irrational
is asserting or writing a normative sentence. Because all such sentences must be
construed by the expressivist (qua expressivist) as expressing belief-like-inconsistency-
apt non-cognitive attitudes, the above would-be solution to the inconsistency problem
becomes circular, because it presupposes that expressivism gives an adequate theory of
normative thought and talk. But expressivism is adequate in this regard only if
expressivists can solve the inconsistency problem! They must not presuppose the
adequacy of their own theory in the course of defending it from an object that, if
persuasive, shows the inadequacy of expressivism!

4 Beyond constitutive functions: agent-irrationality
4.1 The basic idea
But consider what work the notion of a constitutive function is doing in the above
account. It has already been shown that appealing solely to the NFCF property doesn't
explain belief-like inconsistency all on its own. Having followed the natural progression
of the argument showing this, I'm now led to believe that the NFCF property isn't even a
component of belief-like inconsistency.

One way to begin to test this hypothesis is to ask whether any irrational states of
mind instantiate NFCF. The answer is yes: not merely some, but indeed all of these states
have this property. Now ask: are any or all states with NFCF irrational? Certainly not all!
Witness inconsistent wishes and inconsistent desires. But there are some that are, such as
the belief that Charlie Jr. is a teddy bear and the belief that Charlie Jr. is a pickle, the
intention to go to the store at one o'clock today and the intention to stay at home all day,
the belief that stealing is wrong and the belief that stealing is okay, and many others.

There are, then, states that exemplify the NFCF property but are not irrational;
and all states that are irrational exemplify the NFCF property. Expressivists need for it to be the case that belief-like inconsistency aptitude is something that non-cognitive attitudes are capable of instantiating, at least in principle. Some non-cognitive attitudes, say intentions, plans, and preferences, are, arguably, irrationality apt. Why not, then, omit NHCF from the expressivist story about belief-like inconsistency aptitude altogether, and argue that belief-like inconsistency aptitude is identical to, reducible to, or explicable by appeal to, irrationality aptitude? I'll call accounts that do this agent-irrationality accounts.

One version of such an account is this: a state of mind is inconsistent in the belief-like sense just in case (and because) its owner is irrational just by virtue of being in it. This account is simple and elegant as well as being, arguably, extensionally correct.

4.2 The circularity problem
Expressivists who take up this strategy explain belief-like inconsistency by relying on a normative concept: irrationality. Expressivists might insist that being irrational is a naturalistic property. Though he is not an expressivist, Michael Smith (1995: 186) argues that a rational agent is simply an agent with a particular psychological makeup. Since—let's assume—all psychological features are naturalistic, and rationality is a psychological feature, rationality is a naturalistic feature. Likewise, I take it, for irrationality.

This, though, is misleading. To be rational or irrational with respect to one's thoughts is to be in a certain psychological state, but also, essentially, to be a good, virtuous, or desirable, or else bad, vicious, or undesirable, agent by virtue of being in this state of mind, which can, but need not (logically), lead to actions that inherit the normative status of the thoughts leading up to them. To the extent that psychology is a natural science, it has no business making claims about goodness, badness, virtue, vice,
and the like. The psychologist can always give a complete naturalistic description of a
person's psychological state, and still leave us wondering whether it's somehow sub-par,
bad, or irrational, for an agent to be in this state. Even psychiatric discourse doesn't reach
into the normative domain, instead restricting itself to non-normative claims, which are
often about health and sickness. But being in an unhealthy state of mind is not the same
thing as being in an irrational state of mind. When I get confused when trying to prove
some logical theorem, and thereby start forming irrational beliefs, I'm not necessarily in
an unhealthy state of mind. Putative explanations of an agent's being in an irrational state
of mind that omit the normative aspect of such states are at best incomplete.

