

Why Believe This or That or Anything at All?

An instrumental account of epistemic normativity

JAMES BERNARD WILLOUGHBY

A thesis submitted in fulfilment of the degree of Doctor of Philosophy in the School of
Philosophy, Research School of Social Sciences at The Australian National University

February 2023

Declaration

This thesis is solely the author's work. No part of this thesis has previously or is currently being submitted for another degree. To my knowledge, any help received in preparing this thesis and all sources have been duly acknowledged.

James Bernard Willoughby

Student Number: U5518699

School of Philosophy, RSSS, The Australian National University

Date: 8 February 2023

To Tara, you were more patient and encouraging than is reasonable. Love.

Acknowledgments

It is cliché for a reason; the list of people I want to thank is enormous; I do not remember everyone; I apologise to those I missed. For me, philosophy is a community collaboration. I have had many conversations about the material in this thesis, heard many objections and workshopped many solutions. Hopefully, this effort has improved the work. I know it has improved my life. I feel utterly grateful to be a part of my little philosophical community.

First and foremost, Alan Hájek has been an excellent supervisor. If you have the chance to be supervised by him, I recommend you take the opportunity.

Thanks to the rest of my supervisory panel: Katie Steele, Frank Jackson, Matthew Kopec and to Thomas Kelly for supervising me during a visit to Princeton.

Thanks to the many philosophers that volunteered their time with comments or influential conversations: Aidan Ryall, Brandon Yip, Brian Garrett, Brian Hedden, Brian Talbot, Bruno Ippedico, Carl Joseph Brusse, Charles Côté-Bouchard, Colin Klein, Daniel Kilov, Daniel J. Singer, Daniel Stoljar, Domi Dessaix, Donald Nordblom, Edouard Machery, Elizabeth Harman, Erick Llamas, Garrett Cullity, Geoffrey Brennan, Gideon Rosen, Hayden Wilkinson, Hilary Kornblith, Jake Stone, Jane Friedman, Jennifer Munt, Jeremy Strasser, Jesse Hambly, Josef Holden, Julia Staffel, Justin Mendelow, Kim Sterelny, Michael Yuen, Michelle M. Dyke, Nicholas Carroll, Noga Gratvol, Pamela Robinson, Phil Dowe, Philip Pettit, Preston J. Werner, Rachael Brown, Ross Pain, Sarah McGrath, Shalom Chalon, Shang Long Yeo, Shauna Winram, Stephen Francis Mann, Szymon Bogacz, Theodore Murray, Timothy L. Williamson, Toby Charles Penhallurick Solomon, Victoria McGeer, Yehezkel Symonds, Zachary Swindlehurst.

And thanks also to audiences at the Charles Sturt University Philosophy Seminar 2022; the seminar at the University of Colorado, Boulder 2019; the Australasian Association of Philosophy Conference 2019; the Rocky Mountain Philosophy Conference, 2019; Reasons and Rationality: themes from Hilary Kornblith's work 2022; and many ANU Philsoc seminars. Not to mention, the ANU morning tea crowd!

Thanks also to *Philosophers Imprint* and *The Canadian Journal of Philosophy*. Chapter 2 is substantially based on Willoughby, J. B., (2022) Embedded Epistemic Instrumentalism: An Account of

Epistemic Normativity, *Philosophers' Imprint*, 22: 7. doi: <https://doi.org/10.3998/phimp.745>.

Thanks to Thom Fulton and Anya Tchoupakov for editorial services on this paper. And Chapter 3 is substantially based on Willoughby, J. B. (forthcoming). Against Universal Epistemic Instrumentalism. *The Canadian Journal of Philosophy*. I thank Joan Keyes for editorial services on that paper.

Abstract

We are bombarded with epistemic standards. Respect your evidence. Don't believe in contradictions. Don't arbitrarily change beliefs. But how do such standards get their normative force? Why should we respect our evidence, for example? In this thesis, I offer an answer: embedded epistemic instrumentalism. It is the potent and persistent usefulness of conforming to the epistemic standards that gives rise to their normative force. I claim that there is an uncontroversial core to this theory, facts we will all agree on and facts that are integrated well with other things we know. These facts are the ontological foundation of the theory. I then argue for a recipe for responding to the main source of objections, intuitive arguments. The result is that we have an extremely well-motivated theory, that is immune from the main source of criticism, or so I argue. These properties make embedded epistemic instrumentalism stand out from the crowd. I offer two relevant interpretations of this theory, error theory and non-analytic natural realism. I defend the error theory interpretation from a charge that epistemic error theories are self-undermining. And finally, I defend the theory from the charge that I have failed to identify any actual normativity in my theory of epistemic normativity.

Table of Contents

1. Introduction	1
1.1. A Moment of Weakness.....	1
1.2. Identifying The Project.....	4
1.2.1. The question.....	4
1.2.2. The area.....	5
1.3. Defining Important Concepts.....	7
1.3.1. Epistemic standards.....	8
1.3.2. Epistemic normativity	8
1.3.3. Epistemic reasons	9
1.3.4. Epistemic.....	9
1.4. The Gist of My Position	14
2. Embedded Epistemic Instrumentalism.....	16
2.1. Conforming is Useful	18
2.2. The Embedded Goal Model.....	20
2.2.1. The Embedded Health Goal Model	22
2.2.2. The Embedded Epistemic Goal Model.....	25
2.3. The Too-Few-Reasons Objection	27
2.4. The Advantages of Embedded Epistemic Instrumentalism.....	33
2.5. Embedded Epistemic Instrumentalism and Other Topics in Epistemology	39
2.5.1. Different types of epistemic instrumentalism.....	39
2.5.2. Evidence.....	40
2.5.3. Demarcating the epistemic from the non-epistemic	42
2.5.4. The epistemic standards.....	44
3. Against Universal Epistemic Instrumentalism	48
3.1. Epistemic instrumentalism and the universal response.....	48
3.2. Cheap Universal Epistemic Instrumentalism.....	51
3.2.1. Schroeder's cheap reasons	51
3.2.2. The otiose objection.....	52
3.2.2.1. Cheap reasons create a lacuna.....	52
3.2.2.2. Filling the lacuna makes cheap reasons otiose	53
3.3. Rule Universal Epistemic Instrumentalism	55
3.3.1. Rule-based reasons	55
3.3.2. The collapse dilemma.....	56

3.3.2.1.	Avoiding the collapse: realistic rule universalism.....	56
3.3.2.2.	Avoiding the collapse: non-act based instrumental reasons.....	58
3.3.2.3.	Accepting the collapse: the otiose objection, part deux.....	59
3.4.	Non-Authoritative Universal Epistemic Instrumentalism	61
3.4.1.	Non-authoritative reasons	61
3.4.2.	Non-authoritative universal epistemic instrumentalism is in tension with itself....	63
3.4.3.	The distinction does not help with too-few-reasons cases	64
3.5.	Common Goal Universal Epistemic Instrumentalism	69
3.5.1.	Common goal reasons.....	69
3.5.2.	Côté-Bouchard’s common goal of acting for genuine reasons	70
3.6.	Universalism’s Future	71
4.	Why Accept Embedded Epistemic Instrumentalism?	75
4.1.	The Evidence For	76
4.1.1.	The uncontroversial foundation of embedded epistemic instrumentalism.....	76
4.1.1.1.	Explaining the fact that conforming is extremely useful	77
4.1.1.2.	Fitting-in with evolution.....	77
4.1.1.3.	Explaining the manifestations of epistemic normativity.....	79
4.1.1.4.	The uncontroversial core of embedded epistemic instrumentalism	79
4.1.2.	Embedded epistemic instrumentalism’s model-invariance.....	80
4.1.2.1.	Contrasting theories	80
4.1.2.2.	Model-invariant theories vs. model-bound theories.....	82
4.1.3.	Summary of the case for embedded epistemic instrumentalism.....	83
4.2.	Embedded Epistemic Instrumentalism’s Immunity	83
4.2.1.	Intuitive counterexamples and intuitive desiderata.....	84
4.2.2.	The master response.....	87
4.2.2.1.	The assumed explanation	88
4.2.2.2.	Alternative explanations	88
4.2.2.3.	A recipe for generating alternative explanations	90
4.2.3.	An example response	93
4.3.	The Not-All-Judgments Objection.....	95
4.4.	The Third-Factor Response Objection.....	98
5.	Two (Relevant) Interpretations of Embedded Epistemic Instrumentalism	100
5.1.	A Taxonomy of Normative Theories.....	100
5.1.1.	On the term ‘natural’	104
5.2.	Some Additions to Parfit’s’ Taxonomy.....	106

5.2.1.	Token non-reductivism and type non-reductivism.....	107
5.2.2.	A distinction between non-natural realist theories	108
5.3.	Two-theses Classification of Theories	111
5.4.	A Distinction Between Error Theories.....	113
5.5.	The Two Relevant Interpretations	114
5.5.1.	Embedded Epistemic Instrumentalism as non-analytic natural realism.....	114
5.5.2.	Embedded Epistemic Instrumentalism as an Error Theory	115
6.	Epistemic Error Theory and the Self-Undermining Objection.....	117
6.1.	Introduction To the Self-Undermining Objection.....	117
6.2.	Epistemic Error Theory	118
6.2.1.	My formulation.....	118
6.2.2.	Other formulations.....	120
6.2.2.1.	Cuneo's.....	121
6.2.2.2.	Streumer's	123
6.2.2.3.	Husi's	124
6.2.2.4.	Case's	125
6.2.2.5.	Nagel's	125
6.2.3.	Summary of terms.....	126
6.3.	The Self-Undermining Objection	127
6.3.1.	The general version.....	128
6.3.2.	Cuneo's versions	130
6.3.3.	Other versions	131
6.4.	The Not-So-Bad Reply.....	132
6.5.	Case's Retorts.....	135
6.5.1.	There is no non-reasons-suggesting alternative.....	135
6.5.2.	The non-reasons implying alternative doesn't explain the practice.....	138
6.6.	Denying Spirits Exist	139
6.7.	A Deep Flaw in the Self-Undermining Objection.....	142
7.	Normative Metaphysics is Unsatisfying	144
7.1.	A Difference Intuition.....	144
7.2.	There is No Satisfying Metaphysics of Normativity (for me).....	146
7.2.1.	Semantic, conceptual or constitutive-driven metaphysical accounts	146
7.2.2.	Metaphysically extravagant theories	151
7.2.3.	The No-Metaphysics-Feels-Right Conjecture	153

1. Introduction

1.1. A Moment of Weakness

I want to tell you a story. I do not claim that the story is true—memory has a way of turning a random series of dice rolls into a narrative. But it is close to the truth. I am going to go over some philosophy in the story. But I will go over it again properly, more accurately and more soberly, in the following sections and chapters. Here, I will just gesture at the things you need to know for the story to make sense, giving a vague impression rather than a sharp picture.

There was a comment on my honours thesis that stuck with me. For those not familiar with the idiosyncrasies of Australian undergraduate degrees, an honours thesis is part of an optional extra year of study, where students can dip their toe into the turbulent waters of research. Mine was a ten-to-fifteen-thousand-word thesis on why *burdens of proof* do not belong in philosophical practice. As part of that paper, I defined an ‘epistemic reason’ as ‘whatever gets you to truth’ (a view that is very different to the one I offer in this thesis). For this story, think of an epistemic reason as the normative pressure to believe in certain ways. The thought was that whatever leads you to the truth creates normative pressure for you to believe in a certain way. The comment on my thesis was that this conception of epistemic reasons, of epistemic normativity more generally, was under-motivated. The comment stuck with me because I agreed but I did not know what else to say.

I started my PhD program about a year later. The original project, the ancestor of this project, was about the different ways one can come to accept a theory. It seemed to me that full belief was only one of a vast array of subtly different attitudes one could have, and that once one realises this, a lot of skeptical problems are dissolved. But, much to the chagrin of my supervisors, I am interested in everything and I cannot let anything go. So, like an academic dilettante, I thought about many different things that were not part of my thesis. In particular, I could not let that comment about my conception of epistemic reasons from my honours thesis go. It made me vaguely uncomfortable, like a sore tooth; I could not stop toying with it, like a sore tooth.

A few weeks after starting my PhD, I came across an influential paper: Thomas Kelly’s ‘Epistemic Rationality as Instrumental Rationality: A Critique’ (2003). The first thing that paper taught me is that in suggesting that ‘epistemic reasons are whatever helps you get to the truth’, I was offering an instrumental conception of epistemic rationality. Instrumental rationality is means-end rationality. If one wants to live a long time, one should eat lots of vegetables. *Eating lots of vegetables*

is one way to increase one's chances of living a long time. *Eating lots of vegetables* is a means to one's ends. By suggesting that epistemic reasons where whatever means that gets one to the end of true belief, I was characterizing epistemic rationality as a type of instrumental rationality.

The second thing that the paper taught me is the main motivation to reject an instrumental conception of epistemic rationality, what came to be known as **too-few-reasons cases**. Instrumental reasons can be fickle. You have instrumental reasons to eat lots of vegetables because it promotes your health. But, sometimes, people do not care about being healthy. Perhaps they are nihilistic. Perhaps they have other priorities. Perhaps their lives are so full of misery.... But if one does not care about being healthy (and let's pretend that vegetables are not delicious in their own right) then it is not obvious that one has any reasons to eat lots of vegetables. Instrumental reasons can blink in and out of existence with the whim of people's psychologies. Epistemic reasons, so philosophers seem to judge, are never fickle. You always have them.

A problem for an instrumental conception of epistemic rationality is that there will be cases where people do not care about believing truly, and where, at least on first pass, none of their other goals are promoted by believing truly. In these cases, at least on first pass, there will not be any instrumental reasons to believe truly. But, so philosophers appear to judge, people always have epistemic reasons. So epistemic reasons are not instrumental reasons.

Let me quote Charles Côté-Bouchard's summary of one of Kelly's too-few-reasons cases (2015, p. 340):

Spoiler Alert Max missed the finale of his favourite television series, which aired last night and revealed whether character X ends up dying or surviving. Given his passion for the show, Max really wants to avoid learning the answer to that question before he watches the rerun tonight after work. But to his dismay, one of his colleagues—who is convinced that Max did watch the show last night—comes up to him in the morning shouting: 'I knew it! I told you character X would die!'

The standard interpretation of this case is that Max only has instrumental reasons to believe a falsehood, believe against his evidence, but intuitively he still has epistemic reasons to believe the truth, to believe in proportion with his evidence. So epistemic reasons cannot be instrumental reasons. Max has epistemic reasons to believe in proportion with his evidence but no instrumental reasons to believe in proportion with his evidence. In this sense, there are too few instrumental reasons to properly account for all the epistemic reasons, at least on first pass.

About a month or so after I started my PhD, my father was diagnosed with a terminal case of cancer. His body was riddled with it. And there was no significant chance that he would survive the disease. So I paused my PhD program, travelled to my hometown, and looked after him as he rapidly approached non-existence. While looking after him, I lived through a too-few-reasons case.

My father's imminent demise scared him, on occasion. I had given him some short lectures on the Epicurus' view of death.¹ It did not help him much. Sometimes he would get really scared, truly terrified. I recall sitting in his room. He was on the end of his large bed, in the middle of a panic attack. He was rocking back and forth. After about thirty seconds of the most intense panic, he started repeating to himself: "I will be okay. I will be okay. I will be okay...." Just to be clear, he was not going to be okay and it was obvious. But from the outside (I did not press him on it) he seemed to believe that he would be okay for a few moments. Just enough to make the panic ease. After a minute or two he was back to normal. I guess that if I had asked him about his death at this time, he would have accepted that he was going to die soon.

This was a better too-few-reasons case than the current ones in the literature. Firstly, my father had a competent and vigilant guardian looking after him. If he made any dumb decisions based on the false belief, I could intervene. So the chance that a false belief would interfere with his other goals was especially low. Secondly, he only flouted the epistemic standards, believed against the evidence, for a few quiet moments in his room. The chance that any important decision would have to be made in those few minutes was incredibly low. So, again, the chance that a false belief would interfere with his other goals was especially low. Without involving gods or unrealistic psychologies, it is hard to imagine a case where a false belief had less chance of thwarting any goals an agent might have.

If any moment would show the deficiencies of an instrumental conception of epistemic rationality, it was this moment. But two important things strike me when I reflect on it. The first is that, despite this being as close to a perfect too-few-reasons case as one could get in the actual world, my father had an abundance of instrumental reasons to conform to the epistemic standards. He was an especially proud man and especially proud of his intellect. He was the type of smart that was obvious within a few minutes of talking to him. And he loved that fact about himself. He

¹ My, what I am sure is an imperfect, understanding of the view: if you are around, you are not dead; if you are dead, you are not around; so you should never worry about being dead; the only times you could worry about it, you are not dead!

would hate to think that he believed against the evidence merely to make himself feel better. And he would hate for other people to know he believed this way. In addition, he hated to be patronised, pitied. He hated being vulnerable, having to depend on other people. So a plethora of his goals were thwarted in the few brief moments that he believed against the evidence. There were more than enough instrumental reasons in this too-few-reasons case.

The second is that, while I can note that he believed against the evidence, that he believed falsely, I do not feel any pressure to offer rebuke, blame or criticise his actions. I can note that his belief was not true, that it lacked evidential support, but this note lacks any real normative sting. I do not think he did anything *wrong*. When conforming to the epistemic standards has no instrumental benefit, when believing against the evidence has no practical effect, I cannot see any genuine normative part of the world compelling, urging, indicating that one really *ought* to conform, ought to believe with the evidence. If it has no other effect, it just does not matter what you believe.

The two reflections are the basic elements of the epistemic instrumentalism that I characterise and defend in this thesis. I wonder if I believe them merely because I am trying to process an emotionally intense moment, because I am trying to offer an apology for what will be interpreted as a moment of weakness from my father. Perhaps. Indeed, I suspect this moment is only one among a thick soup of events that *caused* me to believe epistemic instrumentalism. But this is all in the context of discovery. This is a part of what led me to epistemic instrumentalism.

For the rest of the thesis, let me switch to the context of justification. I aim to convince you, to convince the philosophical community, that epistemic instrumentalism is true, or at least that it is close to the truth. If I fail, I hope this thesis at least furthers the debate. I hope it clarifies some issues, states an instrumental position so clearly that the problems with it are all the more obvious. I will be satisfied if I contribute, even in a small way, to philosophical progress on understanding epistemic normativity.

So let us begin.

1.2. Identifying The Project

1.2.1. The question

You ought not to believe ‘the earth is flat’. Why?

The first thought many people have when I put this question to them is that the earth is in fact not flat, so that is why you ought not to have the belief ‘the earth is flat’. Importantly, I am not asking you to give me some standard that the belief flouts. There are many potential answers that explicitly or implicitly refer to some standard: you ought not to believe ‘the earth is flat’ because ‘the earth is flat’ is not true; it is not supported by the evidence;² it is not supported by your other beliefs;³ it is not produced by a reliable belief-forming process;⁴ it is not derivable from some secure foundation;⁵ you do not know it.⁶ You might hold that only one of the above is the correct answer or that some proper sub-set are correct answers or that some other answer is correct. But all these answers are driving at the same idea. The idea is about what I call the **epistemic standards** (more on this below). The idea is that the epistemic standards, whatever they are, demand that you not believe ‘the earth is flat’. Any answer that refers to a standard provokes another question: why believe in accordance with that standard?

How does a belief being false (or having any property) make it such that you ought not have that belief? How is reality set-up such that beliefs (or credences, or acceptances or endorsements or any doxastic state) can be wrong to have (or bad or inappropriate or incorrect or whatever term you want to use to denote that you ought not to have it)? Whatever terms you want to use, whatever philosophical framework you want to work in, how does reality tell *against* having specific representations and, for that matter, how does reality tell *for* having specific representations? This is the question I aim to answer in this thesis.

1.2.2. The area

We are engaging in metaphysics. More specifically, we are engaging in what some call ‘meta-epistemology’. Meta-epistemology is the epistemic equivalent of the more familiar meta-ethics. Where meta-ethics asks questions about the metaphysics of moral normativity (or normativity more generally) we are asking questions about the metaphysics of **epistemic normativity** (more on this below, too). In answering why we should have beliefs (representations or what have you) that conform to certain standards, I take myself to be giving an account of epistemic normativity. I take myself to be offering a description about what reality is actually like and when I argue for that description, I take myself to be offering evidence that the description is accurate.

² As an evidentialist holds: e.g. (Conee & Feldman, 2004)

³ As a coherentist holds: e.g. (Quine & Ullian, 1970)

⁴ As a reliabilist holds: e.g. (Goldman, 1967)

⁵ As a foundationalist holds: e.g. (Rene Descartes, 'Meditations on First Philosophy' in Cottingham, Stoothoff, & Murdoch, 1985, pp. 1-62).

⁶ As knowledge-first advocates hold: e.g. (Williamson, 2000)

This is important because the target of my investigation is not a concept or the meaning of some term in English. To this end, I interpret appeals to intuitions as appeals to evidence, in line with Kelly's description of their use in (2005). However, I am open to different ways appeals to intuition may be evidence. They may be evidence via our understanding of the meaning of our terms or our concepts. They may be evidence of how we represent reality and that representation itself may be evidence for how things are. That being said, they may be evidence in some other way (Kornblith, 2007; Williamson, 2007). I am ecumenical about how intuitions are evidence.

This is not to *insist* that there is only one project in philosophy or one use of intuitions. I am also ecumenical about aims in philosophy. I see the different projects in philosophy like the difference between writing poetry and describing a tree. Many poems contain assertions about trees. Many of those assertions are false. They do not correspond to reality. But I do not consider this a mark against poems or a mark against accurate descriptions of trees. Ditto for different aims within philosophy. If your aim in doing philosophy is to tell a story that conforms to as many of our intuitions as possible while being internally consistent, then I just consider that a different project to mine.

Indeed a similar point can be made about the notion of truth I am working with. I am working with a correspondence notion of truth. When I claim that epistemic instrumentalism is true (more humbly, the theory is approximately true or the model is fairly accurate) I mean it in a correspondence way. This notion of truth is not taken for granted in philosophy. Some philosophers hold that normative domains operate with a different notion of truth. Thomas Scanlon (2014) and Derek Parfit (2011) have been interpreted as offering this kind of truth pluralism. Herman Veluwenkamp (2017, p. 751) cites both:

Both philosophers argue that there are domains of discourse in which true judgements do not have ontological implications. In the scientific domain true judgements are about reality and these judgements have ontological implications. By contrast, Parfit maintains that in the normative domain these truths have “no positive ontological implications” and are not “about metaphysical reality” (2011 Vol. II, p. 479 & 747). And for Scanlon, normative truths “need no natural or special metaphysical reality in order to have the significance that we commonly grant them” (2014, 52).