The last substantive point in this section is this. Expressivists, I acknowledge, are
in a position to agree with everything I've said in the previous few paragraphs but then go
on to admit that the agent-irrationality account is correct, and circular in a non-vicious
way. The story might read as follows: belief-like inconsistency is explained by agent-
irrationality; agent-irrationality is a normative property, but we expressivists have a
working account such “properties”, namely we say that they exist only in a deflated
sense. No one can accurately and helpfully say what these properties are in any other
sense, but we expressivists can tell you what it is to ascribe such a property using our
usual treatment of normative thought and discourse.

The final attempt to solve the inconsistency problem on behalf of expressivism
I'm going to look at avoids this form of circularity. But rather than buying expressivism a
normative logic for a reasonable price, it takes advantage of the seller, bordering on theft.
The buyer is Gibbard. I've lumped together his and my account in the same section in
order to do a heads-up comparison between the two. The conclusion I draw is that
mentalism, which is the new-age psychologism and which constitutes the parts of expressivism that are within applied philosophy of language is far more plausible when devoid of the classic expressivist theory of normative judgment: non-cognitivism.

5 Where should mentalists start?
To explain why 'p' is inconsistent with '~p', you have to start somewhere. Some people say that the contents of these sentences, or what they are about— proposition <p> for the first sentence, proposition <~p> for the second—can't both be true, where 'can't' refers to conceptual impossibility, and that this is why the two sentences are inconsistent. But, as Paul Horwich might put it, philosophers of language and logic who do this are trying to explain logic with truth, which is a grave mistake, at least according to him (Horwich 1990). He thinks this because he is convinced that everything there is to know about the nature of truth is exhausted by the set of instances of the schema 'proposition <p> is true just in case p', and the facts about the truth predicate's semantic function.

But if truth conditions do explain consistency conditions, as they are supposed to on the Fregean view, there must be something more to knowing the nature of truth than the things Horwich identifies this nature with. This is so because knowing that <p> is true just in case p, and that <~p> is true just in case ~p, and knowing the semantic roles of the truth predicate, simply says nothing about why 'p' and '~p' are inconsistent sentences. It doesn't even tell you that they are inconsistent sentences. On the other hand, if sentential inconsistency is the result of the conceptual impossibility of one item of content being true in conjunction with another, then truth and falsity are crucial elements in the correct explanation of sentential inconsistency, which mentalists explain in terms of the belief-like inconsistency of the psychological states expressed by the offending sentences. In the
remainder of the chapter, I'll show that if—if—this new brand of psychologism can be made to work, it won't work when our theory of normative judgment is non-cognitivist.

5.2 Gibbard's new account
Gibbard's (2012) idea is to start with a set of primitive—unexplained—conducive to explaining the phenomenon of belief-like inconsistency in a way that is amenable to expressivism. Doesn't this already sound not only ambitious, but highly suspicious, too? The idea is to posit everything you need to explain the stuff you want to explain. But the stuff you want to explain may be a creature of your imagination. So, to show us that it's more than that, you concoct more stuff, which you say isn't the sort of thing that can be explained, even in principle. Suppose I want to explain why, contrary to your beliefs, at least some people survive the death of their bodies (brains included). You have no reason to believe such an explanation is possible. Then I stipulate that Cartesian souls exist, and that, moreover (as Descartes did not insist) they are immortal. And, being the same person just is having the same soul. But I make sure to clarify that the notion of an immortal Cartesian soul is one of my primitive notions—that I, nor anyone else, can explain this notion. You will and ought to be suspicious of my proof for the possibility of survival after the death of one's body. Unless I add the qualification that I'm only showing you how post-brain death survival is possible in a broad sense of possibility, say logical or metaphysical possibility.

Similarly, once qualified as a mere “possibility proof” (as Gibbard puts it in *Thinking How to Live*), showing that expressivism could be true, his new view demands attention. Here it is, then.

First, expressivists are supposed to start with a primitive notion of one
psychological state's “excluding” another. Since this is a primitive notion, I can't give you a definition, so I'll illustrate with a pair of examples. Adam's belief that dogs bark excludes his (or anyone's) belief that dogs don't bark. Micheal's intending to go to the store at a particular time excludes his also staying at home at that time. On top of exclusion, Gibbard adds another primitive notion: an agent's “disagreeing” with a state of mind, where the form of disagreement in question permits non-cognitive attitudes to be apt for disagreement. Putting the two primitive notions together, expressivists get something like this: Adam's belief that dogs bark excludes his belief that dogs don't bark. Exclusion is a symmetric relation, so the second belief excludes the first, too. When Adam believes that dogs bark, he disagrees with the belief (no matter whose) that dogs don't bark. Disagreement, too, is a symmetric relation, so that the belief that dogs don't bark disagrees with the belief that they do bark. Since Adam has both beliefs, he disagrees with himself, which makes his beliefs inconsistent in the belief-like way.

5.3 The inconsistency aptitude problem
Liam make-believes an impossible fiction, say the proposition that some embodied individual, Jones, is in two different places at the same time. Liam's states of make-believe don't exclude one another because states of make-believe aren't exclusion-apt. So why are beliefs? I'll call this the exclusion aptitude problem.

Of course, the expressivist can say that Adam's having inconsistent beliefs amounts to his disagreeing with himself, whereas Liam's having inconsistent states of make-believe does no such thing. But, again, why? Hopefully, expressivists won't say that it's because the former, but not the latter, pair exclude one another. But do they have the resources to do any better? Exclusion and disagreement are both supposed to be
indefinable, after all.

But perhaps the solution to the inconsistency aptitude problem has the form of a non-vicious circle. Or it's an exercise in heel-digging: some states are exclusion and disagreement apt, and others aren't, and that's all there is to it. Moreover, there is a parallel version of the problem for any competing theory of belief-like inconsistency anyway, and these versions are no easier to contend with than the one confronting non-cognitivist mentalism (expressivism). Can't we, for example, point out that a person's desiring x and also desiring ~x isn't irrational whereas his having intentions with these contents is? Can't cognitivist mentalists be asked why Adam's beliefs are problematic whereas Liam's states of make-believe aren't?

In short, yes. In the next subsection, I'll introduce my account and briefly explain how it solves the inconsistency aptitude problem and explains belief-like inconsistency in the process.

5.4 A correctness-conditional solution to the inconsistency problem
The correctness-conditional solution to the inconsistency problem says that for a set of psychological states to be inconsistent in the belief-like sense, the agent to whom they belong must be necessarily guaranteed to be in an incorrect state of mind just by virtue of having these states, and that these states are inconsistent in the belief-like sense because of this necessary guarantee. Simple or complex psychological states that lack this property are not inconsistent in the belief-like sense. A belief is correct just in case it's true belief, and incorrect just in case it's false. This account, then, denies that there is

59 I have no incarnation to take this account home with me. It's more of a gift for mentalists. (And it's not a Trojan Horse!)
60 More would need to be said to apply this account to non-classical logics.
such a thing as a belief-like inconsistency-apt non-cognitive attitude, because non-
cognitive attitudes are never correct or incorrect, though they may be virtuous or vicious,
smart or stupid, reasonable or unreasonable, ambitious or underachieving, etc. Because
correctness and incorrectness play such a substantial role in this account, proponents of
this account will have to deny the adequacy of deflationary or minimal conceptions of
truth, such as Horwich’s (1990) minimalism.

5.5 A correctness-conditional solution to the inconsistency aptitude problem
Since incorrectness explains belief-like inconsistency, the inconsistency aptitude problem
becomes the problem of explaining why beliefs have correctness conditions whereas
other attitudes don't. A belief’s truth conditions just are its correctness conditions, but why
isn't the same true of instances of make-believe? If you make-believe that I have a dog,
your instance of make-believe happens to be false, and, of course, there's nothing wrong
with make-believing something false—even if you know it's false. You are not, just by
virtue of make-believing a false proposition, in an incorrect state of mind. Why, then,
would you be in such a state were you to believe I have a dog?