Veluwenkamp's interpretation of these claims is that 'truth' for normative discourse is different to 'truth' in scientific discourse (I return to this point several times in this thesis). I am suspicious that such a claim is accurate about normative discourse. But the official position I take here is, yet again,

an ecumenical one. Maybe, to my surprise, all or most metaphysicians of normativity operate with a non-correspondence notion of truth. If that is the case, this thesis is just a different theoretical task to what metaphysicians usually engage in. When it comes to epistemic normativity, what does correspond to reality? With the scientific sense of truth in mind, why should you not believe ‘the earth is flat’? If theories of normativity like Scanlon’s and Parfit’s really are ontologically uncommitted, then it is not clear that anything they say is relevant to the answer to my questions, with the scientific notion of truth in mind. It is as if one of my claims is inconsistent with the assertions of a poem. It is not a mark against either. So for most of this thesis, I just treat these enterprises as different tasks.

I want to be clear that my ecumenical position is genuine. I am not trying to insinuate that there is *really* only one project going on in philosophy, and it is the one continuous with the sciences. I doubt that, if surveyed, the majority of philosophers would be aware there is a different notion of truth operating in philosophy or metaphysics; I suspect most philosophers would reject such a characterisation of their own work. But that is irrelevant. The projects and the aims of our projects are up to us. I want to know the properties of epistemic normativity that correspond to reality, in a scientific-like way. Other philosophers might have different projects in mind.

Nonetheless, in chapter 7, I will relax my focus on epistemic normativity in a scientific-like way. In that chapter, I explain how I integrate what I take to be true into my worldview. What I say there has implications for non-ontologically committed claims about normativity. That being said, for most of this thesis, I take philosophical investigations with aims other than a scientific-like understanding of our world, with a notion of truth other than a scientific-like one, to be orthogonal to my investigation.

1.3. Defining Important Concepts

To further clarify the topic, and to facilitate discussion, I will define several concepts in this subsection. Importantly, I never want any argument in this thesis to depend on a particular choice of terms. If my interlocutors want a term, they can have it! Two chapters of the thesis, chapters 5 and 6, are the result of this semantic and conceptual ecumenicalism.⁷ In those chapters, I offer two ways to interpret the theory I put forward in this thesis, depending on the semantics or conceptual commitments you find yourself with. In this section, I take myself to just be *keeping the books* in one

⁷ Indeed, a theme in this thesis is that the version of epistemic instrumentalism I argue for requires few other commitments, whether those commitments are about semantics or other topics in philosophy.

particular way, among many ways we might do it. The key terms I will define are **epistemic standards**; **epistemic normativity**; **epistemic reasons**, and **epistemic**.

1.3.1. Epistemic standards

I want to first discuss the concept ‘epistemic standards.’ You should believe in proportion to your evidence. Suppose you are thirsty and you believe there is a glass of water in front of you. If you suddenly receive evidence that the glass is not full of water, but poison, you should update your beliefs in such a way that respects that evidence. ‘Believe in proportion to your evidence’ is an example of an epistemic standard. As I stated above, other epistemic standards have been proposed: ‘believe coherently’, ‘have beliefs caused by a reliable process’, and ‘believe only what you know’, among others. A central concern in epistemology is describing, defending and rejecting various purported epistemic standards. In this thesis, I am not taking a position about what the correct epistemic standards are. I intend everything I say in this thesis to be compatible with every reasonable way of filling-in the details about the epistemic standards. So let’s call the standards that philosophers are aiming to state, **the epistemic standards**. I leave it to the reader to fill in the details about what those standards actually are. When called for, I will use ‘believe in proportion to your evidence’ as the example of an epistemic standard, although I invite the reader to substitute their preferred epistemic standards in the relevant examples.

1.3.2. Epistemic normativity

The epistemic standards, whatever they are, have a certain normative force. Compare ‘believe in proportion to your evidence’ to a standard I selected on a whim: ‘have beliefs only about blue objects.’ Today you will attempt to believe in proportion to your evidence. Indeed, you might expend considerable energy doing so. You might help other people conform to the epistemic standards. You might rebuke other people for not conforming. You might even feel guilty for not conforming to the standards, yourself. But almost no one will even try to have beliefs only about blue objects, help others to do so, rebuke others for flouting the blue standard nor feel guilty for flouting it themselves. The epistemic standards have a pull. They compel certain behaviour in a way that arbitrarily selected standards do not.

Similarly, the epistemic standards are like the standards of etiquette or the standards of chess, but they compel much more often and much more forcefully than do the standards of etiquette and chess. For non-professional chess players, not conforming to the chess standards, at worst, opens one up to playful criticism: “so you prefer checkers?” said with a sardonic tone. But not

conforming to the epistemic standards is in some cases met with serious condemnation. In other cases, it can be grounds for medical intervention and incarceration. The epistemic standards compel in a way many other standards do not. **Epistemic normativity** is the thing, whatever it is, that makes the epistemic standards compel.

1.3.3. Epistemic reasons

In the relevant literature, philosophers tend to talk about epistemic normativity by using the term ‘epistemic reasons’.⁸ I follow this convention although I do not intend anything of substance to depend on this choice of terms. I could make my point equally well with talk of ‘value’ or ‘goodness’ or any other term that attempts to refer to the genuinely normative. But in reasons-talk, we can say the epistemic standards produce epistemic reasons in a way that ‘have beliefs only about blue objects’ does not.

So we can rephrase the meta-epistemic question at the heart of this thesis: How do the epistemic standards produce epistemic reasons?

1.3.4. Epistemic

‘Epistemic’ is a vague term. I use it to refer to the fact that the types of standards, reasons and normativity that I am talking about in this thesis have to do with beliefs, with knowledge, with truth, with evidence, with things like these, as opposed to types of standards, reasons and normativity that has to do with beauty or being kind; in this thesis, I care primarily about epistemic things as opposed aesthetic or moral things. Another contrast to the epistemic is the pragmatic. ‘The pragmatic’ is a term for things to do with reaching your ends. Importantly, I am not claiming the epistemic is actually different to the aesthetic, the moral or the pragmatic. Indeed, I am claiming that the epistemic reasons are actually a type of instrumental reasons; according to my story, it is partly because we have aesthetic and moral ends, or even just desires, that the epistemic standards get their normative force. Nonetheless, this is supposed to be somewhat surprising. I only care about the instrumental connections between the epistemic and the aesthetic, the moral, etcetera,

⁸ Although not everyone. Barry Maguire and Jack Woods (2020) hold that the epistemic standards only produce non-authoritative reasons. ‘Authoritative reasons’ are what I am calling ‘reasons’, used to denote genuinely normative considerations that speak for or against something. I am not exactly sure how to interpret ‘non-authoritative reasons’. On one reading of Maguire and Woods, ‘non-authoritative reasons’ are produced by all standards, including the standards of chess (Maguire & Woods, 2020, pp. 217-218). On this reading, non-authoritative reasons are just considerations according to some sense of correctness, according to a standard. So in Maguire and Woods terminology, this thesis is defending an instrumental answer to this question: ‘why do the epistemic standards produce authoritative reasons much more robustly than chess standards?’

instrumentally. In this thesis, I am first and foremost interested in standards, reasons and normativity to do with beliefs, with knowledge, with truth, with evidence, and with things like these.

But the term ‘epistemic’ has come under fire. Stewart Cohen argues that use of the term in philosophy is problematic (Cohen, 2016). His argument is that the term functions as a specialised technical term in philosophy but because it has no agreed upon definition, it is problematic in philosophical theorising. I agree with Cohen that ‘epistemic’ is not a term of standard English, and that there is no theory independent way of defining the term (Cohen, 2016, pp. 839-847). However, I do not hold that we need to jettison use of the term in philosophy.

I have two responses to Cohen’s position. Firstly, Cohen’s critique depends on a distinction between terms of natural language and technical terms (Cohen, 2016, p. 844). Terms of natural language, in this case English, are terms that a clever English speaker would understand if used in a sentence. Cohen holds that a clever English speaker would not understand the sentence: “Is this belief epistemically justified?” But, he holds that such a clever English speaker would understand the sentence “Is this belief rational?” For Cohen, technical terms are something else.

Cohen anticipates an objection:

Perhaps technical expressions can have more in common with natural language expressions than I am allowing. How do natural language expressions get their meaning? Certainly, a large part of the story is that they acquire their meaning by being used in a particular way by a community of speakers. Can we say that a technical term like ‘epistemic’ acquires its meaning by being used in a particular way by a very specialised community – the community of epistemologists? If so, then when epistemologists talk about epistemic justification, what they are talking about is determined by how they have been using that expression [footnotes removed] (Cohen, 2016, p. 849).

The objection is that perhaps the term ‘epistemic’ is like a natural language term, but for a smaller population than clever English speakers, perhaps epistemologists. Cohen rejects this possibility with two arguments. The first argument is:

If this account of how ‘epistemic’ acquires its [sic] meaning is correct, then it is hard to make sense of standard practices of epistemologists. It is not uncommon for an epistemologist to give a talk about the nature of epistemic justification before a general philosophy audience. But if the proposed analogy with natural language is correct, then just as one should not give talks in English to non-English speakers, so epistemologists should not be giving talks about epistemic justification to non-epistemologists. Since their

audience is not part of the linguistic community whose conventions of use give ‘epistemic’ its meaning, the audience will not know what they are talking about. Similarly, interdisciplinary-minded epistemologists should not be enlisting the aid of non-philosophers for understanding the nature of epistemic justification. A cognitive psychologist, for example, would be unable to understand what the target notion is. (Cohen, 2016, pp. 849-850)

The argument is that the behaviour of epistemologists provides evidence that ‘epistemic’ does not have its meaning in virtue of a similar mechanism to how natural language terms acquire meaning—but operating on the population of epistemologists rather than the population of English speakers. A problem with this argument is that it is not obvious that ‘epistemologists’ are the right population to focus on. If I am preparing a talk on epistemic reasons, I would expect an audience of philosophers to understand the distinction between epistemic and moral or aesthetic. However, if the talk was for legal scholars, say, I would not expect them to know the distinction. Of course, my expectations may be idiosyncratic. But if philosophers tend to have the expectations I have, if philosophers tend to expect other philosophers to understand an epistemic/aesthetic distinction but not expect legal scholars to understand such a distinction, then that is evidence that the term ‘epistemic’ is like a natural language term, but for philosophers as opposed to English speakers at large.⁹ At the very least, this is a possibility not properly addressed by Cohen.

His second argument against the idea that the term ‘epistemic’ is like a natural language term, but for a smaller population, is the following:

Perhaps, the most serious problem for this approach is that it allows that there is a fundamentally important normative notion that cannot be expressed in natural language. It would be very surprising for such a notion to be in this way ineffable. This would be akin to the word ‘moral’ not existing and there being no way to express what morality is. It is important to see that the point is not simply that there is no single natural language word that picks out the relevant property. According to the proposal we are considering, there is no way to use the resources of natural language to characterise it. If there were such an ineffable normative notion, it is not clear why we should care about it. Presumably, if it were an important notion, we would have developed a way to talk about it (Cohen, 2016, p. 850).

The idea is that natural language, in this case English, has all the expressions it needs to characterise all the normatively important notions. Cohen claims this but I find no motivation to believe it. I have two concerns. My first concern is that the phrase ‘cannot be expressed in natural language’ is vague and its potential precisifications do not support Cohen’s position. Give me long enough,

⁹ I doubt it is only philosophers that use the word. I suspect linguists also use it, for example.

only using English, I can teach a student to properly use the term ‘epistemic’ in a philosophical context. Perhaps ‘cannot be expressed in natural language’ is meant to imply that there is no sharp definition of the term ‘epistemic’. But this is true of a lot of terms; it is hard to give sharp definitions in English of ‘probability’, ‘life’, ‘mind’ and ‘good philosophy’. So not being amenable to sharp definition is not evidence that ‘epistemic’ acquired its meaning differently to other natural language terms. Indeed, if anything, it is evidence in favour of the idea that ‘epistemic’ has a similar development to other natural language terms. Perhaps ‘cannot be expressed in natural language’ is supposed to imply that there is no single phrase in natural language that is synonymous to ‘epistemic’. But a lot of academic investigations involve creating new concepts that are not synonymous to a single phrase in English, for example the concept ‘quantum entanglement’ and ‘probability’. I am not sure why we should expect normative investigations to be different. So, not only is the phrase ‘cannot be expressed in natural language’ vague but it is not clear what precisifications would make this property evidence that the term ‘epistemic’ did not acquire its meaning like a natural language term (but with a much smaller population than all English speakers).

My second concern is that there are counterexamples. I take the distinction between formal and substantive equality¹⁰ to be an important normative notion. It is hard to give a sharp definition of it that everyone will agree with. If this distinction is part of standard English, I would bet it is a recent addition. But give me long enough, only using English, I can teach a student to properly use the distinction. I hold similar views about the terms ‘instrumental value’, ‘intersectional analysis’ ‘sexual harassment’ and ‘probability’. Indeed, I suspect a lot of terms we take to be essential normative notions are more recent additions to English than we suspect, for example, the term ‘facts’:

How old is the notion of “facts” in Anglo culture? To many speakers of English, the question may seem strange. Haven’t “facts” always been there? Aren’t they simply a “fact of life,” a notion as necessary and fundamental as time and place? To these people the news might come as a shock that there are no “facts” (in the plural) at all in Shakespeare. The word fact in the singular does occur in Shakespeare’s works, but its meaning is different from that familiar from modern English discourse. The older meaning of fact, found *inter alia* in Shakespeare, is defined by the OED as “a thing done or performed”.... (Wierzbicka, 2006, p. 41).

¹⁰ To get the gist of this distinction: imagine that the penalty for a crime is a \$10 000 fine. The penalty is applied uniformly to all offenders. You might hold that this is *formally* equal because all offenders suffer the same punishment. But it is not *substantively* equal because rich offenders will barely notice the fine but poor offenders will face significant hardship.

Anna Wierzbicka argues that the term ‘facts’ as we use it today, was a technical philosophical innovation of the enlightenment, that found its way into everyday use (2006, p. 42). The term ‘facts’ is not an isolated incident. Even the term ‘reasonable’ in its modern usage is a much more recent addition to English than you might expect:

Clearly, between Shakespeare’s time and our own, a change has occurred in Anglo ways of thinking, and this change has found its reflection in the semantics of the English language. And it is not just a change in the prevailing conception of what is and what is not “reasonable.” Rather, it is a change in the basic kit of conceptual tools used for making judgements and evaluating actions and events. The word reasonable with its modern family of interrelated senses is a new tool in the kit (Wierzbicka, 2006, p. 105).

I am not in a position to assess the accuracy of Wierzbicka’s claims.¹¹ But I consider the idea that ‘many terms, even important normative notions, were cultural innovations’ very plausible. And I see no evidence that we are done innovating. So I reject the claim that it is implausible philosophers would identify an important normative notion not found in English (or any natural language, currently). Just as Cohen’s preferred term ‘reasonable’ was added to English at some point, so too may the term ‘epistemic’ be added. Someone had to start adding such notions to English and I doubt we have completed the project.

That being said, I am sympathetic to the worry at the heart of Cohen’s point. I take the central worry is that the term ‘epistemic’ is quite vague. So engaging in very specific theorising only led by our intuitions about what is and is not ‘epistemic’ does not look like a reliable method for producing accurate models. Where Cohen and I diverge is he implicitly is much more comfortable with using our intuitions about terms of standard English to engage in such theorising—he believes ‘rational’ is an acceptable term to guide philosophical theorising, for example (Cohen, 2016, p. 845). However, terms of natural language are also often vague. They also admit many precisifications. They also create a lot of confusion. And they often cause a lot of disagreement that results in irresolvable clashes of intuitions. Many of the very same problems Cohen points to when using the term ‘epistemic’ to guide theorising will also be present when using many natural language terms like ‘freedom’ ‘good’ ‘equality’ and even ‘rational’. The fact that these are terms of standard English, as opposed to some more idiosyncratic jargon, does not help if they are still vague and still lead philosophers into confusion.

¹¹ Importantly, I do not need to endorse the entire argument from Wierzbicka (2006). So long as the claims about *when* terms like ‘facts’ and ‘reasonable’ developed their modern meanings in English are broadly accurate, I can offer my objection to Cohen’s position.

This brings us to my second response to Cohen's position. The arguments in this thesis are not based on having some semantic competence or having some special rational insight into essences. I never argue that the meaning of some term (or concept or essence) is thus so the metaphysics must be thus. Cohen's worry is that the fact we lack a good grasp on 'epistemic' means our theorising is led astray. He may very well be correct about this. If one's arguments are full of intuitive reactions to cases about the classification 'epistemic' then that is motivation to worry. Arguments like: 'intuitively, practical concerns are not *epistemically* relevant; therefore....' are undermined by Cohen's concerns. But I am not arguing this way in this thesis. My arguments are as forceful as they are, no matter what terms one uses, no matter the actual meaning of words. This is why I am comfortable stipulating concepts. And why I am happy to give them up if we want to use different terms. So Cohen's worries about 'epistemic' do not apply to my project.

1.4. The Gist of My Position

With these concepts in place, I can now offer an overview of this thesis. We start with the question:

How do the epistemic standards produce epistemic reasons?

In chapter 2, I offer an answer. It is the potent and persistent usefulness of conforming to the epistemic standards that gives rise to epistemic reasons. I note some of the ways it is useful to conform to the standards. I create a model of this usefulness; I make an analogy between reasons to be healthy and reasons to conform to the epistemic standards. I defend this model from the central objection to epistemic instrumentalism. This model and the associated explanation of epistemic normativity and my response to the central objection constitutes my theory of epistemic normativity. I call this theory 'embedded epistemic instrumentalism'. I argue that my theory has advantages over close rivals. And then I explain how this theory interacts with other topics in epistemology, such as the characterisation of evidence.

In chapter 3, I criticise the standard form of epistemic instrumentalism, which I call 'universal epistemic instrumentalism'. This criticism is important because universal epistemic instrumentalism shares a lot of similar motivations with my account. However, I argue that the way universal epistemic instrumentalism responds to the central objection to epistemic instrumentalism creates very implausible theories. My aim is not merely to identify some objections to these theories but to bury the theories in well-nailed coffins.

In chapter 4, I then argue for embedded epistemic instrumentalism. Importantly, my argument is not an argument from exhaustion (cf. Huemer, 2005). I am not saying that all other answers have problems, so this is the answer we are left with. I suspect that there will be problems with all philosophical accounts of epistemic normativity, even mine. So the mere fact that I can identify some problems with other accounts does not make my account stand out from the field. Rather my strategy is to point out that there is an uncontroversial core to embedded epistemic instrumentalism, facts we will all agree on and facts that are integrated well with other things we know. These facts are the base of embedded epistemic instrumentalism. I then argue that embedded epistemic instrumentalism has a recipe for responding to the main source of objections to it, intuitive arguments. The result is that we have an extremely well-motivated theory, that is immune from the main source of criticisms, or so I argue. These properties make embedded epistemic instrumentalism stand out from the crowd.

In chapter 5, I offer two relevant interpretations of embedded epistemic instrumentalism, error theory and non-analytic natural realism. I explore the implications of such interpretations and how they relate to other theories. I conclude that the interpretations amount to similar claims about the world. The only difference is how one thinks the metaphysics is related to the semantics.

In chapter 6, I defend the error theory interpretation from a charge that epistemic error theories are self-undermining. This is important because one of the reasons I can be ecumenical about certain semantic facts is that I am happy for philosophers to interpret embedded epistemic instrumentalism as an error theory. So this chapter establishes that such an interpretation does not undermine the evidential value of the arguments I make in favour of embedded epistemic instrumentalism.

In chapter 7, I defend my theory from the charge that I have failed to identify any actual normativity in my theory of epistemic normativity. My claim is that every metaphysical theory of normativity is open to the same charge. So it is not a damning criticism.

2. Embedded Epistemic Instrumentalism

Conforming to the epistemic standards is an extremely useful undertaking. It is useful in many ways and in many contexts. Conforming is *robustly* useful. The fact that conforming is robustly useful explains the normativity of epistemic standards. This is embedded epistemic instrumentalism in a nutshell. The rest of the thesis explains, expands, and defends this position.

The theory that usefulness explains epistemic normativity is called **epistemic instrumentalism**. Epistemic instrumentalism is familiar (Côté-Bouchard, 2015; Dyke, 2021; Foley, 1987, 1993; Giere, 1989; Kitcher, 1992; Kornblith, 1993; Laudan, 1986, 1990a, 1990b; Leite, 2007; Nozick, 1993; Quine, 1969; Schroeder, 2007)¹² but not popular. We can lay at least part of the blame for this lack of popularity on the standard way epistemic instrumentalism is developed. Standard developments have two flaws: first, they assert that the epistemic standards exert normative force always and everywhere, that epistemic standards have universal application. This **universal epistemic instrumentalism** is not very plausible;¹³ second, they focus on the way the epistemic standards can promote a particular end, the end chosen varying with the account. One proposal is that people have a universal end that is always promoted by conforming to the epistemic standards (Côté-Bouchard, 2015; Foley, 1987). Another proposal is that any arbitrarily chosen end will always be promoted by conforming to the epistemic standards (Kornblith, 1993; Schroeder, 2007). However, the strongest form of epistemic instrumentalism does not focus on a supposed universal end, or an arbitrarily chosen end, or even any particular way such an end is promoted by conforming to the epistemic standards. Rather, the strongest form of epistemic instrumentalism focuses on the plurality of ends and the plurality of ways conforming to the epistemic standards promotes those ends.

As I noted above, the development of epistemic instrumentalism was heavily influenced by Kelly (2003), offering what came to be known as the **too-few-reasons objection**.¹⁴ Again, the gist of the objection is that philosophers judge that the epistemic standards have normative force in response to all proposed cases, even in response to cases where conforming to the epistemic standards does not appear to promote any goals. Philosophers judge as if the normative force of

¹² The details of my account differ, sometimes dramatically, from previous instrumental accounts.

¹³ I criticise this type of epistemic instrumentalism in chapter 3.

¹⁴ Also see: (Kelly, 2007). For a similar argument see: (Siegel, 1989, 2019). The ‘too few reasons’ objection was coined by Schroeder (2007, chapter 5) and followed by Côté-Bouchard (2015). Sharadin calls it the ‘Universality Challenge’ (Sharadin, 2018, p. 3793).

epistemic standards is universal.¹⁵ However, according to epistemic instrumentalism, epistemic standards only have force when a person has a relevant end. At least on first pass, ends are not universal. Therefore, instrumentalists cannot say the epistemic standards have universal force, implying that the theory cannot be correct. In response, instrumentalists try to shoehorn instrumental reasons into playing a universal role, claiming that one way or another epistemic instrumentalism *does* deliver universal normativity. This is the wrong response and yet it is this response that has shaped the standard development of epistemic instrumentalism. My account is much simpler and offers a much more compelling response to the too-few-reasons objection, or so I will argue.

There are many theoretical motivations for holding an instrumental account. While I will present the case for embedded epistemic instrumentalism in more detail in chapter 4, I will make some brief comments here. As Michelle M. Dyke asserts, instrumental accounts avoid the host of difficulties associated with *sui generis* normative accounts (Dyke, 2021, p. 2). Epistemic instrumentalism does not require any dubious metaphysics or epistemology, nor does the theory require fundamental normative properties or any epistemic miracles (cf. Enoch, 2011, pp. 172-173). As Kelly argues, these accounts conform with the loose collection of motivations dubbed ‘naturalism’ (Kelly, 2003, pp. 614-616). Epistemic instrumentalism coheres with scientific practice in that everything it postulates can be studied scientifically. As Hilary Kornblith explains, these accounts are not just theories engaging in mere semantics; they make substantial claims about the world (Kornblith, 1993, pp. 359-363; also see section 7.2.1). Regardless of what you call it, the usefulness of conforming to the epistemic standards is an important fact about the world that should be studied (I return to this point in chapter 7). This is not something that only matters because we happen to refer to it with the term ‘epistemic normativity’. If you are a philosopher who is motivated by such things, then this style of theory will appeal to you. Even ignoring the theoretical motivations, my version of epistemic instrumentalism is based on a few uncontroversial facts and is able to explain what it needs to explain. As such, I hold it is a compelling answer to the question: why conform to the epistemic standards?

The plan: in section 2.1, I identify some of the ways that conforming to the epistemic standards is useful; in section 2.2, I argue that this usefulness explains the normative force of epistemic standards; in section 2.3 I respond to the too-few-reasons objection; in section 2.4, I articulate

¹⁵ In my experience, philosophers do judge in this way. However, it would be worth checking experimentally just how uniform this judgement is. For this paper, I am happy to concede that philosophers judge this way. If they don’t have too-few-reasons judgements, all the better for my theory.

some advantages of this account over its competitors; in section 2.5, I gesture at how embedded epistemic instrumentalism relates to other topics in epistemology.

2.1. Conforming is Useful

In this section I characterise certain ways that conforming to the epistemic standards is useful. My understanding of the ways conforming is useful is strongly influenced by previous work on epistemic instrumentalism. However, while previous instrumental theories tend to focus on one specific way that conforming is useful (attempting to respond to the too-few-reasons objection) my theory holds that conforming is useful in many related, though somewhat distinct, ways. In a sense, embedded epistemic instrumentalism is a disjunctive theory.

First, conforming promotes having accurate representations of the world. Accurate representations help with almost everything you do in life. Want to write? It helps to know the locations of the relevant computer keys. Want to eat? It helps to have true beliefs about what is nutritious, and more importantly, what tastes great. Want to get a promotion? It helps to have accurate credences about the chances that currying favour with your boss will pay off. Conforming helps you accurately represent the world. Accurate representations help you do almost everything you want to do. This is the basic way conforming is useful. Broadly speaking, you could subsume all the different ways conforming is useful identified below in this basic way. As such what follows is not a series of distinct ways that conforming is useful, but rather a series of varying points of emphasis, demonstrating the broad usefulness of conforming. It is a mistake to only focus on one, missing the forest for the trees.

Second, conforming is useful because there is a chance that accuracy will matter in the future. Mark Schroeder bases his theory on this fact (Schroeder, 2007, chapter 6).¹⁶ Odds are you do not currently care what the state beverage of Delaware is, but at some point in the future it might be a question on a quiz show, with the correct answer winning you a million dollars. There is always a chance that accuracy will matter. So rather than directly helping now, the chance that accurate beliefs will help in the future provides instrumental reasons to conform. To build on Schroeder's point: typically, the cost of conforming is small. In many instances, having beliefs that conform takes less effort than having non-conforming beliefs (try to believe you are not reading now, for example). Furthermore, we are indifferent to the content of many beliefs so we may as well

¹⁶ For criticism see: (Côté-Bouchard, 2015, pp. 343-345; Sharadin, 2018, pp. 3796-3801) and section 3.2.

conform. This results in a wide variety of cases where you ought to conform because conforming might help in the future, and the costs of conforming are negligible.

Third, conforming now can help you conform in the future. Hilary Kornblith and Adam Leite develop instrumental theories based on the idea that conforming to the epistemic standards as a rule is important (Kornblith, 1993; Leite, 2007).¹⁷ They are thinking that in many cases it is important to conform. But in some cases, we may have no direct instrumental reasons to conform. Their explanation for why it is important to conform even in cases where nothing much is at stake is that being *able* to conform in the future requires practice. If you are in the habit of flouting the standards, you may not be able to conform when stakes are high. As such, it is better to practice when nothing else is at stake so that when you need to conform, you are able to conform.

Fourth, you have goals that require conforming. Some authors have suggested that all people share some common goal. Côté-Bouchard argues that everyone necessarily has the goal of acting on the basis of genuine reasons and this requires that you conform as a rule (2015, pp. 350-352).¹⁸ Richard Foley asserts that people may simply prefer conforming (1987, p. 11).¹⁹ While I doubt the universality of any single goal, there are some reasonably common goals that demand conformance. For example, many people want others to trust their judgement and they aspire to deserve that trust. Such a goal requires people to conform to the epistemic standards. If people regularly flout the epistemic standards, then they are not worthy of being trusted. Indeed, the general usefulness of conforming and the fact that many common goals require it may be developmentally connected. Conforming is so useful that we can imagine natural evolutionary processes, be they biological or cultural, furnishing us with an appetite for conforming the way it furnished us with an appetite for calories. So, while I doubt the universality of Côté-Bouchard's and Foley's claims, there are some fairly common goals that require conforming to the epistemic standards.

Fifth, many social elements of our world require conformance or systematic deception. As the level of deception required is beyond most people (or they simply do not want to be deceptive) they are required to conform to achieve their social goals. For example, many people want to act in a way that makes those around them comfortable. The more you flout epistemic standards, the more people will feel uncomfortable. Imagine someone refusing to believe that there was a cup in

¹⁷ For criticism see: (Côté-Bouchard, 2015, pp. 348-450; Lockard, 2013, pp. 1712-1716) and section 3.3.

¹⁸ For criticism see: (Sharadin, 2018, pp. 3794-3796) and section 3.5.

¹⁹ Also see (Sosa, 2003, p. 157). For criticism see: section 3.5.1.

front of them when it is clearly visible or placing their hand in a fire because they do not believe it will hurt this time, etc. Perhaps there is some genius liar who could flout the standards while convincingly pretending otherwise, but for most of us, the only means we have to make those around us comfortable is to conform. I concede that there are some subjects such as politics, religion, or public scandals, where one can ignore epistemic standards with seemingly little repercussion but start denying there are cars on the road and your life will change very quickly. Similarly, people do not want to be patronised; they do not want to be ignored; they want to be taken seriously. These social goals often require conforming to the epistemic standards.

Sixth, some people are part of groups with epistemic goals that require conformance. Michelle M. Dyke holds that the fact people are part of communities that have epistemic goals is the basis of epistemic normativity (Dyke, 2021).²⁰ While I reject the theory overall, I agree that communities have goals, and some people want to further those goals. For example, I am part of the philosophical community. This community is trying to understand the world. I want to help. This requires me to conform to the epistemic standards so that I can communicate and convince other members of the community, to help make progress. There are many communities like this: businesses, not-for-profit organizations, governments, etc. Many people want to help their communities and conforming is required to achieve some community goals.

In summary, there are at least six related ways conforming is useful. I suspect there are more but I take it these six ways are relatively uncontroversial. Moreover, they provide instrumental reasons in a lot of cases so even if these are the only ways conforming is useful, they are enough to support my theory. Importantly, we begin to see one way that the standard developments of epistemic instrumentalism are deficient. By focusing only on one way conforming is useful, they miss the plurality of other ways. We can utilise this plurality to deliver a more compelling version of epistemic instrumentalism.

2.2. The Embedded Goal Model

I have identified some ways conforming is useful. But this is not much of a theory, as it is not clear how the plurality of ways conforming is useful explains epistemic normativity. In particular, on first pass, instrumental reasons seem unlike epistemic reasons. Typically, we think of instrumental reasons as fickle. Sometimes you have instrumental reasons to go for a walk. It might, right now,

²⁰ For criticism see section 2.4.

help achieve some explicit goal or valuable end. But in our complex environment whether you have reasons to go for a walk will vary over time. Right now you might be stressed and have reasons to walk, while later tonight, when you are trying to fall asleep, you will not have those reasons. This fickleness is normal for instrumental reasons, but it is a problem for an instrumental account of epistemic normativity. Typically, we think of epistemic reasons as robust. If you see a cat on a mat, then you have epistemic reasons to believe there is a cat on the mat, regardless of how tired you are, whether you are stressed or, at least on first pass, your goals. Typical instrumental reasons are much more sensitive to context than typical epistemic reasons. A challenge for the instrumentalist is to use a theory that typically produces fickle reasons (instrumentalism) to account for a phenomenon that produces robust reasons (epistemic standards).

In this section, I will present a version of epistemic instrumentalism that does just that. I will present my theory as a version of pragmatic epistemic instrumentalism, concerned with achieving your explicit goals. However, I will have something to say about how to modify this account for other versions of epistemic instrumentalism as well. The central move is to posit a type of intermediary goal—a goal that will help you achieve a number of other goals (or valuable ends), whatever they might be. This intermediary goal is relatively stable, so even as your other goals (or valuable ends) change, the intermediary goal stays the same. This intermediary goal gives rise to robust epistemic reasons.

My account offers a different explanation for the stability of epistemic reasons to the standard version. Standard accounts aim to identify some universal relation between conforming to the epistemic standards and promoting at least one of your goals. The universality of the relation explains the robustness of epistemic reasons. In contrast, my account focuses on the aggregate. We are shifting from a particle physics-like focus to a statistical mechanics-like one. The robustness of epistemic reasons does not come from specific interactions between instrumental reasons and goals, but from the robustness of the aggregate of many such interactions. While the difference is subtle, I am offering an alternative explanation of epistemic normativity to the standard account. I am claiming that the robustness of epistemic normativity comes from the forest, not the trees.

To introduce the embedded goal model of epistemic normativity, I employ an extended analogy. The instrumentalist can easily account for less mysterious normative phenomena, like the reasons

you have to be healthy. So, initially, I will build an instrumental model of health normativity. Then I will apply this model to the case of epistemic normativity.

2.2.1. The Embedded Health Goal Model

You should eat vegetables. You should exercise. You should relax. You should do these things because they are healthy. You have health reasons to do them. If you do not do these things, you increase your chances of having health problems—which is bad. It is not bad in some all things considered way, or some mysterious abstract way. Rather, being healthy is crucial for achieving most of the goals you have in life.

You might have explicit health goals: think of my-body-is-a-temple people. Those people place great importance on being healthy for its own sake. Being unhealthy hinders the achievement of their explicit health goals. This is one way it is bad to be unhealthy.

However, being unhealthy is bad regardless of whether you have explicit health goals or not. You have all sorts of other goals that require, or are greatly promoted by, being healthy. You may want to achieve big things in a career. Being healthy helps. You may want a loving relationship. Being healthy helps there too. Maybe you just want to bowl ripping leg-spinners or dunk a basketball or drive a golf ball or run long distances or dance the lambada. Yet again, being healthy helps. You may or may not have the explicit goal, but it is still in your interest to be healthy. Let's call being healthy **an embedded goal**, because achieving it promotes the achievement of a variety of other goals.

Embedded goals give rise to more robust reasons than their explicit counterparts. Explicit goals depend on facts about human psychology, which is rather fickle. For example, when I get home from work, I sometimes want to watch TV while at other times I want to go for a walk. The goal of going for a walk creates reasons to do other things, such as reasons to put on sunscreen (so I do not get sunburnt). These derivative reasons are fickle. They only exist while I have the goal of going for a walk, and whether I have that goal or not is incredibly variable. Of course, it is a matter of degree, but reasons based on explicit goals typically inherit the underlying fickleness of human psychology. Contrast this with embedded goals, which depend on more stable facts: facts about what it takes to achieve many different goals. For example, even when I have no desire to go for a walk, I still have reasons to go for a walk based on the embedded health goal. Walking is good

for my health and my health is good for achieving many different things. Reasons based on embedded goals are less fickle than their explicit counterparts.

Embeddedness comes in degrees.²¹ Consider the following two types of embedded goals. Most modern people have the embedded goal of earning money. It is useful for achieving any number of other goals. However, we can imagine some prehistoric woman, striking stones together to make a knife. She has no reasons to earn money because the institutions are not in place that would make money useful. However, she has a closely associated embedded goal, the goal of gathering resources. Indeed, almost any creature that has goals, needs to gather resources to achieve those goals. So ‘gathering resources’ is a more embedded goal than ‘making money’ because it exists in all contexts where the goal ‘making money’ exists, and additional contexts as well.

A goal can have a high degree of embeddedness, but still not give rise to reasons always and everywhere. Consider the well-embedded health goal. Suppose someone whose every moment of life brings them extraordinary pain and their only goal is to stop the pain as soon as possible. Achieving this goal will necessarily require their sudden death. They choose to decrease their pain accordingly. This is an unhealthy choice, but it is the only reasonable choice they have. Even the well-embedded health goal does not give us reasons to be healthy always and everywhere.

Nevertheless, notice how we typically treat being healthy as if it is good, always and everywhere. For example, being healthy is usually taken as sufficient reason to explain an action. We might ask: ‘why are you running or eating that or relaxing?’ The reply: ‘it is healthy’ is usually taken as a sufficient reason. Similarly, we usually consider everyone to be bound by the normative power of being healthy. You should eat your vegetables, and you should exercise and, you should relax. We do not usually add the caveat ‘...if you want to be healthy’. Lastly, we usually consider it a great harm to deprive children of the ability to be healthy, even if we thwart their preferences. Like all sensible humans, some children want to eat chocolate all day. However, letting a child fulfil that preference is inappropriate, in the absence of an extenuating circumstance. And if a child does not want to eat any vegetables, under usual circumstances, we would coerce the child into eating them. In general, we can treat promoting the goal of being healthy as good always and everywhere and we will rarely encounter any issues because it is such a well-embedded goal.

²¹ ‘Degree’ in a non-technical sense. In this paragraph, I describe a partial ordering. I leave the work for how to best formally model this ‘degreed’ relation between embedded goals for another time.

An important point of contrast between explicit goals and embedded goals is possible complexity.²² To understand this contrast, imagine a child playing his first game of chess against an adult. Suppose that both the child and the adult have the goal of winning the game. As a matter of fact, in the early game you should develop your pieces, castle, and control the centre spaces. The adult understands this, and has the explicit goal ‘develop my pieces, castle, and control the centre spaces’. The child is not even aware that you can castle in chess, so cannot have this explicit goal. Nevertheless, the child still has the embedded goal ‘develop my pieces, castle, and control the centre spaces’ because embedded goals are based on facts about what will help the child win in a wide variety of contexts, not on the cognitive limitations of the ignorant child’s mind. So embedded goals can be extraordinarily complex. They can be based on information a person does not possess. They can even be unimaginable. Embedded goals can be much more complex than their explicit counterparts.

Returning to our case study, consider the complexity of the goal being healthy. My best guess is that being healthy amounts to some complex combination of having a long and well-functioning body and mind. This complexity might pose a problem for formulating an explicit health goal because explicit health goals must be cognitively tractable. But there is no complexity constraint on embedded health goals. Depending on how the facts fall, the health goal might be a very complex goal indeed. It might be disjunctive. It might be inconceivable. It might even be pluralistic. This last possibility deserves some attention. The embedded health goal might be pluralistic like so: ‘achieve longevity to the extent that it does not overly reduce functionality and achieve functionality to the extent that it does not overly reduce longevity.’ Such a goal could be pluralistic in the sense that, from a health perspective, there is no single correct way to trade off longevity and functionality; there is no single notion of ‘overly reduce’. Indeed, extreme cases aside, when faced with a choice between longevity and functionality, it is hard to say what is the healthier choice for a third party to make. Therefore, the embedded health goal might have this messy complexity.

To summarise the Embedded Health Goal Model: there are explicit health goals and there is an embedded health goal. The embedded health goal is more robust because it depends on facts concerning what contributes to us achieving our other goals as opposed to facts concerning our psychology. We have a degreed notion of embeddedness: the more embedded a goal the more contexts where it provides reasons. However, just because a goal has a high degree of

²² This possible complexity avoids a problem that plagues some other accounts of epistemic instrumentalism: see section 2.4.

embeddedness does not imply that it provides reasons always and everywhere; there will be cases where a highly embedded goal provides no reasons. Finally, embedded goals can be very complex because their complexity is not constrained by our cognitive limitations. This means that people can have goals that are highly disjunctive, inconceivable, or even pluralistic.

2.2.2. The Embedded Epistemic Goal Model

Returning to epistemic instrumentalism, I am going to argue that people have an embedded **epistemic goal**—a goal to conform to the epistemic standards. Some people might have an explicit epistemic goal, the counterpart of my-body-is-a-temple people. Philosophers may be a prime candidate, I suppose. These epistemic champions explicitly aim to have beliefs that conform to epistemic standards. However, as suggested by our Embedded Health Goal Model, explicit epistemic goals are not required to account for the robust and widespread normative force of the epistemic standards.

Almost all of us have a variety of goals. As I claimed above, achieving the epistemic goal helps. For example, many people want to be well respected members of their community. It is hard to be respected if your testimony is false most of the time. We have short term mundane goals, like getting coffee and food. These require us to locate coffee and food. We have long-term goals concerning things that will happen many years in the future, often requiring us to accurately anticipate how people will react to various changes to their circumstances. And we will have any number of goals in the future—ones we have not even considered yet. These will likely require us to have beliefs that conform to the epistemic standards too. So, the epistemic goal is an embedded goal, intimately connected to the vast array of other goals we find ourselves possessing.

The epistemic goal has a very high degree of embeddedness. Even if we imagine aliens living on the planet Seti-546—if they have a variety of goals, and they use representations to guide their behaviour as we do—then they are going to have the embedded epistemic goal. Suppose there are animals without goals *per se*, like coral. Even then they will have something like goals to eat and to reproduce, so even they will have something like the embedded epistemic goal too. It is in something like their interest to have accurate representations of the world. The epistemic goal is among the most embedded goals we have.

The well-embedded epistemic goal provides a robust basis for epistemic normativity. Unlike relying on explicit goals, the facts that give rise to the embedded goal are consistent across time.

Unlike most explicit goals, the embedded goal can be very complex, very disjunctive, inconceivable, and pluralistic. We have not settled what the epistemic standards are and how they trade off against one another. Nonetheless we still have the embedded epistemic goal to conform to the epistemic standards.

Importantly, the embedded model is not committed to any particular source of normativity. I have presented the embedded model paired with pragmatic normativity, ultimately concerned with achieving explicit goals. However, the embedded model pairs with most sources of normativity, be it pragmatic, moral, aesthetic or some bespoke source. For example, the embedded model of epistemic normativity is compatible with holding that only morality generates genuine normativity. Even if morality alone generates genuine normativity—to achieve what morality demands of you, to treat people with respect or to maximise the good—you are required to achieve a wide variety of goals, more than enough variety to embed the epistemic goal (cf. Clifford, 1877 [1999]). If a source of normativity gives rise to a rich set of ends, then it is compatible with the embedded model of epistemic normativity.²³

Recall our central question: how do the epistemic standards produce epistemic reasons? My answer: conforming to the standards promotes having accurate representations. Having accurate representations is broadly useful in almost any context. Moreover, human cultures and psychologies have developed in such a way to reinforce and enhance the usefulness of conforming. So humans find themselves in an environment where conforming to the epistemic standards is persistently and enormously useful. Epistemic reasons are the persistent instrumental reasons we have to conform to the epistemic standards. The persistent web of epistemic reasons is epistemic normativity. Today you may want to be a respectable epistemic peer. Tomorrow you may want to discover a vaccine for a pandemic. Conforming helps achieve those things. But the specific individual goals or valuable ends you have do not really matter. More importantly, you have a plurality of goals or ends and conforming promotes them in a plurality of ways. This creates a robust normative pressure that permeates our lives. This is the normative grip that the epistemic standards exert on us.

There are a couple of caveats that restrict the scope of the embedded position. The first caveat is that ‘being useful’ does not imply ‘being the most useful’. A knife is useful for many tasks but there

²³ Strictly speaking, the embedded goal model is compatible with there being *sui generis* epistemic normativity. But the explanatory power of the model tells against this possibility (see section 2.4 and 4.1-2).

may be more useful tools to use for any particular task. Nevertheless, we find ourselves reaching for knives to help complete many tasks. Similarly, for embedded epistemic instrumentalists, usefulness is the thing that compels us to attempt to conform to the epistemic standards, metaphorically reaching for the standards to help complete many tasks.

The second caveat is that the strength of epistemic normativity in any particular case, including how it trades off against other types of normativity, is controversial. So the instrumentalist is not committed to the relevant instrumental reasons being particularly strong in any particular case. It might be that some of the instrumental reasons to conform are exceptionally weak. Moreover, instrumentalists are not committed to there being no instrumental reasons *against* conforming to the epistemic standards. It might be that in some cases, conforming incurs a large cost (imagine a malicious demon offers you the choice between conforming to the epistemic standards or saving the world).

With those two caveats in mind, instrumentalists need to hold that it is more useful to conform to the epistemic standards than the standards of chess in general, similar to the way a knife is more useful in general than a vegetable peeler. But in most cases, conforming to chess standards is not that useful (ditto other types of standards that philosophers tend to hold are not genuinely normative). Indeed, the epistemic standards are so broadly useful that to comply to the chess standards—say, do not leave your queen hanging—it helps to conform to the epistemic standards first and recognise that a move will leave one’s queen hanging. So this is not a severe constraint (I return to this point in section 2.5.4).

2.3. The Too-Few-Reasons Objection

So what is the problem? For epistemic instrumentalism, epistemic normativity requires conforming to the epistemic standards to be useful. According to the embedded epistemic goal model, while conforming is widely useful, it is not useful always and everywhere. In cases where conforming is not useful, you have no epistemic reasons to conform. These are known as **too-few-reasons cases**. For many, this is a deal breaker. They hold that the epistemic standards exert normative force always and everywhere. So, they suppose, we need some further fact about the epistemic standards, something beyond usefulness, to explain epistemic normativity in these cases.

Before examining a specific too-few-reasons case, I will point out a weakness with the objectors’ position. That conforming is widely useful is uncontroversial (even if there are questions about

how wide and how useful). As such, in a great many cases, non-instrumentalists are positing a glut of reasons (I expand on this point in section 4.1.1.1). They say that in addition to the instrumental reasons, there are some non-instrumental reasons too. The extra source of normativity might be a form of some *sui generis* normativity,²⁴ some essential property of beliefs, the result of a proper function (Nolfi, 2018), or one of many other possibilities. But this extra source of normativity is posited in addition to the usefulness of conforming because the general usefulness is hard to deny. The problem is that usefulness alone is enough to explain why we would develop epistemic standards. It is enough to explain why we would pay attention to them, why we would track who conforms and who does not, the social esteem generated by conforming, and why we would spend considerable resources on conforming. So, we have an uncontroversial fact that explains most of the relevant phenomena associated with epistemic normativity. Non-instrumentalists need a powerful motivation for postulating some additional source of normativity that in the overwhelming majority of cases is otiose.