In a sense, the answer is built into the very concept of belief. The fact that my
having a false belief implies I've made a mistake is a conceptual truth about belief.
Really, the implication I just spoke of holds true in virtue of the relation between making
a mental mistake and having a false belief. The relation in question is identity. And I don't
know how to answer the question of why S's property of having-a-false-belief-that-P is
identical to itself. It just is.

As in the case of make-believe, it makes no sense to speak of incorrect plans,
intentions, preferences, or any type of attitude that can't be reduced in part or in whole to
belief. That's what makes belief special. That's what makes it the uniquely belief-like inconsistency apt attitude. When Tom tells me that my plan to X is mistaken, I must take what he says as an expression of his belief that my belief to the effect that X-ing is a good idea, or that I have a reason to X, or that X-ing makes sense, or whatever, is false. How else would I be able to understand the essence of this strange claim of his?

Gibbard anticipates an account that's formally a lot like the correctness-conditional account I've been expounding, and has an objection ready, to which I now turn.

5.6 Objections to the correctness-conditional account
5.6.1 The bigger mistake objection
Individual beliefs that happen to be false, but could have been true, are just as incorrect as Adam's beliefs about dogs (that they bark, and also that they don't bark). If the problem with Adam's pair of contradictory beliefs is that one of them is false—if this is the basis of the belief-like inconsistency of these beliefs—then it's the same problem as my believing that it's Thursday when in fact it's Tuesday. But surely there is something worse going on with Adam than there is with me. Epistemically, I'm not doing as bad as he is. But, according to this objection, my account implies otherwise.

This objection, however, grossly misconstrues the account it targets. I say that the big problem with Adam's beliefs—the source or basis of their belief-like inconsistency—is not just that at least one is false. This is part of the problem. Fully spelled out, however, the problem is that Adam is necessarily guaranteed to have an incorrect belief just by virtue of having this pair of beliefs. So there is a significant difference in magnitude.
between my believing p when not-p on the one hand, and Adam's believing p and also believing not-p on the other. In the first form of case, the believer gets things wrong but might have gotten things right (assuming p is contingently false), whereas, in the second, he gets things wrong and could not have gotten them right no matter what. Believing necessarily-p when necessarily-not-p, or believing p and not-p (one belief, two items of content), are states that are just as bad to be in as having two separate beliefs that contradict one another. An agent with any belief set (no matter the size of this set) that is guaranteed as a matter of logical necessity to have at least one false member, and therefore guaranteed to make it true that its owner is making a mistake—is in an incorrect state of mind—just by virtue of being its owner, is as bad as any other belief set with this property. Belief sets containing one or more false beliefs, but which don't have this property, aren't as bad nor are they inconsistent.

5.6.2 The smaller mistake objection
This objection can be run as the following argument, which draws on a pair of cases and an intuition I for one don't have:

Case1: I mistakenly believe today is Thursday.

Case2: Allan is working on a horribly complicated math proof, and he makes a tiny, very much understandable misstep in a delicate portion of the proof.

My belief could have been, but isn't, correct, whereas Allan could not have been right about the step he made a mistake on. But—and this is the intuition I don't have—his mistake is more excusable than mine, because, after all, his epistemic challenge is far
greater than mine. But the correctness-conditional theory says that Allan's mistake is bigger, not smaller, than mine.

My response to this is fairly simple. The response is a conditional strategy: to make the revisionary move only if necessary. Insofar as the intuition driving the above argument is widespread, I'm willing to say that people who have this intuition should re-evaluate it, because it is misleading in the following way. It is crucial to disambiguate the badness of an agent's coming to believe a certain false proposition from the level of quality of the belief itself. Maybe—maybe—it's true that just by virtue of my coming to mistakenly believe that it's Thursday, which is something I could easily rectify, I'm a lot more epistemically irresponsible than Allan, or that my lack of epistemic virtue is far greater than his, or whatever. If these things are true, so be it, I’m inclined to say. But when the focus is on my belief itself, and on Allan's set of beliefs rather than Allan as an epistemic agent, the intuition that comes to my mind is that his beliefs, and indeed all beliefs that could never be true, are, in themselves, always bigger mistakes than incorrect beliefs that might have been true under different circumstances. Those who vehemently disagree may see me as biting the bullet on this issue, though I do not. This cost, if it is one, is less severe than the costs of Gibbard's view, to which I now return.