The non-instrumentalist's position is a little too convenient. If you hold that there is a value, independent of usefulness, in conforming to the epistemic standards then you are suggesting an extremely fortunate coincidence took place. We learned to value conforming because conforming is so useful, but we also happened to find that conforming is valuable in an unrelated way. This is suspiciously convenient.²⁵ As such, making the too-few-reasons objection, suggesting that we need a non-instrumental theory of epistemic normativity, commits one to controversial claims from the outset. The too-few-reasons objection is not an objection from a neutral position. That does not mean, however, that instrumentalists should ignore it. Too-few-reasons cases are supposed to provide the extra motivation required for positing an additional source of epistemic normativity. They are commonly cited as the motivation for rejecting epistemic instrumentalism. So the too-few-reasons objection is worth serious consideration.

Recall that, Côté-Bouchard offers a too-few-reasons case (2015, p. 340), paraphrasing Kelly (2003):

Spoiler Alert Max missed the finale of his favourite television series, which aired last night and revealed whether character X ends up dying or surviving. Given his passion for the show, Max really wants to avoid learning the answer to that question before he watches the rerun tonight after work. But to his dismay, one of his colleagues – who is convinced that Max did watch the show last night – comes up to him in the morning shouting: ‘I knew it! I told you character X would die!’

²⁴ There are many philosophers who make such claims. Cowie cites (Kelly, 2003; Parfit, 2011; Scanlon, 1998; Wedgwood, 2007 cited in Cowie, 2014, p. 4004). The classic *sui generis* view of normativity is (Moore, 1903).

²⁵ I return to this point and greatly expand on it in 4.1.1.

In Spoiler Alert, supposedly, Max has no goals that are promoted by conforming to the standards and believing the testimony of his colleague. Therefore, he has no instrumental reasons to conform to the standards. And yet, philosophers judge²⁶ that he still has epistemic reasons to conform. So, epistemic instrumentalism alone cannot provide a complete account of epistemic normativity. This is a typical too-few-reasons case. It is hard to underestimate the influence these cases, and the objection based on them, has had on the literature.

In response, this type of everyday too-few-reasons case is not a problem for embedded epistemic instrumentalism. If Max is anything like a typical person, he has a wide range of goals that are promoted by conforming to the epistemic standards. For example, if he were to not believe his colleague, then his colleague would feel very awkward. Imagine a conversation in which Max maintains that he does not believe character X died. His colleague would feel distrusted, perhaps feel bad about themselves, or perhaps anger towards Max. The colleague might even lose respect for Max. For a lot of people, they do not want their colleagues to feel distrusted or bad about themselves or angry. Most of us care greatly about how much we are respected. And this is just one among many social goals that are promoted by conforming. If Max has the usual complex web of goals, there are plenty of instrumental reasons to conform. In an everyday case, Max has the embedded epistemic goal.

This fact that there are instrumental reasons in standard too-few-reasons cases is evidence for epistemic instrumentalism. Epistemic instrumentalism claims that epistemic reasons are instrumental reasons. Too-few-reasons cases were designed specifically to be cases where intuitively there are epistemic reasons but no instrumental reasons. But in standard too-few-reasons cases there remain many instrumental reasons to conform to the epistemic standards. This is exactly what one would expect if epistemic instrumentalism were true.

The above response is available to every instrumentalist. Once the embedded model is characterised, it is easy to see that Max will have many relevant goals that are promoted by conforming to the epistemic standards. But nothing in this response actually depends on the embedded model. And yet this response is not commonly offered. I suspect that philosophers can anticipate the next epicycle of the dialectic, idealised too-few-reasons cases (see section 4.2.1 below). Suppose Max no longer has the normal range of goals, but rather only one: to believe

²⁶ Again, let us suppose. See n. 15.

character X did not die. Place Idealised Max in the same circumstances, that is in the circumstances where all the epistemic standards point to the fact that character X died. Philosophers still judge that Idealised Max has epistemic reasons. But this Idealised Max does not have any goals that are promoted by conforming. So, I cannot claim Max has relevant instrumental reasons.

The standard response is to propose universal epistemic instrumentalism (Côté-Bouchard, 2015; Kornblith, 1993; Leite, 2007; Schroeder, 2007), which I criticise in chapter 3. Supposedly, some way or another, conforming to the epistemic standards *always* promotes at least one goal you happen to have, whether because we all share a common goal that is promoted, or because conforming promotes every goal. This is where standard developments of epistemic instrumentalism start focusing on the way that conforming promotes specific individual goals.

There are various reasons to develop the theory in this universal direction. One such reason is that the orthodoxy holds that the vast majority of philosophers judge that idealised Max has epistemic reasons. Perhaps partially explaining why I am not drawn to universalism, I do not judge that way. To make my judgement a bit more vivid, imagine a **Zen Boltzmann brain**, that is a brain that spontaneously came into existence in a void, lacking any goals we have, lacking a body, so it is essentially causally isolated. Is there some normative force compelling the Zen Boltzmann brain to conform to the epistemic standards? I do not have a strong intuition one way or another but if I were to bet, I would say it does not. I have a similar response to Idealised Max. Perhaps there are more philosophers who judge similarly to me than commonly supposed. But I am not questioning what judgements philosophers make in this thesis. Suppose the orthodoxy is right, that the vast majority of philosophers judge that Idealised Max has epistemic reasons. That is a motivation to develop a universal solution, but it is not a strong one. These judgements do not have much epistemic weight.

In response to idealised cases, I offer an error theory regarding too-few-reasons judgements (note that this is different to the error theory interpretation from chapter 6; this response is consistent with both interpretations of embedded epistemic instrumentalism from that chapter).²⁷ To warm you up to the idea of an error theory, allow me to make the following three observations. First, perhaps no actual agent has ever had just one goal such as: ‘believe *p*’. Indeed, I am not sure it is nomically possible for such an agent to exist. Second, even if there were such an agent, they would

²⁷ A similar response is offered by Sharadin (2018) although he does not have the resources of embedded epistemic instrumentalism to support this response. Also, Sharadin’s positive view has at least one serious issue: see section 2.4.

have a psychology very unlike the vast majority of humans. Third, the judgement we are considering is that the idealised agent has epistemic reasons. This is not a Moorean proposition like ‘I have hands’, which we arguably have more reasons to believe than any philosophical argument to the contrary could provide. The judgement we are considering concerns a technical concept in an extremely unusual context. Even if you take philosophical judgements seriously, this is not a strong candidate for an epistemically impeccable judgement.

The embedded epistemic goal model explains why we would make such judgements. The epistemic goal is so well-embedded that it is an efficient heuristic to treat epistemic reasons as if they applied always and everywhere (just as in the health case). We can talk and act as if epistemic reasons apply always and everywhere because they in fact apply to almost all actual agents almost all the time. The cases where there are no epistemic reasons are so unusual that the heuristic rarely provides the wrong judgement. And even in cases where we do over generalise, such as in response to *Idealised Spoiler Alert*, the mistake has little downside. The only cost is that our abstract epistemic theorising is led astray. But if we need a place to sleep, or we are hungry, or we do not know where our kids are, some damage to our abstract epistemic theorising can be safely ignored. But as we are now fat and warm and safe, we may stop to appreciate how this heuristic is misfiring in very unusual cases and correct our abstract epistemic theorising.

This response was partly anticipated by Thomas Kelly, who rejects it:

The viability of this instrumentalist response, of course, presupposes that there is some shared cognitive goal which might underwrite the existence and intersubjectivity of epistemic reasons. In fact, it is here, I believe, where the instrumentalist conception of epistemic rationality founders: there is simply no cognitive goal or goals, which it is plausible to attribute to people generally, which is sufficient to account for the relevant phenomena. Individuals do not typically have this goal: believing the truth (Kelly, 2003, pp. 623, also see 622, 632-633).

Kelly’s rejection of the error theory response is plausible if the relevant goal in question is an explicit goal, or if we think about some specific goal people have in isolation. Not everyone explicitly wants to conform for its own sake. Getting coffee does not require us to conform to the epistemic standards wholesale. Kelly’s reply lands hard when the instrumentalist is focusing on the trees rather than the forest.

However, thinking about epistemic instrumentalism in this way is a mistake. The plurality of goals people have give rise to the well-embedded epistemic goal. Many people have the embedded epistemic goal because they want to be reliable testifiers, because they want to achieve their long-term plans, because they do not want to be patronised, because they do not want to make people uncomfortable—to present just a few of the long list of goals any typical person will have. As in the health case, more than enough people have the embedded epistemic goal to cause the practice of treating epistemic standards as if they gave rise to reasons that apply always and everywhere.

Too-few-reasons objectors like Kelly are not on firm ground. The proponents of the embedded epistemic goal model are committed to the claim that philosophers are overgeneralizing from a useful heuristic in response to idealised too-few-reasons cases. However, we know that humans use heuristics and we know they over generalise from them, so it is no surprise that we would be doing this in response to cases that involve very unusual agents. Meanwhile, too-few-reasons objectors are committed to the claim that our judgements in response to idealised too-few-reasons cases are accurate. Considering these judgements are the product of messy processes—like biological and cultural evolution as well as on-the-fly idiosyncratic learning processes—and considering the cases involve very unusual agents, this is a more significant commitment than the overgeneralizing one. So, the error theory here is not some last-ditch effort to save epistemic instrumentalism (cf. Sharadin, 2018). Rather it is at least as plausible as (if not more so than) maintaining that judgements in response to idealised too-few-reasons cases are uniformly accurate.

To sum up the response to the too-few-reasons objection: first, the objection is not made from an uncontroversial position. Usefulness is how we learnt to appreciate conforming, so positing an additional source of normativity is suspiciously epistemically convenient. Second, actual too-few-reasons cases that are put to instrumentalists, like Spoiler Alert, are easily replied to: there are plenty of relevant goals. Third, the cases that are not easily replied to, such as Idealised Spoiler Alert, are plausibly replied to with an error theory—an error theory that is at least as likely, if not more so, than the claim that our judgements in response to idealised cases are all strictly accurate. So, the too-few-reasons objector is not in a strong position.

Now I want to get precise about exactly what the embedded epistemic instrumentalist is claiming. Suppose in *Idealised Spoiler Alert*, that Idealised Max believes ‘I have no evidence that character X dies’, flouting the epistemic standards. In response, embedded instrumentalists hold that Idealised

Max has no epistemic reasons against believing ‘I have no evidence that character X dies’. The idealised agent lacks the rich set of goals that give rise to epistemic reasons. Importantly, embedded instrumentalists do not say that Idealised Max *has* epistemic reasons to believe, either.²⁸ Idealised Max requires the embedded epistemic goal, or some other relevant explicit epistemic goal, to have epistemic reasons. But he lacks any epistemic reasons. In terms of ‘rationality’ Idealised Max is neither epistemically rational nor epistemically irrational for believing ‘I have no evidence that character X dies’. I consider Idealised Max pragmatically rational. I would consider him pragmatically irrational if he failed to believe. Although your view will turn on your view of pragmatic rationality. I also consider Idealised Max all things considered rational, though again your view will turn on your view of all things considered rationality. Of course, embedded instrumentalists can still evaluate Idealised Max as if he had the embedded epistemic goal. They can point out that Idealised Max did not believe in accordance with the epistemic standards. They just deny that this fact has normative significance on its own.

2.4. The Advantages of Embedded Epistemic Instrumentalism

Embedded epistemic instrumentalism has many virtues. I articulated three such virtues in the introduction: it avoids metaphysical and epistemic difficulties associated with other accounts; it is compatible with a naturalistic world view; and it is not a theory about mere semantics. To this we can add that embedded epistemic instrumentalism does not require some convenient epistemic coincidence nor does it rest on epistemically dubious judgements (see section 2.3). However, many theories in the vicinity of embedded epistemic instrumentalism share these virtues. Embedded epistemic instrumentalism has other advantages over these accounts. In this section I will identify some of these advantages.

Contrast embedded epistemic instrumentalism with the standard instrumental position, universal epistemic instrumentalism (Côté-Bouchard, 2015; Kornblith, 1993; Leite, 2007; Schroeder, 2007).²⁹ Universal instrumentalists hold that always and everywhere we have epistemic reasons to conform to the epistemic standards. They hold that in every too-few-reasons case, no matter how idealised, there are instrumental reasons to conform to the epistemic standards. As such, standard epistemic instrumentalism also rests on those epistemically dubious judgements about highly idealised cases.

²⁸ This is one of the differences between an instrumental position and a purely pragmatic position. The pure pragmatist holds that all reasons are of the same kind. They do not draw a distinction between epistemic and non-epistemic reasons. Epistemic instrumentalists draw a distinction, holding that there is a special class of reasons, the epistemic reasons.

²⁹ I will criticism this position more completely in the next chapter (chapter 3).

Moreover, universal epistemic instrumentalism has resulted in some complicated positions, which have provoked many criticisms.³⁰ The problems with universal epistemic instrumentalism, which I go over in more detail in the next chapter, stem from the fact that universal instrumentalists are holding on to two ideas:

1. We always have epistemic reasons to conform to the epistemic standards, no matter the situation (this is the universal part).
2. Epistemic reasons boil down to facts about usefulness (this is the instrumental part).

Combining these parts, the universalist must hold that it is always useful to conform to the epistemic standards. But consider our inductive evidence against this. In the vast majority of cases, whether or not something is useful depends on the situation you are in. A screwdriver is useful if you need to screw a screw but terrible for making an espresso. Having a lot of money is useful for buying a house but terrible in a revolution. Consider how surprising it would be to learn that a tool is always useful, to learn that there are no possible situations where that tool would not be a little helpful—extraordinarily surprising. The fickle nature of usefulness suggests we should have a low credence for the idea that conforming to the epistemic standards is always useful. This is why the central fact that motivates epistemic instrumentalism, that conforming is widely useful, is a poor foundation for trying to explain the purported universal truth, that the epistemic standards have normative force always and everywhere. Usefulness is a capricious property, shifting wildly with slight changes to context. It is ill-suited for the universal project. This is why attempts to argue for the universal position require a great deal of philosophical gymnastics and why the universal project will not succeed. Embedded epistemic instrumentalism, by contrast, does not have this internal complexity. It does not have to defend improbable claims. It is a much simpler theory with a stronger reply to the too-few-reasons objection. It is a more internally coherent version of epistemic instrumentalism than the standard.

Moreover, the idea of an embedded goal allows embedded epistemic instrumentalism to address another concern about standard developments of epistemic instrumentalism. Matthew Lockard identifies this additional concern:

³⁰ For criticism of various universal positions see: (Côté-Bouchard, 2015; Lockard, 2013; Sharadin, 2018) and chapter 3.

One serious (and, in my view, underappreciated) difficulty facing such forms of epistemic instrumentalism concerns the level of conceptual sophistication they demand of individuals whose beliefs may properly be regarded as epistemically rational or irrational. Desiring to achieve an epistemic goal presumably requires one to have a conceptual repertoire that includes such concepts as proposition, belief, and truth. Young children and higher animals seem capable of having rational and irrational beliefs despite lacking these concepts—hence despite lacking desires to achieve any epistemic ends whatsoever (2013, p. 1706).

But embedded goals can have a normative grip on us while not being desired or even being unconceived. As a result, this overly-demanding objection is not a problem for embedded epistemic instrumentalism. Indeed, as I mention above, even creatures that have no sophisticated cognition such as coral have something resembling an embedded epistemic goal. So embedded epistemic instrumentalism has at least two serious advantages over standard epistemic instrumentalism: it offers a stronger response to the central objection and it is plausibly undemanding of epistemic agents.

There are some other non-standard, non-universal instrumental accounts of epistemic normativity, however, and embedded epistemic instrumentalism has advantages over them too. One such advantage is that it offers an account of epistemic normativity. Contrast this with **ecumenical epistemic instrumentalism**, as defended by Nathaniel Sharadin (2019).³¹ Sharadin's position is that if there is an epistemic reason to believe something, then there is also an instrumental reason to believe that same thing but that we should be silent about whether the epistemic reason is identical to the instrumental one. This position is ecumenical in the sense that philosophers of all metaphysical stripes—naturalists and non-naturalists alike (see section 5.1)—could accept it. A central motivation for holding this position, according to Sharadin, is that the literature already implicitly identifies the instrumental position with the ecumenical position (2019, Section 5). Much of the literature on epistemic instrumentalism is concerned with whether there are instrumental reasons in all the purported cases of epistemic reasons; essentially a lot of the literature is concerned with the too-few-reasons objection. Sharadin holds that this implies that both instrumentalists and non-instrumentalists already identify the instrumental position with the ecumenical position. The thought is that because much of the disagreement in the literature

³¹ Sharadin's position can be considered universal or non-universal, though he is a non-universalist in other work (Sharadin, 2018)

concerns whether there are instrumental reasons in certain cases, and ecumenical epistemic instrumentalism involves the claim that there always is, then the literature is implicitly just debating ecumenical epistemic instrumentalism. I reject this line of reasoning. Rather, I hold that philosophers believe that answering whether there are instrumental reasons in certain cases will provide evidence for or against non-ecumenical epistemic instrumentalism (the position that epistemic reasons are instrumental reasons). If in all cases of epistemic reasons there are instrumental reasons, then that is evidence that epistemic reasons are a sub-species of instrumental reasons, as this is the obvious explanation for why epistemic reasons always come in tandem with instrumental ones. Alternatively, if there are epistemic reasons without instrumental reasons then that is evidence for epistemic reasons being something else. Moreover, the question of whether there are instrumental reasons in certain cases loses much of its importance once it is divorced from the question of whether epistemic reasons are instrumental reasons. I am (and I suspect many other philosophers in the literature also are) trying to understand the nature of epistemic normativity. Ecumenical epistemic instrumentalism refuses to engage with that concern. While staying silent on that issue may result in a position that is easier to defend, it undermines a key objective of an investigation into epistemic normativity. Embedded epistemic instrumentalism engages with the core concerns of many philosophers who are interested in epistemic normativity in a way that ecumenical epistemic instrumentalism does not.

Dyke also offers a non-universal epistemic instrumentalism. As I outlined above, she holds that communities have epistemic goals, and that epistemic normativity arises because people are a part of these communities (Dyke, 2020). A problem with this position is that she does not identify how being part of a community with some goal gives an individual reasons to achieve that goal. She does not require people to endorse the community or personally share the community's goal, making it unclear what the reason-giving-connection is between individuals and communities. Furthermore, there is evidence that community goals alone do not give rise to individual reasons. I might be part of a community that pursues sexist ends but that does not give me reasons to pursue those sexist ends. Similarly, I might be part of an epistemic community but that does not give me reasons to pursue those epistemic ends. So, Dyke offers no account of how community goals give rise to individual reasons and there is evidence that community goals alone do not give rise to individual reasons. Contrast this with embedded epistemic instrumentalism which offers

reasons that bear on individuals. It is your complex web of goals or ends—including those shared with a community—that give rise to the embedded epistemic goal.

Another problem is that Dyke bites the bullet in response to lone wolf cases, a type of too-few-reasons case, where the agent in question has no community:

On the view I propose, it is indeed true that a person who was stipulated by hypothesis not to be a member any social groups with epistemic ends would not possess any distinctively epistemic reasons for belief. ...

Yet I do not think it is such an absurd result, after all, to suppose that a lone human being (e.g. the only person left on Earth after a catastrophic event) would lack epistemic reasons. Note that we need not deny that the lone person would still have purely *practical* reasons for belief. She would still have purely instrumental reasons to promote her own actual interests. So, for instance, if the last human survivor on Earth aims at surviving for much longer, then she had better learn to form reliable beliefs about where she has just left her provisions of food and water. What she would lack are the kinds of distinctively epistemic reasons for belief that could not be accounted for by any purely individualistic instrumental conception of epistemic reasons (footnotes removed and original emphasis) (Dyke, 2021, p. 12).

I do not find it at all absurd that in a case without a community people have different types of reasons. Perhaps it is important to note that one of the central mechanisms that reinforce the usefulness of conforming is not present when we lack a community. However, this reaction to lone wolf cases is a weakness in Dyke's position.

Firstly, I note that some of the other mechanisms that reinforce the usefulness of conforming—our psychologies and our ability to practice habits—still exist in lone wolf cases and will provide stable reasons to conform to the epistemic standards. Even when they are lone wolves, many people will want to be smart and reasonable, and this requires conforming. Even when they are lone wolves, people have reasons to practice conforming to the epistemic standards. So even in lone wolf cases it will be robustly useful to conform. So on Dyke's view, the robust usefulness created by being part of a community is treated extremely differently to the other ways conforming is made robustly useful.

Secondly, Dyke is explicitly motivated by other too-few-reasons cases (Dyke, 2021, p. 2). This is what motivated her to posit the community-based reasons as the distinctly epistemic reasons. While it is not inconsistent to bite the bullet in lone wolf cases but modify a theory in response to other too-few-reasons cases, the difference in treatment of the cases requires more motivation.

Dyke does hold that lone wolf cases are not typical (Dyke, 2021, p. 12). But neither are successful too-few-reasons cases, the type of case that would require someone to invoke community-based reasons to explain the existence of epistemic reasons. So there is an undermotivated asymmetry in the treatment of the cases. Contrast this with embedded epistemic instrumentalism. It is not motivated by too-few-reasons cases. It starts with an observation about the variety of ways conforming is useful. This observation then implies certain responses to cases but at no stage does the theory rely on people having any particular judgement. So it does not face this tension of explaining why certain too-few-reasons cases provoke changes in our theories and other too-few-reasons cases do not.

There are also non-instrumental accounts that have many of the virtues of instrumental accounts. Consider Kate Nolfi's account of epistemic normativity (2018), which holds that the proper function of beliefs gives rise to epistemic normativity. Here the 'proper function' of an object means the function that was selected-for when the object evolved. We might say the heart has the proper function to pump blood, for example. Suppose the proper function of beliefs is to provide a map to guide your behaviour. Roughly, Nolfi holds that beliefs should conform with epistemic standards because that best achieves their proper function. This shares many of the virtues of embedded epistemic instrumentalism. But a problem with this approach, much like Dyke's approach, is that it has nothing to do with us. For example, at the time and location of writing, I am in a very calorie-rich environment. My body evolved to store a lot of calories. I would prefer it if my body did not store them. I like eating a lot and I want to live a long time. I am in a very different environment to the one in which my ancestors evolved. The proper function of my fat storage system does not match my current environment. The proper function of my fat storage system is not about me in a relevant sense. Similarly, the proper function of my beliefs is not an important factor when considering what I should believe because it is not about me. Contrast this with embedded epistemic instrumentalism. The goals or ends that I have are about me. They can even explain what might be appealing about a proper function theory. If by fulfilling its proper function, my fat storage system promotes the goal of me living longer, its proper function becomes important. If the proper function does not promote (or hinder) my goals or ends, I can discard consideration of it. Embedded epistemic instrumentalism provides reasons for individuals to have certain beliefs, in a way that Nolfi's and Dyke's accounts do not.

Of course, this is not an exhaustive list of advantages. Some philosophers may even find disadvantages in holding the embedded position. But this is a strong candidate for an account of

epistemic normativity, nonetheless. It has the virtues of epistemic instrumentalism. It has advantages over both standard and non-standard developments of epistemic instrumentalism. It explains how epistemic reasons bear on individuals—on me and you. So embedded epistemic instrumentalism has a lot going for it.

2.5. Embedded Epistemic Instrumentalism and Other Topics in Epistemology

In this section, I will relate embedded epistemic instrumentalism to related topics in epistemology: to different types of epistemic instrumentalism in section 2.5.1, to ‘evidence’ in section 2.5.2; to demarcating the epistemic from the non-epistemic in section 2.5.3; and the epistemic standards in 2.5.4. This section is supposed to offer some suggestions (but only suggestions) about how this theory integrates with the rest of epistemology. With the notable exception of ‘evidence’, the theme of this section is that embedded epistemic instrumentalism is extremely ecumenical between many possibilities.

2.5.1. Different types of epistemic instrumentalism

Côté-Bouchard offers a distinction between types of epistemic instrumental theories: between ‘source epistemic instrumental theories’ and ‘standard epistemic instrumental theories’ (Côté-Bouchard, 2019).³² **Source epistemic instrumental theories** are theories that hold the source of epistemic normativity is instrumental in nature. Embedded epistemic instrumentalism is a source epistemic instrumental theory. **Standard epistemic instrumental theories** are theories that hold the epistemic standards are instrumental in nature. Embedded epistemic instrumentalism is not a standard epistemic instrumental theory. Except for the constraints laid out in this chapter, it is uncommitted about the nature of the standards.

An example of a standard epistemic instrumental theory is epistemic consequentialism (Singer, 2018a, 2018b):

Epistemic consequentialism is the view that ‘the epistemically right (e.g., the justified) is to be understood in terms of conduciveness to the epistemic good (e.g., true belief). Like ethical consequentialism, epistemic consequentialism is a framework for more specific views: epistemic consequentialists might disagree about what epistemic goodness is (popular options include true belief and knowledge, and less-popular ones are justified belief and understanding), what epistemic rightness is (for example, justification, rationality, warrant, or what one epistemically ought to believe/do), and what the conduciveness relation is (for example, maximizing over the long term, or indirectly

³² He uses the terms ‘Source-EP’ and ‘Norm-EP’.

promoting). A simple consequentialist view might be that a belief is right iff, among the available options, having the belief maximises the number of true beliefs (and minimises the number of false beliefs) the holder has in their life-time. (footnotes removed) (Singer, 2018a, pp. 581-582).

Embedded epistemic instrumentalism is compatible with this type of picture. But it is also compatible with denying this type of picture. It all depends on one's conception of what the epistemic standards actually are. Even though it is a source epistemic instrumental theory, embedded epistemic instrumentalism has no implications about standard epistemic instrumental theories.

2.5.2. Evidence

Kelly argues that instrumentalists need a concept of 'evidence' that is not essentially normative (Kelly, 2007). While I may quibble with the wording, I agree that embedded epistemic instrumentalism places constraints on possible understandings of 'evidence'. Recall that our example epistemic standard is: 'believe in proportion to your evidence'. If the instrumentalist interprets 'in proportion to' as some kind of normative relation between evidence and beliefs, then the instrumentalist is invoking the exact type of non-instrumental normativity to understand evidence that they have been avoiding invoking to explain epistemic normativity. Even though it is not strictly inconsistent with embedded epistemic instrumentalism that there is some non-instrumental normative relation between evidence and belief, appealing to this is in tension with the motives of embedded epistemic instrumentalism. So the instrumentalist needs an alternative.

I concede that I do not have a purely non-normative analysis of evidence. But I have an argument that there is some part of the world that constitutes non-normative evidence for the instrumentalist, even though I do not know what part of the world that is. The argument starts by noting that some creatures can do better than randomly guess when making predictions about the future, about the effects of interventions on the world. I am fairly accurate at predicting which keys to tap on my keyboard to make certain letters appear on the computer screen, for example. This is a complex causal process. Some part of this process involves the non-normative sense of evidence that instrumentalists need because we need to appeal to that evidence to explain what is happening.

To see this, suppose two people named Tyler and Macy, are playing poker (Texas no-limit hold'em, if you need specifics). Tyler and Macy are of equal skill. However, Macy is cheating. She has a

carefully-placed mirror so that she can see Tyler's cards. Tyler's cards (or Macy seeing the cards or Macy believing that Tyler has such and such cards, etc.) are evidence for whether Macy will end up with the strongest hand after the last card (the river) has been dealt (even though, on occasion, it will be misleading). Suppose over the course of an evening, Macy wins more poker chips than Tyler (I find it plausible that cheating sometimes works). How would you explain Macy's success? Part of the explanation is that she had access to more evidence. And this caused various things to happen, including more chips to end up in Macy's possession. And when I say 'caused', I mean there are various physical phenomena that are a result of Macy having more evidence; there are more poker chips on her side of the table at the end of the game. A notion of 'evidence' features in explanations of certain physical phenomena. This is the notion of evidence the instrumentalist needs.

To make this point more vivid, suppose two copies of the same algorithm are playing poker, a Macy-algorithm and a Tyler-algorithm. If you give the Macy-algorithm access to the Tyler-algorithm's cards, in the long run the Macy-algorithm is going to win more money. And there is nothing specific about the metaphysics of the algorithms playing poker that we do not understand. In principle, we can follow the complex causal pathways that will lead to the Macy-algorithm to have more chips. And this same thing is happening, albeit with much more complex biological systems, when Macy wins the poker game. The example shows you that there is some notion of evidence that is purely descriptive that explains why algorithms with access to more information will win in the long run and why Macy ends up with more poker chips when cheating.

For a less toy example, note that eyes have evolved independently many times in evolutionary history. When explaining why the eye has evolved independently many times, the fact that eyes give creatures access to a wider range of evidence is important. The access to extra evidence explains why eyes help creatures survive. But surviving is a causal, worldly phenomenon. The evidence in questions involves photons, cells, neurons, etc. The sense in which evidence promotes survival involves a descriptive notion of evidence. This notion of evidence, whatever it is, is the notion of evidence that the instrumentalist will appeal to.

Indeed, the existence of a non-normative notion of evidence is implied by a widely held thesis, normative supervenience. **Normative supervenience** is the thesis that the normative supervenes on the non-normative, so if there are no non-normative differences between two cases then there are no normative differences. We only require a local version of normative supervenience to imply

that some part of the actual world constitutes non-normative evidence. **Local normative supervenience** is the claim that at least in the actual world, if we hold fixed all the non-normative properties, then the normative properties remain the same. Lots of philosophers endorse a much stronger version of supervenience than this. Although he only endorses a local version himself, Gideon Rosen notes that “Supervenience is as close to common ground as we find anywhere in philosophy” (Rosen, 2020, p. 206). Tristram McPherson notes that “... supervenience is taken to be both obvious and uncontroversial” (McPherson, 2021, section 2).

Even if evidence is an essentially normative thing, local supervenience implies that in the actual world, whenever there is a difference in evidence, there will always be a difference in the non-normative things. So long as a creature or machine is sensitive to those non-normative differences, it can learn to track evidential relationships. This is how a robot vacuum cleaner can navigate a living room. This is how biological creatures could evolve a sensitivity to evidence. And this sensitivity to the non-normative differences in different cases of evidence is all the instrumentalist needs.

There is future research in articulating exactly what this notion of evidence is. But for this thesis, the fact that there is a non-normative part of the world that constitutes evidence is enough to put Kelly’s worry to rest.

2.5.3. Demarcating the epistemic from the non-epistemic

In section 1.3.4, I claimed the arguments in this thesis do not need the term ‘epistemic’ to be defined in a very particular way. With embedded epistemic instrumentalism on the table, I wish to return to this topic. Embedded epistemic instrumentalism is ecumenical between many different ways we might demarcate the epistemic from the non-epistemic. It is compatible with drawing a very sharp boundary between the epistemic and non-epistemic. It is compatible with a fuzzy boundary. And it is even compatible with there being only fuzzy differences between epistemic and non-epistemic, with no uncontroversial core. It is compatible because of the structure of the theory. The theory holds that *epistemic* normativity is the usefulness of conforming to the *epistemic* standards. However sharp or fuzzy the epistemic standards are from other standards is how sharp or fuzzy the difference between the epistemic and non-epistemic is.

To offer one such way of thinking about the demarcation, a good model is the model of biological species. First note that there are many controversies about how to analyse ‘species’, despite general

agreement that at some level, all creatures are atoms in the void (Ereshefsky, 2022). So even if we all agree that epistemic reasons are a type of instrumental reasons, that does not imply that we will agree about how to categorise all the different types of instrumental reasons. Second note that there are a variety of strategies used to demarcate species and these strategies will have analogues for demarcating types of reasons. For example, one strategy is to appeal to a certain concept of natural kinds: Homeostatic Property Clusters (HPCs) (Ereshefsky, 2022, section 2.4). Marc Ereshefsky characterises this approach to species:

HPC theory assumes that natural kinds are groups of entities that share stable similarities. HPC theory does not require that species are defined by traditional essential properties. The members of *Canis familiaris*, for example, tend to share a number of common properties (having four legs, two eyes, and so on), but given the forces of evolution, no biological property is essential for membership in that species. For HPC theory, the similarities among the members of a kind must be stable enough to allow better than chance predictions about the properties found among the members of a kind. For instance, given that we know that Lizzy is a dog, we can predict with greater than chance probability that Lizzy will have four legs (Ereshefsky, 2022, section 2.4).

There is a similar strategy for identifying epistemic (as opposed to other) standards. For example, we might say that given that a certain standard is an *epistemic* standard, we can predict that the standard will be about the properties that a representation will have. But appealing to HPCs is only one strategy philosophers might adopt. Indeed, as pluralism about species is an option (Ereshefsky, 2022, section 3) so is pluralism about the epistemic. Depending on our purposes, there might be different ways of grouping standards together. Embedded epistemic instrumentalism is compatible with these possibilities.

For example, Brian Talbot argues that we have truth-promoting non-evidential epistemic reasons (Talbot, 2014). A truth-promoting non-evidential reason is a reason to believe, where such a belief would promote the truth of your beliefs in general but not that belief in particular. For example, suppose it is the case that believing falsely that you are very smart makes it the case you will have more true beliefs in the future. Even if you only cared about truth, you would have a reason to believe you were smart, even if the evidence suggests otherwise. Most importantly for our purposes, Talbot argues that such reasons are *epistemic* in nature (Talbot, 2014, section 6). This is a demarcation controversy. Are such reasons *really* epistemic or not?

Embedded epistemic instrumentalism is ecumenical. It is compatible with Talbot's position that either these reasons are epistemic or they ought to be considered epistemic. Talbot's position fits

well with the metaphysics of embedded epistemic instrumentalism, because all epistemic reasons are instrumental reasons on this picture, so it is not so surprising that there are other types of instrumental reasons that also inform what one ought to believe. However, embedded epistemic instrumentalism is also compatible with the denial that such reasons are epistemic reasons. It might be that the epistemic standards are such that they prohibit sensitivity to truth promoting non-evidential reasons—like a judge ruling that a jury is not allowed to consider certain information in a trial. If an embedded epistemic instrumentalist could establish this about the standards, then such an instrumentalist could deny that such reasons are *epistemic*.

If pressed, I would say that it does not actually matter whether they are called ‘epistemic reasons’ or not. Sometimes epistemologists are going to want to ignore them. For example, Talbot admits that truth-promoting non-evidential reasons are in general less modally secure than reasons produced by evidence (Talbot, 2014, p. 612). So if one were only interested in the most modally secure reasons for belief—reasons shared by all possible agents say—then one would have motive to ignore them. However not all epistemological investigations are directed at modally secure reasons. I can imagine an epistemologist being interested in the types of beliefs that make for a good member of an epistemic community (cf. Zollman, 2010). This type of investigation would be interested in truth promoting non-evidential reasons for belief, in particular the reasons for belief people have that are generated by their role in an epistemic community (cf. Fleisher, 2018). Most importantly, whether these academic investigations are worthwhile does not depend on what we call the associated reasons.

Moreover, possibly, the term ‘epistemic’ is too vague to include or exclude truth-promoting non-evidential reasons. It may be like arguing whether a certain number of grains of sand in close physical proximity constitutes a heap. Importantly embedded epistemic instrumentalism is compatible with the term ‘epistemic’ being quite vague, too. So however one demarcates the epistemic from the non-epistemic, embedded epistemic instrumentalism will be consistent.

2.5.4. The epistemic standards

Embedded epistemic instrumentalism is compatible with a wide range of possible epistemic standards. The only constraint is that conforming to such standards must be useful. However, I predict that any independently plausible epistemic standard will be useful to conform to.

Firstly, the content of epistemic standards is usually something about producing accurate representations. If a standard lacks *any* connection to accurate representations, it is hard to argue that the standard is epistemic in nature. As I claimed above, accurate representations are useful for almost everything else you could do. So a standard about accurate representations, particularly one about reliably producing accurate representations, will tend to be useful to conform to.

Secondly, usefulness is invoked directly to argue for certain standards—for example Dutch-book arguments for probabilism (Ramsey, 1926; Vineberg, 2022).³³ Philosophers identify actual epistemic standards by finding useful standards.

Thirdly, a standard that lacks apparent usefulness puzzles philosophers. Consider value-of-knowledge problems (Pritchard, Turri, & Carter, 2022). These problems are about explaining why knowledge is more valuable than any proper subset of its constitutive elements. For example, why is knowledge more valuable than true belief? Note that we can reinterpret this as a puzzle about usefulness. Why is attaining knowledge more useful than attaining true beliefs? Answering this question answers this particular value problem:

Plato's own solution was that knowledge is formed in a special way distinguishing it from belief: knowledge, unlike belief, must be 'tied down' to the truth, like the mythical tethered statues of Daedalus. As a result, knowledge is better suited to guide action. For example, if one knows, rather than merely truly believes, that this is the way to Larissa, then one might be less likely to be perturbed by the fact that the road initially seems to be going in the wrong direction. Mere true belief at this point might be lost, since one might lose all confidence that this is the right way to go (Pritchard et al., 2022, section 1).

Plato's argument, via Pritchard *et al's* interpretation, is that knowledge is more resilient than mere true belief. Knowledge is less sensitive to new evidence. Sometimes acting well over time requires stable representations. As knowledge is less sensitive to new evidence, it is more stable over time than non-knowledge states. This is why it is useful to have knowledge rather than just true belief.

Fourthly, and I will explain this point in detail in chapter 4, usefulness is invoked to explain why certain standards would evolve. If a standard was not useful, it is hard to tell a plausible story about how humans learn that standard. To summarise, I predict that plausible epistemic standards will

³³ Probabilism is the view that your degrees of belief (credences) should conform to the basic axioms of probabilities. Dutch book arguments are arguments that certain combinations of credences that don not conform to the basic axioms of probability will make certain combinations of bets seems reasonable, even though together those bets will result in a sure loss.

be useful because epistemic standards are about accurate representations, because one way to argue for a standard is to point to its usefulness, because if a standard lacks usefulness then it puzzles philosophers, because we invoke usefulness to offer evolutionary explanations about how humans learn such standards.

However, there is an ambiguity in the phrase ‘the epistemic standards’. One can interpret the phrase descriptively: the standards that humans currently have. One can interpret the phrase abstractly, as referring to the standards an agent that has limitless cognitive resources should conform to. One can interpret the phrase as offering exactly one set of standards. One can interpret the phrase as referring to a collection of rival standards. One can interpret the phrase prescriptively. And there are more interpretations possible. Which one do I mean? I mean them all. This ambiguity is present in epistemology itself. There are debates about exactly what the relevant level of analysis for epistemology is, about how idealised, how descriptively accurate, how normative epistemology is supposed to be (Kornblith, 2019). As I have tried to make this model as uncontroversial as possible, I have not taken a stand on the correct interpretation of what exactly normative epistemology is discussing.

Embedded epistemic instrumentalism is compatible with all these interpretations. If you interpret the phrase descriptively, embedded epistemic instrumentalism explains the normative pull of those actual standards. If you interpret the phrase abstractly, it explains why such an ideal agent would have normative pressure to conform, and why merely human agents might try to ape those standards. If the phrase refers to one standard, then the theory is about that standard. If the phrase refers to multiple standards, then the theory is about those standards. And so on.

A virtue of embedded epistemic instrumentalism is that it can be plugged-in to many different conceptions of epistemology. For example, Susanna Rinard argues that epistemic normativity is of the same *kind* as other types of normativity (Rinard, 2017). Embedded epistemic instrumentalism is compatible with this view because it offers an instrumental reduction of epistemic reasons. Indeed there are a lot of similarities between Rinard’s view and embedded epistemic instrumentalism. This is surprising because Rinard has a quite different aim for her theory to mine (see section 1.2) and argues for her position with different styles of arguments to mine. She is interested in offering a theory of rationality that helps solve difficult cases of what to believe from a first-person point of view and she argues for her view by offering intuitions or judgments in response to cases (Rinard, in press). That being said, strictly speaking, embedded epistemic

instrumentalism is also compatible with denying Rinard's view. As I argued in section 2.5.3, an embedded epistemic instrumentalist may insist that epistemic normativity is a different kind to other types of normativity, by appealing to similar resources that we might use to distinguish species.

Indeed, in general I will appeal to embedded epistemic instrumentalism's ecumenical nature as both a virtue and as a strategy for avoiding criticism. Whether beliefs are voluntary or involuntary, whether epistemic reasons are commensurable or incommensurable or have some degree of incommensurability with other reasons, whether reasons are subjective or objective or both, embedded epistemic instrumentalism is compatible with it all. It is a virtue of the account that we do not need to resolve the many controversies in epistemology to settle what epistemic normativity is.

3. Against Universal Epistemic Instrumentalism

As I claimed in the previous chapter, I am an epistemic instrumentalist. However, there is a certain kind of epistemic instrumentalism, *universal epistemic instrumentalism*, that I find unacceptable. That I find it unacceptable is not so unusual. There is a long list of criticisms against this position. However, I am unsatisfied with existent criticisms. They fail to correctly identify the deep flaws in the universal position. I aim to rectify that.

The plan: in section 3.1, I review material from above: I distinguish epistemic instrumentalism from universal epistemic instrumentalism and summarise some motivations to adopt the universal position. In sections 3.2-5, I characterise various ways the universal position has been developed. In each of these sections, I identify deep problems with each version. Finally in section 3.6, I offer a conjecture that no universal position will succeed.

3.1. Epistemic instrumentalism and the universal response

In this section I want to review the important elements of the instrumental position for this chapter. Recall, **epistemic instrumentalism** is the thesis that you should conform to the epistemic standards because conforming promotes the achievement of your goals or other valuable ends. Beliefs that conform are useful tools. Again, there are some theoretical motivations to be an epistemic instrumentalist. Firstly, conforming to the epistemic standards is in fact a useful thing to do, at least in a lot of cases. So epistemic instrumentalism has a simple explanation about why we should conform to the epistemic standards and why we might learn to pay special attention to them. Secondly, this is not just a theory about mere semantics rather it is a metaphysical claim about the world (Kornblith, 1993, pp. 358-363).³⁴ It explains why epistemic normativity is important, about why it is philosophically interesting, regardless of the content of our terms or concepts. Thirdly, epistemic instrumentalism is an avenue to complete Quine's project of naturalising epistemology (Quine, 1969).³⁵ Epistemic instrumentalism, as it proposes a natural ontology, requires a much less suspicious metaphysics and epistemology than non-natural rivals (for some metaphysical and epistemic concerns about normativity, see: Joyce, 2001; Mackie, 1977; Street, 2006). This is a serious advantage over those non-natural rival theories (for rival theories see: Kelly, 2003; Parfit, 2011; Scanlon, 1998; Wedgwood, 2007 cited in Cowie, 2014, pp. 4004; also

³⁴ I return to this point in section 7.2.1.

³⁵ I return to this point in section 4.1.

see Enoch, 2011; Moore, 1903; Scanlon, 2014). And many philosophers have expressed sympathy for naturalising epistemology with epistemic instrumentalism (Dyke, 2021; Foley, 1987; Laudan, 1990a; Nozick, 1993). The promise of a natural metaphysics is one of, if not the, central motivation for epistemic instrumentalism. It will be useful to have these advantages in mind, because some versions of epistemic instrumentalism posit other types of reasons, which undermine these advantages.

However, recall Côté-Bouchard's case:

Spoiler Alert Max missed the finale of his favourite television series, which aired last night and revealed whether character X ends up dying or surviving. Given his passion for the show, Max really wants to avoid learning the answer to that question before he watches the rerun tonight after work. But to his dismay, one of his colleagues – who is convinced that Max did watch the show last night – comes up to him in the morning shouting: 'I knew it! I told you character X would die!' (Côté-Bouchard, 2015, p. 340).

On first pass, none of Max's goals are promoted by believing his colleague. But the epistemic standards demand that his colleague be believed. We can stipulate the colleague is an excellent testifier, having never led Max astray in the past. The problem for instrumentalists is that philosophers judge that Max has epistemic reasons to conform to the epistemic standards in this case.³⁶ But if none of Max's goals are promoted, there are no instrumental reasons to conform. So the instrumentalist has too-few-reasons to account for the epistemic reasons in this case. Of course, there are many possible too-few-reasons cases. So there will be many cases where the instrumentalist cannot find relevant instrumental reasons. This is **the too-few-reasons objection**.

Also recall that successful too-few-reasons cases are much more unusual than is commonly supposed. Finding relevant instrumental reasons in *Spoiler Alert* is not particularly hard. Supposing Max is anything like a normal human, he will not want to risk upsetting his colleague by not trusting his colleague's testimony. So Max has instrumental reasons to conform to the epistemic standards in *Spoiler Alert*.

I suspect this is not generally pointed out because philosophers can easily anticipate a response: idealising the case more. We can modify the case and stipulate that Max does not care about other

³⁶ At least, let's say for the sake of argument. Without this judgement, the motivation to be universal or non-instrumentalist is weakened. So universalists and non-instrumentalists should welcome it. In my experience, a lot of philosophers will make this judgement. But it would be worth checking experimentally just how uniform this judgement is among philosophers and non-philosophers.

people's feelings. But even in this modified case, Max still has relevant instrumental reasons to conform. If Max is anything like a normal human, he will want to be seen as reasonable, as not weird. Flouting the epistemic standards risks causing Max's colleagues to see him as unreasonable, as weird. But suppose Max does not care about other people's opinions, either. This modified too-few-reasons case is still not successful. If Max cares about getting a promotion or being invited to parties say, there are reasons not to risk others' seeing him as weird because it will lower the chances of him getting the promotion or the invitation. But suppose Max doesn't care about promotions or parties.... I am sure there is a successful too-few-reasons case (for such a case see section 4.2.1). My point is that it will be a lot more unusual, a lot more unrealistic than *Spoiler Alert*. It will have to involve an agent that has a very impoverished set of goals or some other type of unrealistic element. Too-few-reasons cases are not an everyday problem.