5.7 Gibbard's primitive notions
I've helped myself to a notion of incorrectness, and one of conceptual necessity. Whether a state has correctness and incorrectness conditions is a conceptual matter. Is the same true of exclusion aptitude and disagreement aptitude? In a sense it is. Yet there are important differences which favor the correctness-conditional account over Gibbard's.

Exclusion and disagreement are technical, or at least quasi-technical, notions.
Imagine yourself unfamiliar with Gibbard's work. I casually ask you whether your wanting a cup of coffee excludes your wanting to stay away from coffee? You are likely to ask for further clarification on what I mean by the term 'exclude'—in other words, you want to know what concept I'm working with—or perhaps you naively answer yes. Then I tell you that exclusion is a concept I can explain only by getting you to use it, and that desires never exclude one another, but that contrary plans always do. You ask me what the relevant difference between planning and desiring is, and I tell you that it's part of the concept of a plan that plans you can't carry out together exclude one another, and that it's part of the concept of desire that desires you can't jointly fulfill don't exclude one another. At this point, you probably have some grasp of this new concept. But to you, it is a new concept, as it was for me when I first started using it. However, you and I both used the concepts of correctness and incorrectness as children. These concepts weren't stipulated into existence for a specific theoretical purpose. If I tell you that, because today is Thursday, I can't get together with you this afternoon, you will tell me, in one way or another, that I've made a mistake about which day of the week it is, for it is in fact Tuesday. One advantage of my view over Gibbard's is that it relies on a familiar notion as one of its primitives.

But, arguably the notion of conceptual necessity isn't as untainted as that of (in)correctness. In fact I'll grant that it's something you need to read philosophy to get a handle on (even though you certainly don't need to read philosophy to understand that Adam's belief that all dogs bark, and his belief that some don't, amounts to Adam's having made a serious mistake).

But neither is Gibbard's quasi-technical notion of disagreement going to be
intelligible to anyone, unless he or she has a working knowledge of Gibbard's philosophy. When most people think of disagreement, after all, we (implicitly or otherwise) think that in order to disagree with someone (perhaps even ourselves, though we don't typically talk that way) is to believe that something is the case while the other person believes it isn't the case. Gibbard, however, wishes to stretch the notion of disagreement. He says that it's possible, and indeed common, for one party to disagree with a plan, or with a decision, made by another party. I say “let's pack”, you say, “no, let's not”. With my utterance, I've expressed a plan or intention to pack, and with your utterance, you've disagreed with that plan or intention. That's how a Gibbard-expressivist wants to parse the case at hand.

But what does it really mean to say you've “disagreed with a plan”? We do talk that way—sometimes. What we mean, though, is that when someone “disagrees with a plan”, he or she believes the plan to be foolish, unwise, unfeasible, irrational, or just plain bad. I don't know what else could be meant by 'disagreeing with a plan'. The disagreement involved in my saying “let's pack” and your saying “no, let's not” seems to me to be exhausted by the fact that my utterance implicates that I think packing is a good idea while your utterance implicates that you think packing is not a good idea, perhaps even a bad one. Gibbard just stipulates that disagreeing with a plan isn't the same thing as believing it to be ill conceived. The problem with relying on a technical notion of disagreement stipulated into existence in order to solve a host of other technical problems for a highly technical philosophical theory is that all of this is so far removed by what and how people actually think that it becomes a philosophers' invention. Moreover, it becomes totally unclear why agents should worry about disagreeing with themselves in this quasi-yet-overly technical sense. Expressivists are yet again trying to stipulate
themselves out of a mess they've gotten into which the opposition did not stipulate into existence, but rather observed independently of the dialectic.