The strangeness of successful too-few-reasons cases is an issue for the objector (and the universalist, see section 3.6). Epistemic instrumentalism is a metaphysical theory. When the instrumentalist claims that 'epistemic reasons are instrumental reasons' they are claiming something about the world, not merely about the relationships between concepts. To use Derek Parfit's taxonomy, epistemic instrumentalism can be interpreted as a form of non-analytical naturalism (Parfit, 2011, p. 643).³⁷ Think of the instrumentalist's claim like you think about the claim 'water is H₂O'. So the evidential relationship between the too-few-reasons objection, that is the judgements, and epistemic instrumentalism is not simply that there is a counterexample to an analysis, like a Gettier case against a justified, true belief analysis of knowledge (Gettier, 1963). It is something more complicated. And as these metaphysical judgements are about more and more unrealistic scenarios, it becomes harder to justify treating those judgements as true or reliable.³⁸

Rather than question the evidential value of the judgements, or offer an error theory as I have done, **universal epistemic instrumentalists** instead offer an extremely bold reply in response to too-few-reasons cases: in all possible cases there are instrumental reasons to conform to the epistemic standards whenever we judge that there are epistemic reasons. This is the sense in which they are *universal* epistemic instrumentalists. The relevant instrumental reasons are everywhere. The

³⁷ This is one of the two relevant interpretations I offer in chapter 5.

³⁸ For arguments that odd cases make for poor evidence, see: (Kilov & Hendy, 2022; Machery, 2017; Van Inwagen, 1998). Whether you agree with the arguments or not, all I am pointing out here is that the epistemic basis for justifying judgements from esoteric cases is controversial at best, and the instrumentalist is not compelled to adopt that controversial epistemology, so it is questionable whether the cases create a strong objection in the first place. I will argue in section 3.6 that even more than the objector, the universalist needs the objection to carry significant epistemic weight.

fact they offer such a bold reply to such an esoteric set of cases is the first clue that something about this position is problematic. But the issues do not end there. In the next four sections I will characterise four different versions of the universal position and the serious problems with each version.

3.2. Cheap Universal Epistemic Instrumentalism

3.2.1. Schroeder's cheap reasons

The first theory I will consider is defended by Mark Schroeder (2007, Chapter 6). Consider the situation in *Spoiler Alert*. Even if Max does not know of any goals that will be directly promoted by conforming, there is some chance that the correct information will be important later. Suppose later that day, Max's crush wants to speak about the spoiler. Knowing about the spoiler will increase the chances that the conversation with his crush will go well. Suppose later that day, Max is listening to the radio and a competition comes on requiring knowledge of that spoiler. Knowledge of that spoiler may help Max win the competition. Who knows what the future will hold for Max?

Cheap universalists posit a glut of instrumental reasons. The central idea is that any possibility of promoting your goals gives rise to instrumental reasons and there is always *some* possibility that conforming to the epistemic standards will promote your goals. So instrumental reasons are very cheap. Cheap enough that, supposedly, there will always be relevant instrumental reasons in any too-few-reasons case.³⁹

³⁹ You might worry that 'some chance' does not do enough to account for epistemic reasons. Rather, you might hold, the conforming beliefs need to actually be helpful, not just have *some chance* of being helpful. This concern is about the distinction between objective and subjective reasons. For an introduction to the distinction see: (Vogelstein, 2012, p. 240). For the classic case involving this distinction see (Jackson, 1991, pp. 462-463). The gist of the distinction is this: suppose there is a cup of clear liquid in front of you. You are thirsty. You believe the cup is full of water. But, in fact, the cup is full of acid. There is a sense in which you have reasons to drink the liquid—you are thirsty, and you believe it's water; these reasons are your subjective reasons. There is a sense in which you have reasons to avoid drinking the liquid—it is in fact acid; these reasons are your objective reasons. Schroeder's idea is that you have subjective reasons to conform to the epistemic standards. This is the little bit of epistemic normativity in too-few-reasons cases.

You also might worry that cheap reasons are too cheap, that they do not provide enough normative force to account for epistemic normativity in too-few-reasons cases. The problem with pressing this worry is that the weight of epistemic reasons in too-few-reasons cases is controversial. There is no consensus about how weighty epistemic reasons are, how they trade-off against other reasons, or even if epistemic reasons are the only relevant reasons for determining the best all-things-considered belief. So an objection based on this worry would be rather weak because its foundations are so controversial.

I doubt that there are cheap reasons in every possible case. We can modify Spoiler Alert, making it extremely unrealistic. Suppose Max only cares about working through the day and watching his show; he does not care about anything else. Also, Max believes with certainty that an all-powerful god will intervene to ensure that conforming to or flouting the epistemic standards will not impact any of Max's goals and we can stipulate that this belief is true. In this case, his goals are not affected by not conforming to the epistemic standards. So there will not even be cheap reasons to conform. But even if there is a response to this modified case (such as this case is too unrealistic to take seriously) there is still a deeper flaw with the position.

3.2.2. The otiose objection

My objection has two parts: 1. positing cheap reasons creates a lacuna and 2. filling the lacuna makes cheap reasons otiose. Let's call this **the otiose objection**.

3.2.2.1. Cheap reasons create a lacuna

Philosophers make too-few-reasons judgements, that is they judge there are reasons to conform to the epistemic standards even when there are no apparent instrumental reasons to conform. Cheap universalism vindicates these judgements by claiming that there are cheap epistemic reasons in those cases, so philosophers are judging correctly.

But cheap reasons are not rare. Almost every standard in almost every situation has *some* possibility of promoting your goals. For example, there is some possibility that negligently swinging a chainsaw around your head will promote your goals; perhaps your crush will be so impress by your reckless behaviour that he agrees to go on a date with you. For another example, as you are reading this, there is some possibility that flouting the epistemic standards will promote your goals—let's call the associated reasons **cheap anti-epistemic reasons**. Possibly, our world is governed by a malicious demon; possibly, flouting the epistemic standards pleases the demon; so we *in fact* have cheap anti-epistemic reasons to flout the epistemic standards. However, we do not make too-few-reasons judgements about cheap anti-epistemic reasons. Indeed, we do not make too-few-reasons judgements in the vast majority of cases where we have some cheap reasons (Schroeder, 2007, chapter 5).⁴⁰ So, by the lights of cheap universalism, some of our judgements are mistaken. Perhaps

⁴⁰ Again, let us grant this to the universalist and the objector. If philosophers do make too-few-reasons judgements in response to other cases of cheap reasons, such as cheap anti-epistemic reasons, this would be evidence that too-few-reasons judgements are not particularly important for epistemic normativity. If we made too-few-reasons judgements

we judge correctly that there are reasons in too-few-reasons cases. But we judge incorrectly that we do not have reasons, in cases where we have all sorts of cheap reasons, such as cheap anti-epistemic reasons. So cheap universalism only vindicates our too-few-reasons judgements by positing a massive error in our judgements about when reasons do not exist. If you were motivated to adopt universalism to vindicate judgements, to ‘save the appearances’ as it is sometimes called, this is a problem.

Moreover, and more importantly for my objection, there is also a gap in the cheap universalist’s story that needs explanation. If we do not normally judge cheap reasons to exist, but we do normally judge cheap epistemic reasons to exist, what is special about cheap epistemic reasons? This lacuna is a deep flaw in the position.

3.2.2.2. Filling the lacuna makes cheap reasons otiose

To fill this lacuna, cheap universalists need to explain why we make the relevant judgements in too-few-reasons cases. But as those judgements were the motivation to posit cheap reasons in the first place, explaining why we make those judgements with some other feature of reality leaves nothing left for cheap reasons to explain. Filling the lacuna makes cheap reasons otiose.

For example, Schroeder—responding to a slightly different problem—tries to fill the lacuna (2007, Chapters 5, 6, 7). His explanation is extremely complicated. The gist is that reasons we have in virtue of a task itself are more important than any old cheap reasons. We have cheap epistemic reasons in virtue of believing alone. The reasons to conform to the epistemic standards are special in this way, so we pay extra attention to them. But the same is not true for our cheap anti-epistemic reasons. Those reasons we do not have in virtue of believing alone. So we make too-few-reasons judgements about cheap epistemic reasons but not cheap anti-epistemic reasons, because the former are reasons we have in virtue of believing alone.

I reject this explanation. But suppose it is right. If we have reasons to conform to the epistemic standards in virtue of believing alone, then we do not need to posit cheap reasons. In too-few-reasons cases, as people are believing, they should already have relevant instrumental reasons to conform to the epistemic standards. Whether there are additional cheap reasons is beside the point.

about all cheap reasons, it would make as much sense to be a universalist about the cheap reasons to negligently swing a chainsaw around your head as it would to be a universalist about epistemic reasons. In this case, universalism would not be an important part of the theory of epistemic normativity.

Schroeder has undercut the original motivation for positing cheap reasons by explaining too-few-reasons judgements with these in-virtue-of-the-task-itself special reasons.

This is not a coincidence. If cheap universalists explain why we treat cheap epistemic reasons differently to other cheap reasons, then that explanation also responds to too-few-reasons cases. This explanation cannot depend on epistemic reasons being cheap because we do not track all kinds of other cheap reasons, such as cheap anti-epistemic reasons. So a successful explanation of why we treat cheap epistemic reasons differently makes positing cheap reasons redundant. Cheap reasons do not actually move the debate forward at all because they require an independent response to too-few-reasons cases. So cheap reasons are otiose.

I have found this point difficult to communicate so it is worth summarising the argument. This is the otiose objection:

P1: Positing cheap reasons is supposed to reply to the too-few-reasons objection by asserting that there are always cheap reasons in every too-few-reasons case and we are tracking these cheap reasons with our too-few-reasons judgements.

P2: Positing cheap reasons incurs an explanatory debt: why do we track cheap epistemic reasons but not other cheap reasons, such as cheap anti-epistemic reasons?

P3: While there are many possible explanations, a successful explanation cannot depend on the cheapness of epistemic reasons because both types of reasons are equally cheap.

P4: Any successful explanation is also an independent response to the too-few-reasons objection because it will explain why we make the too-few-reasons judgement about epistemic reasons but not (most) other types of reasons.

Therefore,

C1: cheap reasons are otiose. They are supposed to reply to the too-few-reasons objection, but they require an explanation that constitutes an independent reply to the too-few-reasons objection.

Therefore,

C2: cheap universal epistemic instrumentalism is not the correct theory of epistemic reasons.

This is a serious problem for cheap universalism, one that cannot be resolved by rejecting a case, denying a judgement or finagling the theory. The too-few-reasons objection is a challenge to

explain why we treat epistemic reasons differently to other instrumental reasons. The answer is not that epistemic reasons can be cheap.

3.3. Rule Universal Epistemic Instrumentalism

3.3.1. Rule-based reasons

Note that up until now we have been putting forward a version of universalism analogous to act-utilitarianism. We have been looking for a way that conforming to the epistemic standards promotes at least one goal in every case. But we could offer a version of universalism analogous to rule-utilitarianism. Adam Leite offers such an account (2007).⁴¹

Why conform to the epistemic standards? Leite offers this answer (for him, conforming to the epistemic standards means believing in proportion with your evidence):

...it serves the overall satisfaction of one's goals (cognitive and perhaps otherwise) to accept a system of norms according to which believing on good evidence is epistemically laudable even if one lacks good instrumental reasons for so believing in the particular case (Leite, 2007, p. 463).

Rule universalists hold that we should adopt a policy of conforming to the epistemic standards, even if in particular cases conforming to the epistemic standards does not promote any specific goals. Considering *Spoiler Alert*, suppose Max does not have any traditional instrumental reasons to conform. The rule universalist maintains that Max would still have other instrumental reasons to conform to the epistemic standards as a rule. Supposedly, maintaining that rule will promote Max's future goals. So, supposedly, the rule provides instrumental reasons for Max, and instrumental reasons in all possible too-few-reasons cases.

I also doubt rule epistemic instrumentalism can provide relevant instrumental reasons in all possible too-few-reasons cases. We can imagine contrarian humans, or aliens if you prefer, whose only goal is to not conform to the epistemic standards at any time or as a rule. It is hard to see how these people have relevant instrumental reasons to conform. But even if there is a response to such a case, there is a deeper flaw in this position.

⁴¹ Leite is not himself an instrumentalist. However, Hilary Kornblith endorses a similar view, with epistemic normativity being based on what makes for a good epistemic policy, rather than on what achieves your specific goals in any particular case (Kornblith, 1993). So even if Leite does not endorse the view, other philosophers do.

3.3.2. The collapse dilemma

Rule universalism is like rule consequentialism; it is open to similar objections.⁴² Brad Hooker outlines an objection to rule consequentialism (2016, section 8). Rule consequentialism either collapses into act consequentialism or it is incoherent. Following a rule either maximises the good or it does not. If following the rule maximises the good, then act consequentialism also holds that you should follow the rule, in which case act consequentialism and rule consequentialism are equivalent in an important sense. If following a rule does not maximise the good, then the rule consequentialist seems committed to not maximising the good which at least appears incoherent.

There is a similar, although not exactly the same, issue with rule universalism. Suppose you are faced with a choice about whether to conform to the epistemic standards. If there are instrumental reasons to conform to the epistemic standards, then rule universalism and **act universalism** (let us call the alternative) are equivalent. They both tell you to conform. If we do well by following a rule, then that is what act universalism tells us to do. If there are no instrumental reasons to conform to the epistemic standards, then rule universalism is advocating that we conform without any instrumental reasons. While this may not be strictly incoherent, there is something odd about it.

Indeed, we can put a finer point on **the collapse dilemma** for rule universalism. The first horn: rule universalism collapses into act universalism. Suppose there are instrumental reasons for following a rule. The act universalist can respond to those reasons as much as the rule universalist can. So, we have not gained any ground by shifting to rule universalism. The second horn: rule universalism is not an instrumentalist theory. Suppose there are no instrumental reasons to conform as a rule. If there are reasons to conform in this case, they are not instrumental reasons. This implies epistemic instrumentalism is false. So, at least on first pass, adopting either horn is not a viable gambit for universal instrumentalists.

3.3.2.1. Avoiding the collapse: realistic rule universalism

Hooker offers a response on behalf of rule consequentialism (2016, section 8). His strategy is to include some pragmatic considerations when deciding on the correct rules. So, for example, when assessing the best rules, we should consider how hard they are to follow, to learn and to enforce. These considerations limit the complexity of rules that we can have. We cannot have a rule that

⁴² For an introduction to rule consequentialism see: Hooker (2016).

says: 'maximise the good' because it would be very hard to know what maximises the good in every situation and it would be very hard to enforce. Rules that are too specific will also be a problem. A rule like "whenever it costs you almost nothing, help an 83-year-old woman named Mabel across the street" might be easier to know when to apply, but rarely come up. So, according to Hooker, we should have rules that look more familiar. Rules like, 'don't lie', 'be nice to people', 'be charitable', etc. This list of rules would be longer than 'maximise the good' but would be more enforceable and easier to apply. So, the best rules will look more like the ten commandments rather than a general imperative to maximise the good.

Most importantly, for Hooker, once we accept the rules, the right action to take in any particular case is to conform with the rules. So, if the rules say 'don't lie', then the right action to take in any situation is not to lie.⁴³ Note that following the rule will not be the same as maximising the good. Perhaps there is a situation where lying could maximise the good. But, according to Hooker, rule consequentialists hold there is some reason to conform to a rule even when it does not maximise the good.

You might be tempted to make a similar response on behalf of rule universalists. Rule universalism does not collapse into act universalism. There will be some occasions where there are no instrumental reasons to conform to the epistemic standards. Rule universalism will still direct you to conform to the epistemic standards because conforming to the standards is what makes a belief right.

No instrumentalist should be satisfied with Hooker's response. Most importantly, Hooker's response to the collapse dilemma distinguishes rule consequentialism from act consequentialism by giving up on consequentialism. It holds that conforming to a rule gives rise to right action. When applied to rule universalism, the analogous claim is that conforming to the epistemic standards gives rise to right belief, regardless of whether there are instrumental reasons to conform or not. This response amounts to giving up epistemic instrumentalism. Of course, if there are instrumental reasons, act universalists can respond to those. But if there are no instrumental reasons, then rule universalism is no longer instrumentalist. While Hooker's response avoids the charge of inconsistency (which is how he originally characterised the second horn of the collapse

⁴³ There is a complication here. You might wonder what happens if not lying would lead to the deaths of one thousand people or even one person? We can put in a caveat at the end of all rules: ...unless doing so leads to catastrophe. Of course, this caveat creates more problems for rule consequentialism, problems I am not sure I can solve let alone solve in this essay. So, let's put aside catastrophes for this discussion.

dilemma) a Hooker-style response runs head-first into the second horn of the collapse dilemma, as I characterise it for rule universalists.

3.3.2.2. Avoiding the collapse: non-act based instrumental reasons

David Gauthier and Edward F. McClennen have postulated a special type of instrumental reasons. (Gauthier, 1984; McClennen, 1997). McClennen and Gauthier hold that sometimes there is a special type of instrumental reasons to conform to the rules that you have adopted, even when the act of conforming does not promote any of your goals in the traditional instrumental way. We might call this type of instrumental reasons **resolute-choice instrumental reasons**.

To explain resolute-choice instrumental reasons, let's review McClennen's analysis of the story of Odysseus tying himself to the mast of his ship. For those that do not know the story: when people hear the song of the Sirens they are tempted to go to the Sirens. If you go to the Sirens, they will kill you. However, Odysseus wanted to hear the song for himself, so he ordered his crew to tie him to the mast of the ship. When he heard the Siren song, he begged to be released. He wanted to go to the sirens. But his men refused. When the ship passed the Sirens, Odysseus no longer wanted to go to them.

McClennen's analysis of the story is that when Odysseus is not hearing the Sirens, he has a preference to not go to them. As Odysseus hears the Siren song, he has a preference to go to them. Here is the issue for decision theorists: at the time Odysseus is hearing the Sirens there are no traditional instrumental reasons to not go to the Sirens. Odysseus wants to go. Indeed, we can stipulate that at the moment he is hearing the song, that is all he wants to do; Odysseus does not care about his future self or his past self, about continuing to live. All he wants is to go to the sirens. So it seems all the instrumental reasons Odysseus has at that moment point in favour of going to the Sirens.

But McClennen disagrees. He argues that one has instrumental reasons to be 'resolute' (McClennen, 1997, p. 231). 'Resolute' means "regimenting future choice to the originally adopted plan." (McClennen, 1997, p. 231). So, according to McClennen, Odysseus has instrumental reasons to follow through with the plan of not going to the Sirens, even though when he hears the Sirens his only preference is to go. McClennen defends this type of resolute choice by noting that the resolute chooser does not have to expend resources on ensuring his future cooperation

(McClennen, 1997, pp. 324-235).⁴⁴ If Odysseus were a resolute chooser, he would not have to waste the resources of getting tied to the mast. Instead of tying himself to the mast to survive, he could have achieved the same result more efficiently by being a resolute chooser. I note that Gauthier has a similar view (Gauthier, 1984). The thought is that by being a resolute chooser, you give up short term gains for overall more gains. This is where the instrumental reasons to stick to your plans come from. The reasons to be a resolute chooser are resolute-choice instrumental reasons. If we just evaluate the act in isolation, Odysseus will have no instrumental reasons to not go to the Sirens. But once we stop focusing solely on the specific act, other types of instrumental reasons become apparent, at least, so say McClennen and Gauthier.

We can accept this view for the sake of argument. We can accept that the Max from *Spoiler Alert* should be a resolute chooser. But this does not help respond to too-few-reasons cases. Firstly, we can stipulate that Max never formed a plan to conform to the epistemic standards. It is just something he did a lot of the time. Secondly, we can stipulate that there is no overall gain for choosing to conform to the epistemic standards. There are still going to be too-few-reasons, even with the addition of resolute-choice instrumental reasons. Too-few-reasons cases have the wrong structure for a view like McClennen's or Gauthier's to help. So even if we accept that you should be a resolute chooser, rule universalists still do not have enough instrumental reasons.

3.3.2.3. Accepting the collapse: the otiose objection, part deux

Rather than trying to avoid the collapse dilemma, the universalist might just accept that rule universalism collapses into act universalism. If you think of rule universalism as a concession to too-few-reasons objectors, then accepting the collapse is a problem. If the rule universalist concedes that there are cases where there are no instrumental reasons to conform to the epistemic standards, then the fact that rule universalism is a type of act universalism makes the concession incoherent. Act universalism just is the claim that there are always some instrumental reasons to conform to the epistemic standards.

But rather than a concession, you might be tempted to view rule universalism as a change of emphasis. Traditionally, act universalism focuses on how a belief that conforms to the epistemic standards can promote your goals. You can think of rule universalism as making a similar point

⁴⁴ This true only if acting by force of will alone is itself costless. As someone struggling to submit their thesis, this does not feel true. Each minute I must force myself to keep writing takes a large emotional toll. Let us set this issue aside because I will argue that, even if this picture is true, it does not help the instrumentalist respond to too-few-reasons cases.

but focusing on how conforming in particular cases can help cultivate good habits. Rule universalists are worried that each deviation from the epistemic standards is a risk that you will deviate at some other point. And they are hopeful that each time you conform to the epistemic standards you increase the chances you will conform in the future. Indeed, in this way rule universalism complements act universalism. Sometimes conforming matters because the belief is important. Sometimes conforming matters because the conforming is important. Sometimes both are important. Moreover, this move forces too-few-reasons cases to be rather extreme to rule out all relevant instrumental reasons. A successful too-few-reasons case must rule out any possibility that conforming in one case will affect the chances of conforming in the next case. These too-few-reasons cases will not look very realistic. This is a promising way to construe rule universalism.

But universalists are not satisfied with pointing out that successful too-few-reasons cases are esoteric. They could have done that well before invoking rule-based reasons. Rather they want to claim that *in all possible* cases there are instrumental reasons to conform. This change-of-emphasis rule universalism inherits all the problems of act universalism—in particular the otiose objection. Considering a modified *Spoiler Alert*, imagine that Max is about to die and that he has good evidence that he is about to die. All he wants is to finish his day at work (say his day is ending in a minute) and go home to watch his show, spoiler free (say his home is right next door to his work). Does he have instrumental reasons to believe his spoiler-uttering colleague? Rule universalism says yes. Conforming in this case increases the chances of conforming in the next case. But Max is about to die. It is *extremely* unlikely that there are going to be any important cases after this.⁴⁵ So the probability that conforming will actually matter is extremely low. At the same time, there are many equally cheap reasons in this case, even other rule-based reasons. Perhaps blowing raspberries in this case will help in the next case where blowing raspberries is the right action to take. But we do not make the too-few-reasons judgements in response to reasons to blow raspberries. So the rule universalist, who considers their theory to be a version of act universalism, has the same explanatory lacuna as cheap universalism. We need some explanation about why we treat rule-based epistemic reasons so differently to other rule-based reasons. The fact the reasons are rule-based cannot be the basis of the explanation because there are plenty of equally cheap rule-based reasons we do not offer the too-few reasons judgement about. So we can offer the full otiose objection as presented in section 3.2.2 against this version of rule universalism, too.