A similar point can be made about exclusion. Return to the earlier case in which I ask you whether your wanting more ice cream whilst also wanting to refrain from more ice cream implies that you have a pair of attitudes that exclude one another. Thinking about my question, you might ask the following yourself: “Do you mean that I can't have more ice cream, and also refrain from having more ice cream? If so, then yes, my wants obviously exclude one another.” But we know that joint unrealizability doesn't explain belief-like inconsistency. Yet joint unrealizability is precisely how we'd want to understand exclusion in this sense, but for Gibbard's work.

In short, Gibbard has deeper and more serious issues to contend with by comparison with the correctness-conditionalist mentalist.
Chapter 6: Conclusion

In the previous chapter, I've shown that three approaches to solving the inconsistency problem on behalf of non-cognitivist mentalism, that is, expressivism, aren't up to the task. I explored the shortcomings of these theories, which inspired me to suggest a positive account, one which lacks those shortcomings, but rejects non-cognitivism. The instructive failures of the non-cognitivist accounts lend some plausibility to my conjecture that the root of the problems these accounts can't solve is non-cognitivism itself. For the non-cognitive attitudes with which expressivists want to identify moral (and normative more generally) stances need to be inconsistency apt in the belief-like sense, rather than being inconsistency apt in some other sense. There is a sense in which my wanting another cup of coffee, and also wanting to stick to my plan to reduce caffeine consumption, are inconsistent: I know I can't do both things, so I need to further reflect
on what I'm going to do. But I'm not irrational just by virtue of having this or any set of unrealizable desires. On the other hand, whenever I have inconsistent beliefs, I am irrational to some extent. So, too, is it with having inconsistent moral stances. At the very least, then, the non-cognitive attitudes expressivists aim to identify with moral beliefs need to be apt for irrationality. Plans and intentions (also preferences) arguably have this property, so they are natural candidates to play the role of moral beliefs in an expressivist theory. But, unless these states are subject to a sufficiently robust belief condition, it is unclear why anyone besides philosophers who have antecedent reasons to accept expressivism—whether as true, or merely possible—should think that non-cognitive attitudes are inconsistency apt in the belief-like sense. Were I to count myself among the ranks of the expressivist camp, I would not be nearly as comfortable with, when all is said and done, in effect stipulating that one or more of these types of attitude have this property, as other expressivists may be.

Some antecedent reasons to accept expressivism may include Moore’s (1903) open question argument, which, if sound, casts doubt upon versions of realism that identify normative properties with some subclass of natural properties, along with the metaphysical and epistemological challenges confronting heavy-duty realism, which, if insurmountable, cast doubt upon this version of realism, which I believe to be the correct metaethical view. I therefore argued that these challenges are not insurmountable. I did not argue against naturalistic realism, not via the open question argument or any other. I instead started with what Enoch calls the just-too-different intuition, which, if you have it, makes it very difficult for you to take naturalistic realism seriously, and ought to incline you to take expressivism and heavy-duty realism seriously, for these are the two
metaethical views that do justice to this intuition do a greater extent than any I am familiar with.

When all is said and done, among expressivism and heavy-duty realism, I take heavy-duty realism to be more intrinsically plausible for reasons I have already given and won't repeat. In addition, I have a strong inclination to construe this brand of realism as the closest thing to a commonsensical view we're likely to get in metaethics. The metaphysical and epistemological challenges confronting heavy-duty realism are not damaging enough to justify a move away from this form of realism. To the extent that it is indeed the closest thing to a commonsensical view among its competitors, the reasons we have to reject it on the basis of the aforementioned challenges are not sufficient reasons, and are far from decisive. Moreover, expressivism has its own challenges to contend with, not the least of which is the inconsistency problem. The likelihood of a forthcoming expressivist solution to this problem I take to be small. There is, as far as I can predict, nothing for expressivism in the offing on the subject of the nature of the foundations of the logic of normative thought and talk. Since this thought and talk is governed by logic no less than any other kind, I take myself to have strong reasons to remain a heavy-duty realist in the face of the expressivist challenge.
References