⁴⁵ I imagine there is *some* chance, or at least a positive subjective probability, that in the next few minute an important decision will need to be made. Again, I am considering these chances to be extremely low, extremely cheap.

Therefore, rule universalism either is not a form of epistemic instrumentalism or is open to the otiose objection. It is not a step forward for universal instrumentalists, despite looking so promising.

3.4. Non-Authoritative Universal Epistemic Instrumentalism

3.4.1. Non-authoritative reasons

Recently there has been a lot of interest in a position about epistemic normativity that distinguishes between two types of reasons, *authoritative* and *non-authoritative* reasons⁴⁶ (Côté-Bouchard & Littlejohn, 2018; Maguire & Woods, 2020; Cowie, 2019 cited in A. Murphy, 2021). None of the philosophers who defend this view defend it as a version of epistemic instrumentalism, so nothing I say here is necessarily a criticism of the actual view. Nonetheless, we can entertain this view as a version of universal epistemic instrumentalism. Let's call this hypothetical position **non-authoritative universal epistemic instrumentalism**. In this section I articulate two criticisms. First, the central distinction at the heart of this position is not compatible with universal epistemic instrumentalism, so this hypothetical position is incoherent. Second, this distinction does not provide a good explanation for the too-few-reasons judgements.

But before I make my criticisms, let me introduce the view. As I said above, non-authoritative universal epistemic instrumentalism relies on a distinction between different types of reasons: authoritative and non-authoritative reasons. To identify the difference between these types of reasons Barry Maguire and Jack Woods point to the different functions of these reasons (2020, pp. 225-230). **Authoritative reasons** contribute to what you ought to do, all things considered. For example, 'that it will cause your imminent death' is an authoritative reason against jumping off a tall building; everything else being equal, it tells against jumping off a tall building. **Non-authoritative reasons** contribute to when something is correct, according to a standard. For example, it might be that castling early in a game of chess is correct from the standard of winning games of chess, but it might not be what you ought to do from an all-things-considered point of

⁴⁶ Maguire and Woods use the terms "authoritatively normative" and "nonauthoritatively normative" to refer to different types of reasons (2020, p. 217). Côté-Bouchard and Littlejohn use the terms "thinly normative" and "robustly normative" to refer to the same distinction (Côté-Bouchard & Littlejohn, 2018, p. 154). Other terms that are used to point to one side of the distinction are "genuine", "substantive" and "generic" (Maguire & Woods, 2020, n. 13).

view, say if you were teaching the game to someone, perhaps you should not introduce castling until they understand the other rules.

This distinction is supposed to offer a convincing reply to too-few-reasons cases (Maguire & Woods, 2020, pp. 239-241). In these cases, there are non-authoritative reasons that make it epistemically correct to believe in certain ways even if there are no authoritative reasons that make it all-things-considered rational to believe in that way. So, in *Spoiler Alert*, it is epistemically correct for Max to believe the spoiler, but he lacks authoritative reasons to believe it, so all-things-considered he ought not believe the spoiler. According to non-authoritative universal epistemic instrumentalism, there are still reasons in all too-few-reasons cases, they just lack authority. So we can, supposedly, vindicate the too-few-reasons judgement while preserving the core of an instrumentalist position, that instrumental reasons are doing the normative heavy lifting.

What gives rise to relevant standards? There are many possibilities. Here are two. First, it might be that humans can collectively or individually declare standards. This is implied by Barry and Woods (2020, pp. 217-218). They hold that we have non-authoritative reasons to play games in certain ways, like moving the king one space at a time during chess. This implies that we can generate non-authoritative reasons as easily as we can create games. So in a similar way that children can declare the standard of a novel game on the playground, humans may be able to declare standards and thereby create non-authoritative reasons. Second, it might be that the relevant standards were created by evolution (Nolfi, 2018). Perhaps beliefs have a proper function. A ‘proper function’ is the function that beliefs were selected for when they evolved. The relevant standards for belief, a proponent of the view might hold, is whatever fulfils that proper function. In this way, evolution may have bestowed upon us a suite of relevant epistemic standards, and these standards generate non-authoritative reasons, the reasons we are tracking when we make too-few-reasons judgements.

Whatever story you tell about how the standards are generated there is one constraint. It cannot be that the standards, and therefore the non-authoritative reasons, are generated by usefulness alone. Why? Because the distinction is supposed to provide extra explanatory resources that the traditional instrumentalist does not have; the one thing you cannot claim is that it is useful to conform to the epistemic standard in genuine too-few-reasons cases. So the non-authoritative reasons have to come from somewhere else. This constraint creates a problem for universal epistemic instrumentalism.

3.4.2. Non-authoritative universal epistemic instrumentalism is in tension with itself

Non-authoritative universal epistemic instrumentalism is either not universal or not instrumentalist. Either way, the position is in tension with itself. Are non-authoritative ‘reasons’, reasons? Answering either ‘yes’ or ‘no’ is incompatible with universal epistemic instrumentalism.

Suppose the non-authoritative instrumentalist answers: ‘no, non-authoritative ‘reasons’ are not reasons.’ This implies that even though there are non-authoritative ‘reasons’ in too-few-reasons cases, there are no reasons, instrumental or otherwise. The non-authoritative ‘reasons’ are supposed to vindicate too-few-reasons judgements; they are supposed to be the things we are tracking with our judgements. So if the things we are tracking are not reasons, then the universal instrumentalist is wrong; it is not true that there are always instrumental reasons in these cases. Sometimes there are *only* non-authoritative ‘reasons’ which on this horn of the dilemma, are not reasons, instrumental or otherwise. So answering ‘no’ is incompatible with the universal-part of universal epistemic instrumentalism because it suggests the relevant reasons are not universal.

Suppose the non-authoritative instrumentalist answers: ‘yes, non-authoritative ‘reasons’ are reasons.’ This implies that the reasons in too-few-reasons cases are not instrumental reasons. Recall that the only constraint on the origin of non-authoritative reasons is that they do not come from being instrumentally useful. That constraint creates a problem on this horn. Sometimes in too-few-reasons cases there are *only* non-authoritative reasons, which on this horn of the dilemma, are not instrumental reasons. So answering ‘yes’ is incompatible with the instrumental-part of universal epistemic instrumentalism because it implies the relevant reasons are not instrumental reasons. This dilemma shows that the hypothetical position ‘non-authoritative universal epistemic instrumentalism’ is in tension with itself. Either the position is not universal or it is not instrumentalist so in neither case is it universal epistemic instrumentalism.

Indeed, this dilemma is a version of the collapse dilemma I offered against rule universalism. The non-authoritative distinction does not introduce any new resources to explain the short fall of instrumental reasons that too-few-reasons cases point to. So either it will collapse into the other instrumental views, where there are no relevant reasons in too-few-reasons cases or it draws upon

non-instrumental reasons making it not an instrumentalist position. Making a distinction does not help you find the extra instrumental reasons universalists need.

3.4.3. The distinction does not help with too-few-reasons cases

In section 3.2.2, I argued that positing cheap reasons incurs an explanatory debt that makes cheap reasons otiose. Something similar happens in the case of non-authoritative reasons. However, there are important differences between the philosophers who advocate for cheap reasons (Schroeder, 2007) and those who advocate for non-authoritative reasons (Côté-Bouchard & Littlejohn, 2018; Maguire & Woods, 2020; Cowie, 2019 cited in A. Murphy, 2021). Cheap reasons were posited to explicitly handle too-few-reasons cases (Schroeder, 2007, chapter 6) whereas non-authoritative reasons have other motivations and only incidentally are supposed to respond to too-few-reasons cases.⁴⁷ So unlike cheap reasons, even if non-authoritative reasons fail to produce a response to the too-few-reasons objection, they may not be completely otiose; they may play other important theoretical roles.

Nonetheless, non-authoritative reasons fail to produce a convincing response to the too-few-reasons objection. The response is that in the relevant cases there are still reasons, just not authoritative ones. So our judgements are correct. But as with cheap reasons, non-authoritative reasons are not rare and this leaves the position open to the otiose objection.

Consider games. We may have non-authoritative reasons to play games in certain ways, like moving the king one space at a time during chess (Maguire & Woods, 2020, pp. 217-218). But I take it as common ground that there are enough instrumental reasons to account for the reasons to move the king one space—that is we do not think there are too-few-reasons for the normativity of games, like chess and Twister.⁴⁸ Indeed, if we can arbitrarily produce standards, say according to the Made-Up Standard: it is correct that you should now be hopping on one leg, then non-authoritative reasons can be extremely common.

⁴⁷ For example, Maguire and Woods are primarily concerned with providing a response to objections based on the idea that you cannot weigh epistemic reasons with other instrumental reasons (2020, p. 214). So it is possible that while non-authoritative reasons fail to produce a convincing response to the too-few-reasons objection, they might provide a convincing response to these other objections. Although it also is possible that the otiose objection generalises.

⁴⁸ Again, let's grant the universalist and the objector (see n. 15, 26).

But as I said above, perhaps the epistemic standards, and therefore the epistemic non-authoritative reasons, are produced in some other way, by evolution or some other mechanism. Nonetheless, I can offer a schema for generating the same problem.

1. The first step is to locate some standard produced by that same mechanism that we do not make too-few-reasons judgements about. In the case of a proper function, we might point to the proper function of my hand, for example. Suppose the proper function of my hand is, at least in part, to dextrously manipulate objects with my fingers.⁴⁹
2. The second step is to generate a too-few-reasons-like case, a case where we have no instrumental reasons to achieve the relevant standard. For our proper function example, imagine I need to fall asleep to live, and dextrously manipulating objects with my fingers would interfere with me falling asleep. Even though according to the standard generated by the proper function I have non-authoritative reasons to dextrously manipulate objects, I have no instrumental reasons to do that.
3. The third step is to note that we do not make too-few-reasons judgements about the too-few-reasons-like case. For our example, I do not need to posit extra reasons in this case for me to dextrously manipulate objects with my fingers. Indeed, I bet that for most standards, once it is no longer useful for me to achieve that standard, in general people judge that there are no relevant reasons to achieve the standard. There are exceptions, such as ethical and epistemic standards, but these are very rare, or at least I conjecture.
4. Then we note that we are making a mistake somewhere. If we are tracking the non-authoritative reasons in canonical too-few-reasons cases by making too-few-reasons judgements, then we are making a mistake in not making similar judgements about non-epistemic standards in too-few-reasons-like cases. One way or another we are treating the standards generated by whatever mechanism generates the relevant standards differently. Again, part of the challenge too-few-reasons judgements offer is to explain what is special about epistemic standards. And that there are non-authoritative reasons to conform to them is not what makes them special; if there are

⁴⁹ This, probably, is not the proper function of a hand. Evolution is an awesomely complex process. Features of the body can be selected for many different, even competing, functions, including trade-offs between those functions. But, to the extent the proper function theory makes sense, I take it that there is some much more complicated true story that follows similar lines to the one I tell above. I leave it to defenders of a proper function theory to tell the true, detailed story about proper functions.

any non-authoritative reasons then there are heaps of them, many that we regularly judge not to exist.

Strictly speaking this schema leaves open a possible response. Non authoritative instrumentalists could identify a mechanism that generates epistemic standards but no other standards we fail to make too-few-reasons judgements about. I doubt it is possible to make such a response plausible because such a response is heavily constrained.

Firstly we have a pretty good idea that these standards evolved, either in a biological process as Sharon Street argues (2006, 2009) or perhaps in some cultural process. As the example above suggests, evolution does not only produce standards that we make too-few-reasons judgements about. So we need an additional mechanism to make this response. But whatever additional mechanism is proposed, it has to fit in with the evolutionary story about how we actually learned the standards.

Secondly, we also need a story about why humans are tracking standards produced by that mechanism. There are heaps of standards. And there are heaps of mechanisms by which standards are produced. There needs to be some explanation about why the particular mechanism is special, such that we developed too-few-reasons judgements about the standards it produced.

Thirdly, supposing there are beliefs, the mechanism needs to be extremely modally robust. Why? As I noted in section 3.1, successful too-few-reasons cases are already extremely odd. So to the extent that you are using those odd cases to motivate universal epistemic instrumentalism, you are already aiming to develop a theory that produces epistemic reasons in all (or perhaps almost all) cases where agents believe. Otherwise, you are not a universal instrumentalist. If the mechanisms or standards are not modally robust (given the existence of beliefs) then there will be too-few-reasons cases with no non-authoritative reasons. If standards do not robustly exist given beliefs, we can imagine a human on a far planet or some proto-human in the deep past or some engineered human in our future, who has beliefs but does not have the relevant environment to produce epistemic standards. Non-authoritative universalism will have to reject the idea that there are epistemic reasons for those agents in these cases, even though they are structurally similar in all other respects to too-few-reasons cases. Therefore the mechanism needs to exist robustly, so that

when we imagine a proto-human in a too-few-reasons case (etc.) the relevant standards already exist.

Fourthly, it is plausible that humans make too-few-reasons judgements about other standards such as moral, aesthetic or legal standards (another issue where some experimental evidence would be most welcome). This additional mechanism needs to also give rise to those standards, too. If it does not, then we can issue a similar objection but this time about the class of mechanisms that give rise to the standards we make too-few-reasons judgements about.⁵⁰

So the possible response available to non-authoritative instrumentalists requires some mechanism that plausibly generates our epistemic standards, that is plausibly integrated with an evolutionary story about how we learnt about those standards, that is special in some respect so humans pay attention to the standards produced by that mechanism (but that nevertheless does not give rise to authoritative reasons), that exists robustly across modal space and robustly produces those standards (given the existence of beliefs) and that produces other standards we make too-few-reasons judgements about. These requirements on the response are very constrained, so constrained that I doubt there is a plausible story to tell that checks all the boxes.

Setting aside such a response, this argumentative schema is a problem for anyone that holds the non-authoritative distinction somehow vindicates our too-few-reasons judgements. Perhaps we are right to judge that there are reasons in too-few-reasons cases, because we accurately track the non-authoritative reasons but then we are wrong to judge that there are no reasons to conform to the other standards generated by the relevant mechanism. One way or another, we cannot save all the appearances.

Moreover, again setting aside the possible response to my argumentative schema, once we make

⁵⁰ Strictly speaking there is a similar response available to this point too. Perhaps some mechanism gave rise to only epistemic standards. And perhaps some other mechanism gave rise to only moral standards. And perhaps some further mechanism gave rise to only.... And none of these multiple mechanisms produce any standards that we do not make too-few-reasons judgements about.

But my point about the unlikeliness of such a story is even more problematic for this multiple mechanisms response. To be plausible and not run afoul of the objection, none of the mechanisms can produce standards that we do not make too-few-reasons judgements about. All the mechanisms need to integrate with the stories about how we learned about the relevant standards, evolution or otherwise. They all need to be in some sense special so we can explain why we would track them with too-few-reasons judgements. They all need to be in some sense modally robust so they give rise to non-authoritative reasons in all too-few-reasons-like cases. They all need to be in some sense unified, to explain why we would use a similar judgement to track the standards. So this multiple mechanisms response is many times more unlikely than the one mechanism response. This is why I suggest the proposed mechanism needs to give rise to the other standards we make too-few-reasons judgements about or it will be vulnerable to the objection.

the distinction between authoritative and non-authoritative reasons, we still need an explanation about why philosophers treat epistemic standards differently to other standards. Whatever that explanation is, it cannot depend on non-authoritative reasons because those reasons are too common. For example, perhaps the relevant difference is that conforming to the epistemic standards is almost always useful.⁵¹ Most of the time we have authoritative reasons to conform so it is natural to expect that we pay special attention to the epistemic standards. But this just is the non-universalist instrumentalist response to too-few-reasons cases (Sharadin, 2018; section 2.3). I agree the overwhelming usefulness explains why we would make too-few-reasons judgements, an efficient if mistaken overgeneralisation. But then there is nothing left for the distinction to explain in too-few-reasons cases. Granted that the distinction makes too-few-reasons judgements true (at the cost of implying anytime we do not make too-few-reasons judgements in response to too-few-reasons-like cases we are mistaken). But the explanation of why we make the judgement in response to epistemic standards but not in response to other standards has nothing to do with tracking the non-authoritative reasons. Indeed, this story implies that a fortunate coincidence took place. The reason we treat epistemic standards as special is because of the overwhelming (but not universal) usefulness of conforming. But the too-few-reasons judgement is made true by the presence of unrelated non-authoritative reasons. This is like Bertrand Russell's case of the broken clock fortunately telling you the right time (Russell, 1948 [2009]). Sure, your belief is true, but you just got lucky. It is not really vindicating the judgement. Indeed positing that this fortunate coincidence took place is a cost for a theory.

This problem is a version of the otiose objection. Too-few-reasons judgements tell you that something is special about epistemic normativity. The purported explanans of the judgements, non-authoritative reasons, do not tell you what is special about epistemic normativity. And an additional explanation about what is special, such as the overwhelming usefulness of conforming to epistemic standards, makes the purported explanans otiose, at least when it comes to responding to the too-few-reasons objection.⁵²

⁵¹ This was suggested by an anonymous reviewer of the paper version of this chapter.

⁵² I remain uncommitted about whether you can combine non-universal embedded epistemic instrumentalism to explain too-few-reasons cases with the non-authoritative distinction to explain other cases. Although as I argue above, this combination implies that a fortunate coincidence took place where our judgements were caused by something (usefulness) unrelated to what makes the judgements true (the presence of non-authoritative reasons). So at a minimum, this combination of theories carries this cost.

3.5. Common Goal Universal Epistemic Instrumentalism

3.5.1. Common goal reasons

So far, the versions of universalism we have been considering hold that there are instrumental reasons to conform to the epistemic standards, regardless of the goals you have (setting aside extreme cases). Let us change tack. Now we will consider if there is some universal goal that is promoted by conforming to the epistemic standards. If we all have such a goal, then we will always have instrumental reasons to conform to the epistemic standards.

Richard Foley offers the simplest version of common goal epistemic instrumentalism, that we all have the goal of conforming to the epistemic standards (for him, the epistemic standards are now believing the truth and now not believing a falsehood):

Might not [some person] lack the goal of now believing truths and now not believing falsehoods? Perhaps not. Perhaps this is impossible, at least if the notion of a goal is understood in a suitably broad sense. After all, the vast majority of us are intrinsically curious about the world; we intrinsically want to have true beliefs (Foley, 1987, p. 11).

If we all had that goal, then we would all have instrumental reasons to conform to the epistemic standards.

Indeed, we can make the required sense of having a goal quite minimal. Ernest Sosa holds it is at least plausible that “for any arbitrary belief of ours, we would prefer that it be true rather than not true, other things equal” (Sosa, 2003, p. 157). This minimal claim would be enough to establish common goal universalism. So long as you also held that satisfying preferences gave rise to instrumental reasons, then there would always be instrumental reasons to conform to the epistemic standards because conforming always promotes the achievement of one of your preferences.

This mere preference view has an analogous lacuna to cheap universalism, so it is open to a version of the otiose objection too. However plausible it is to claim we all have a mere preference to conform, it is as plausible to claim that all other things equal, we prefer to win a game of chess rather than lose. But chess reasons and epistemic reasons are treated very differently. So, the mere preference view still needs that additional explanation that cheap universalism needed. What is the difference between epistemic reasons and other instrumental reasons? And the explanation cannot depend on there being a mere preference for conforming because it is as plausible to say we have a mere preference for many goals, but we do not judge the resultant reasons to exist the way we

judge epistemic reasons to exist. A mere preference does not explain the difference between epistemic reasons and other instrumental reasons.

That being said, it is not likely that we all have a mere preference for conforming. Humans exhibit a massive amount of variation (Hull, 1986). As Nathaniel Sharadin argues (2018, pp. 3794-3796), common goal universalism posits an unlikely uniformity in nature.⁵³ For universalists, it will not do if the ‘vast majority’ of people intrinsically value truth or prefer truth. Universalism requires that every single person prefers truth. And it is hard to not believe that there are some *avant-garde* libertines, students of Nietzsche or just psychologically damaged people who are utterly indifferent to truth other things equal. Perhaps there are enlighten monks without preferences:

Imagine a Zen Buddhist monk who has credences but no preferences. Gazing peacefully at the scene before him, he believes that Mt. Everest stands at the other side of the valley, that K2 does not, and so on. But don’t ask him to bet on these propositions, for he is indifferent among all things (Eriksson & Hájek, 2007, p. 194).

In general, humans are varied. Not all of us have two eyes. Not all of us want to keep on living. Not all of us like our children. There are even some humans who do not enjoy philosophy! So we need incredibly strong evidence to believe a specific universal uniformity in humans, especially when the supposed uniformity is merely a contingent fact about our psychology.

3.5.2. Côté-Bouchard’s common goal of acting for genuine reasons

Côté-Bouchard employs a different strategy for establishing a common goal (2015). The strategy is to argue that agents necessarily have a goal that is promoted by conforming to the epistemic standards, namely the goal to act for genuine reasons. This is a reasonable strategy to try because the element of necessity is exactly what the above versions of universalism lacked.

Rather than analyse Côté-Bouchard’s argument, I will just grant the central claim. Let us grant that ‘agents’ constitutively have the goal to ‘act for genuine reasons’ and satisfying that goal often requires the agent to conform to the epistemic standards (Côté-Bouchard, 2015, pp. 350-351). Of course, conforming does not always promote that goal. We can imagine that by trying to conform, we end up acting for only apparent reasons because the evidence was misleading. In response, Côté-Bouchard claims that *as a rule* conforming to the epistemic standards promotes the goal of

⁵³ Sharadin calls common-goal embedded epistemic instrumentalism ‘the special interests strategy’.

acting for genuine reasons (2015, p. 351). As I have argued above in section 3.3.2, rule universalism faces the collapse dilemma. That is reason enough to reject this position.

But there is a new problem with this position, too. No constitutive claim could make common goal universalism plausible. The constitutive claim is supposed to be evidence that almost all humans in fact have the requisite goal. Côté-Bouchard's argument is supposed to proceed: almost all humans are agents; (constitutively) agents desire to act for genuine reasons; so almost all humans have a desire to act for genuine reasons. However, this is a roundabout way to argue that almost all humans have a goal. If a lot of humans did have this goal, there would be a lot more evidence of this fact, such as a lot of people declaring that they have such a desire or that they are working towards it, or that something thwarted their attempts to get it, or that attaining it is just around the corner, just like they do with widespread goals to be healthy and wealthy. I suspect that most humans have the goal to continue breathing. When the goal to continue breathing is thwarted, even temporarily, it is newsworthy; many people have said something to the effect, "I almost choked on that chip!" I take it the goal to act for genuine reasons is supposed to be even more widespread than the goals to be healthy and wealthy, even more widespread than the goal to continue breathing. But almost no one mentions that they have this goal to act for genuine reasons, or mentions anything about this goal. They do not mention it when it is achieved nor when it is thwarted. So even if Côté-Bouchard is right about the constitutive claim, given the lack of evidence that people by and large have a common goal, we should hold that people by and large are just not 'agents' in Côté-Bouchard's sense. So this version of common goal universalism has the same serious problem as all types of common goal universalism: humans don't share a common goal!

3.6. Universalism's Future

Universalism has not been successful. Will it succeed in the future? I conjecture that it will not. The space for a successful universal theory is very narrow, perhaps non-existent. Moreover it starts with a very unlikely claim and it does not have much evidence to support that claim. So I doubt there will ever be a successful universal instrumental theory.

We have analysed four different versions of the universal position. But all these theories are trying to address the too-few-reasons objection with the same broad strategy. The strategy is to find a

source of previously unaccounted for instrumental reasons. Let's call this strategy **extra-reasons universalism**. This strategy has an incredibly narrow window of success.

Firstly, in general, instrumentalists do not lack reasons. The fact that successful too-few-reasons cases must be so unusual implies that, in everyday cases, there are instrumental reasons to conform to the epistemic standards. The problem is that whether something is useful is a contingent fact. So there will be some possibilities where it is not useful to conform. You do not change this by describing an additional source of contingent reasons. Perhaps the possibilities will become more fanciful, involving gods, agents with few goals or aliens but they will still be possible. You need extra *necessary* reasons, not extra *contingent* reasons.

Secondly, to be an instrumentalist theory, the extra reasons need to be instrumental. But as we identify extra reasons originating from sources other than usefulness, such as reasons generated by rules or by standards, the reasons tend to lose their instrumentalist character. And if the extra reasons are not instrumental reasons, then the position is no longer instrumentalist and it loses the theoretical advantages of that position. A problem for extra-reasons universalism is that there is not some vast untapped reservoir of instrumental reasons. This is why the collapse dilemma often manifests. By pushing to identify extra reasons, universalists end up positing extra non-instrumental reasons or fail to posit any extra reasons at all.

Thirdly, the extra reasons cannot be ubiquitous. If the extra reasons are everywhere then you cannot explain why we make the too-few-reasons judgement about epistemic reasons specifically. This is what gives rise to the otiose objection. And indeed, positing ubiquitous extra-reasons implies we are making a massive error by not having the too-few-reasons judgement more generally. So if your universalism posits extra reasons that are ubiquitous, then your position requires an independent response to the too-few reasons objection, making the extra reasons otiose.

So a successful extra-reasons universalism posits some type of extra necessary reasons, that are instrumental in character, that are not ubiquitous. That is a tall order. So the question that now faces us: is it possible to thread the needle by identifying some relevant extra necessary non-

ubiquitous instrumental reasons or is there some other strategy that might be successful? I doubt it.

Universalism makes this claim: conforming to the epistemic standards is *always* useful. There is a lot of inductive evidence against this (see section 2.4). For almost every set of standards,⁵⁴ conforming to them is not always useful. Conforming to the rules of Twister does not help plan a typical funeral. For every tool, using it is not always useful. Wielding a knife does not help lead a relaxing meditation. And because the epistemic standards are not obviously different to other standards and tools, we should set our credence very low that conforming to the epistemic standards is *always* useful.

The main, perhaps only, source of evidence for universalism is our too-few-reasons judgements. But you must be extraordinarily optimistic about the reliability of our judgements to consider them especially evidentially weighty. The content of these judgements might be semantic,⁵⁵ in which case you must believe that the content of our concepts is an incredibly accurate guide to metaphysical reality, even in very unusual cases, to overcome the initial implausibility of the universalists claim. Alternatively, the content of these judgements might be metaphysical,⁵⁶ in which case you need to believe our metaphysical insight is incredibly accurate, even in very unusual cases, to overcome the initial implausibility of the universalist's claim. Finally, the content of the judgements might be some third thing, in which case you must believe this hitherto unspecified faculty is incredibly accurate, even in very unusual cases, to overcome the initial implausibility of the universalist's claim. Without believing these judgements are especially evidentially weighty, a belief that I bet is not particularly widespread, the universalist is stuck with an implausible claim and little evidence to move the debate forwards.

In this way, the universalist depends on too-few-reasons judgements even more than the objector. The objector is arguing that epistemic normativity is something other than instrumental normativity because the epistemic standards always give rise to epistemic reasons, but conforming is not always useful. It is not so unlikely that epistemic normativity is something other than instrumental normativity and it is quite likely that conforming to the standards is not always useful. So the objector does not need the judgements to be that evidentially weighty. On the other hand, the universalist needs the judgements to tell us that conforming to the epistemic standards is *always*

⁵⁴ Why not all?—I anticipate some gerrymandered set of standards like: 'do only what is useful'.

⁵⁵ Broadly speaking, this is consistent with the method of cases presented by Frank Jackson (1998).

⁵⁶ Broadly speaking, this is consistent with the method of cases presented by Timothy Williamson (2007).

useful, a wildly unlikely claim. So I doubt universal epistemic instrumentalism will ever be successful.

4. Why Accept Embedded Epistemic Instrumentalism?

In this chapter, I argue that there is a lot of evidence in favour of embedded epistemic instrumentalism. Most of my positive arguments for embedded epistemic instrumentalism apply to almost any instrumental theory (section 4.1). And most of my defensive arguments apply to almost any non-universal instrumental theory (section 4.2). However, I take the preceding chapters to motivate focusing only on embedded epistemic instrumentalism.

My argument is different from the usual arguments for instrumentalism. The usual arguments assume some kind of reflective equilibrium. Philosophers note that epistemic instrumentalism is broadly extensionally adequate—it delivers similar verdicts as our intuitions in response to many thought experiments. Also they hold that the disvalue of its extensional inadequacy—the few cases where it fails to deliver similar verdicts as our intuitions—is compensated for by the value of other properties such as its parsimony or explanatory power. For arguments like this see: (Côté-Bouchard, 2015; Dyke, 2021; Schroeder, 2007). Using this method, my own judgement is that epistemic instrumentalism does well but is not a clear standout from other accounts; it is just one of the accounts you might settle on using reflective equilibrium.

I aim to do something slightly different from appealing to reflective equilibrium. As I stated above, one of the central motivations to be an instrumentalist is the naturalistic metaphysics and epistemology that it offers (cf. Quine, 1969). This is one of the properties that may compensate for extensional inadequacy. In what follows, I will explain this motivation. Importantly, I am not offering an account of naturalism in general, but rather I aim to identify the features of embedded epistemic instrumentalism philosophers might be referring to when they claim that epistemic instrumentalism is naturalistic.

Nevertheless, without appropriate survey data, my best guess about the central complaint of non-naturalists is that what exactly counts as a ‘naturalistic’ theory is vague at best and confused at worst so there is little value in claiming that epistemic instrumentalism is naturalistic. In response to this worry, I banish any talk of ‘naturalism’ from my arguments in this chapter.

Rather I discuss how embedded epistemic instrumentalism has features that make it more likely to be true than if it did not have those features, all other things equal. As these features make it more

likely to be true, these features are evidence in favour of it.⁵⁷ So rather than engaging in reflective equilibrium, I consider myself to just be pointing at some of the evidence in favour of embedded epistemic instrumentalism. My claim in the second half of this chapter is that the evidential value of objections offered against embedded epistemic instrumentalism is quite low. So this chapter overall makes the case for holding that embedded epistemic instrumentalism stands out in a crowded field; it has a lot of evidence in favour of it and not much evidence against it.

4.1. The Evidence For

In the following sub-sections (4.1.1 and 4.1.2), I identify two features of embedded epistemic instrumentalism that count as significant evidence in its favour. The first is that the central positive claims should be uncontroversial. The second is that embedded epistemic instrumentalism does not demand much from the world and explains a lot of phenomena, a property I call ‘model-invariance’. In section 4.1.3, I summarise the case for embedded epistemic instrumentalism. While other theories may have some of these features to some degree, I claim that the combination and the degree to which the theory has these features provides a compelling evidential basis in favour of this theory.

4.1.1. The uncontroversial foundation of embedded epistemic instrumentalism

In this sub-section, I argue that any reasonably complete account of epistemic normativity will have to involve usefulness as a central element. There are at least three ways usefulness will have to figure in any broad account of epistemic normativity. 1. Conforming is extremely useful; all accounts of epistemic normativity have to explain the usefulness of conforming to the epistemic standards. 2. All accounts will invoke usefulness to explain how humans learned the epistemic standards. 3. All accounts will have to invoke usefulness to explain how the importance of the epistemic standards manifests in our day-to-day lives. So all non-instrumental accounts are really instrumentalist-plus accounts, needing to invoke usefulness to explain various aspects of epistemic normativity. The fact that usefulness is intimately involved in explaining various aspects of epistemic normativity (regardless of what else you add) increases the likelihood, and therefore is evidence that, epistemic normativity is usefulness.

⁵⁷ I am talking loosely. You may think that only beliefs or only propositions count as evidence. But what I say can be translated. If you hold that only beliefs count as evidence, then I am saying the belief ‘embedded epistemic instrumentalism has this feature’ is the evidence. But let us suppress this complication because it causes awkward expressions and does not matter for the discussion.

Importantly, I am not claiming that other theories cannot explain these phenomena. Of course, non-instrumentalists can combine their theory of epistemic normativity with instrumental normativity to produce a hybrid account, as Thomas Kelly does (2003, pp. 634-637). I am just claiming that a relatively complete account of epistemic normativity will have usefulness as a central theme. The only question is whether there is any need to posit anything more than usefulness to account for the relevant phenomena. It is the austerity of embedded epistemic instrumentalism that is controversial. But let us set the austerity aside for this sub-section and focus on the non-controversial core of embedded epistemic instrumentalism.

4.1.1.1. Explaining the fact that conforming is extremely useful

As I claimed in section 2.1, conforming to the standards is useful in at least a wide range of circumstances. The instrumentalist claims that this usefulness is what makes the epistemic standards compel. It is the thing that we are tracking with philosophers' semi-technical talk of 'epistemic normativity' and 'epistemic reasons'. That conforming is robustly useful is all embedded epistemic instrumentalism requires from the world; it is the metaphysical base of the theory.

However, the fact it is robustly useful to conform to the epistemic standards is itself something to be explained. Indeed, conforming is so useful that non-instrumental accounts are postulating a fortunate coincidence took place (Cowie, 2014; also see section 2.3). The coincidence is the way usefulness is strongly correlated with an extra element, such as a brute normative fact. It just so happens that in almost all cases, it is useful to conform to the epistemic standards. But, says the non-instrumentalist, it also happens that in all these cases there is some extra element, such as a brute normative fact, that also makes the epistemic standards compel. On non-instrumental accounts of epistemic normativity, in almost all cases, the compelling nature of the epistemic standards is over-determined. Postulating this coincidence is a cost to all non-instrumental accounts. The fact that embedded epistemic instrumentalism does not postulate this striking coincidence is an advantage embedded epistemic instrumentalism has over other accounts.

4.1.1.2. Fitting-in with evolution

Embedded epistemic instrumentalism integrates well with our evolutionary history. It does so in two related ways. Firstly, it is not subject to evolutionary debunking worries. Secondly, non-instrumental accounts also have to appeal to usefulness to explain how we learned about the epistemic standards.

There is a large and complex literature on evolutionary debunking arguments (Bogardus, 2016; Evers, 2015; Isserow, 2019; Joyce, 2006, 2016; Matt Lutz, 2018; Plunkett, Buchak, & Lombrozo, 2019; Street, 2006, 2009; Vavova, 2014, 2015, 2021). So what I say here will not capture all the subtleties of the dialectic.

But as a first pass, evolutionary debunking arguments note that our standards, epistemic or otherwise, evolved. The literature focuses on biological evolution, but the relevant standards might have been produced by cultural evolution, or some complex combination of evolutionary processes. Whatever the exact evolutionary process, there is an issue for a certain type of non-instrumental realist normative theory. Any theory that suggests epistemic normativity is produced by something other than usefulness, has to posit another coincidence. We learned about the epistemic standards because they promoted survival (or more generally, because they were useful for some end that gave rise to a selection pressure in an evolutionary process). But, for non-instrumental accounts, something other than usefulness gives rise to epistemic normativity, such as a basic fact about the epistemic standards or the essence of beliefs or whatever your non-instrumental account posits. Indeed, had some other set of standards been useful, those are the ones we would have evolved to appreciate. So it is extremely fortunate that the useful standards are also the standards that have the extra element, such as brute normative properties. Pointing out that certain accounts of epistemic normativity require a coincidence to take place in our evolutionary history is known as an **evolutionary debunking argument**.

Note that instrumentalists do not have this issue. They say that epistemic reasons are a species of instrumental reasons. Indeed, on the embedded picture I endorse, conforming to the standards is highly useful for a wide variety of ends, including surviving. So it is expected that a process that is sensitive to practices that promote survival would locate the epistemic standards. Rather than providing reasons to doubt, evolutionary considerations support an instrumental account.

Moreover, even the people rejecting the debunking arguments are still assuming that these standards evolved. This is because the only plausible story about how humans learnt these standards involves evolutionary processes in some capacity. And the way evolution learns things is by those things being useful for survival. Indeed, unlike for moral standards, the direct usefulness of conforming to the epistemic standards is hard to doubt. While I accept that we lack specific evidence for any particular story about how certain cognitive capacities evolved (Isserow, 2019), it is nonetheless a good bet that the usefulness of conforming will be an important part of the

evolutionary story of how humans came to recognise epistemic standards. This is another way usefulness will be a theme in an account of epistemic normativity.

4.1.1.3. Explaining the manifestations of epistemic normativity

It is common ground that epistemic standards are important. But consider the ways this importance manifests itself. People feel compelled to conform. People feel ashamed for not conforming. People spend years learning to conform to the epistemic standards in specific disciplines. People call-out other people who are not conforming. People try to justify their own beliefs. People expend massive amounts of resources conforming. These things happen in the world. These things interact with physical, causal systems.

Accounts like a brute-fact account of epistemic normativity or an essence-of-beliefs account will need to be supplemented with something to explain how the brute facts or essences interact with the world to produce all these physical phenomena. Similarly, even if you hold that epistemic normativity is about humans endorsing certain practices or beliefs, like the non-cognitivist does, you still need to explain how humans developed to want to endorse certain epistemic practices. Usefulness will be a part of the supplement. Usefulness reduces to physical facts. Usefulness, through extremely complex causal pathways, can move creatures to conform, rebuke, feel guilty and expend resources in ways brute facts or mere endorsement cannot, alone. This is yet another way usefulness will be a part of any story of epistemic normativity.

4.1.1.4. The uncontroversial core of embedded epistemic instrumentalism

So whatever your account of epistemic normativity, any reasonably broad story will have to involve usefulness. Firstly, conforming to epistemic standards is useful in almost all cases. This fact needs to be explained. Secondly, usefulness is crucial for telling a story about how humans latched onto these standards via evolution. Thirdly, usefulness can explain how the epistemic standards interact with physical, causal systems, the kind of phenomena that demonstrates just how important the epistemic standards are to us. Importantly, I am not saying that no other story of normativity can explain these phenomena. All I am saying is that usefulness will be part of that story, in addition to any other resources that non-instrumentalists can appeal to. This is the sense in which all non-instrumental accounts of epistemic normativity are instrumental-plus accounts. No matter what else is going on, even if you think epistemic normativity is a *sui generis* property that sits on top of the physical world, usefulness will be an important element in the story about how humans engage with this property. The core positive claim of embedded epistemic instrumentalism, that

conforming is broadly useful, should be uncontroversial in the sense that every broad theory of epistemic normativity will need to appeal to usefulness at some point. This uncontroversial core is itself evidence that epistemic normativity just *is* usefulness.

This uncontroversial core is why I think it is very important to work within the instrumentalist framework when investigating epistemic normativity. Even if the instrumentalist's austerity is misguided, the usefulness of conforming to the epistemic standards is so central to any story of epistemic normativity that it is worth investigating. And any story of epistemic normativity that ignores this usefulness will be extremely deficient in explaining how epistemic normativity fits in with the rest of the world. We all need to appeal to usefulness to explain epistemic normativity. The controversy is whether we need to appeal to anything else.

4.1.2. Embedded epistemic instrumentalism's model-invariance

Embedded epistemic instrumentalism is compatible with a wide variety of possibilities. This is its ecumenical nature I argued for in section 2.5. Moreover, embedded epistemic instrumentalism has an appropriately broad explanatory scope. I call a theory that has both of these features, **model-invariant**.

The thought is that the warrant to hold embedded epistemic instrumentalism and the wide explanatory power of embedded epistemic instrumentalism does not depend on other specific models and this constitutes evidence for the theory. In this sub-section, I will argue that embedded epistemic instrumentalism is model-invariant and that this is significant evidence for it. I will first contrast embedded epistemic instrumentalism with constitutivism. I will then make some general comments about why model-invariance is evidence for a theory.

4.1.2.1. Contrasting theories

Embedded epistemic instrumentalism does not need any odd metaphysics, does not need any particular psychological theory, does not need any particular epistemology to be true. It is ecumenical between many possibilities. This is in stark contrast with many other accounts of epistemic normativity.

For example, consider the type of constitutivism advocated for by Ralph Wedgwood (2002). This account holds that the essence of beliefs is that beliefs are good when they are true. Essence-of-belief accounts of epistemic normativity require beliefs to exist in a very particular way. These

accounts require **normativism** about beliefs to be true, that beliefs are a thing that somehow have a normative component baked-in. Wedgwood's essence claim is an ontologically committed claim (Wedgwood, 2002, n. 11) of the type Kit Fine offers in (1994). This is a controversial metaphysical commitment. You do not have to be an eliminativist about beliefs (Schwitzgebel, 2021, section 1.5) to deny normativism, although of course if one is an eliminativist then one cannot accept this constitutivism either. In contrast, embedded epistemic instrumentalism is compatible with belief-desire psychology, with normativism or the falsity of normativism, with Bayesian epistemology or traditional epistemology, with doxastic voluntarism or the denial of voluntarism and so on. It can be true even if our current epistemic and psychological theories are false.

Moreover, embedded epistemic instrumentalism explains a wide range of similar phenomena. Many phenomena exhibit similar properties to humans believing in the actual world: humans with different attitudes say acceptance, degrees of belief or knowing; humans making decisions about what other humans should believe; groups of humans; non-human animals; natural processes like evolution; human processes like markets; and things humans make like algorithms. These phenomena all exhibit a tendency—to a greater or lesser degree—to conform to epistemic standards and it is often good for them to conform. For example, evolution can in some sense learn and it faces a lot of the same epistemic issues that humans face. Indeed, human learning and evolutionary learning is similar enough that we can apply some of our epistemic models to evolution (Bradie & Harms, 2020). These similarities demand explanation. And these similarities are left unexplained by many accounts of epistemic normativity.

To return to our example, Wedgwood's constitutivism only addresses the normativity of beliefs and knowledge. Perhaps one might extend such a theory to degrees of belief, known as credences (for example, the essence of credences is that the more accurate they are the better). But extending the theory to other similar phenomena, like non-human animals, groups, evolution, and algorithms looks less plausible. By using the essence of beliefs as the central explanatory element, this type of constitutivism has a narrow explanatory scope that is mismatched to the range of phenomena that interact with the epistemic standards. So this type of constitutivism has this twofold problem: it only explains a very specific type of phenomenon out of a large range of related phenomena, and it demands a lot from the world, specifically that beliefs exist as the normativist describes them. This is evidence against this type of constitutivism.

4.1.2.2. Model-invariant theories vs. model-bound theories

Speaking more broadly, any account of epistemic normativity that is wedded to a specific epistemology or psychological theory is on shaky foundations. Suppose some account of normativity requires belief-desire psychology or Bayesian epistemology to be literally true (in the way Wedgewood's constitutivism requires normativism about beliefs to be literally true). This is evidence against such an account of normativity.

One way to see this is to consider the metaphysical inferences philosophers are willing to make from scientific models. Even staunch scientific realists concede that many aspects of scientific models may not correspond to reality, that scientific models are only approximately true (Psillos, 1999; also see section 4.2.2.3). Indeed, I agree with Peter Vickers (2013) that what the scientific realist is pointing at to justify their stance is just evidence for the existence of certain entities and that to some extent, the evidence can be subtly different in different cases, so episodes in the history of science need to be carefully analysed individually.⁵⁸ When making ontological inferences from scientific theories, philosophers need to be careful.

As for scientific theories, doubly so for philosophical theories. Philosophers making metaphysical claims—as opposed to exploring the consequences of a theory—should be careful to not just assume some philosophical theory is literally true and infer metaphysical claims heedlessly. This type of inference is not strongly supported by the evidence. The chance that any particular philosophical theory is literally true—as opposed to approximately true—is quite low.⁵⁹ So if the warrant for one's metaphysical or explanatory claims depends on the literal truth of such a theory, then one's claims have very little warrant. We can call such claims 'model-bound'.

Contrast this with embedded epistemic instrumentalism. It is model-invariant. It is applicable to belief-desire psychology as much as to formal epistemic models or knowledge-based models. And it will probably apply to the next epistemic model humans invent. It applies if all these theories are

⁵⁸ The idea that we need to carefully consider the evidence for any ontological inference from a scientific model is a serious problem for David Enoch's argument in: (2011, chapter 3)

⁵⁹ Does this imply that embedded epistemic instrumentalism itself is unlikely to be literally true? Perhaps, but answering this question in detail takes us beyond the scope of this thesis. Nonetheless, there are three important points to note. Firstly, if embedded epistemic instrumentalism has this property 'unlikely to be literally true', then probably all contested philosophical theories have this property, so having this property does not distinguish one theory from another rival. Secondly, there are many ways embedded epistemic instrumentalism may be approximately true and still be an extremely important philosophical theory—this is an upshot of section 4.1. Even if embedded epistemic instrumentalism's austerity is misguided, the robust usefulness of conforming to epistemic standards is still crucial for understanding epistemic normativity. Thirdly, requiring *other* theories to be literally true still lowers the chances that your theory is true, so is still evidence against your theory.

only somewhat close to the truth. It applies even if these theories are radically false. This is what it is to be ecumenical across many possibilities. And, because it is not wedded to any particular epistemology or psychology, its explanatory scope is appropriately wide. It explains why evolutionary processes would find accurate representations, as well as why human societies would enforce having accurate representations, as well as why many humans want to believe accurately. It unifies the similar normative properties found in these phenomena in a way model-bound accounts cannot do.⁶⁰ Embedded epistemic instrumentalism's model-invariance constitutes significant evidence in its favour.

4.1.3. Summary of the case for embedded epistemic instrumentalism

I have focused on the case for embedded epistemic instrumentalism. I pointed to the uncontroversial core of embedded epistemic instrumentalism, the fact that usefulness is important for any account of epistemic normativity. I also pointed out that embedded epistemic instrumentalism is model-invariant; it does not require much from the world and offers explanations for an appropriately wide range of epistemic phenomena. These two features may be related. A possible explanation for why embedded epistemic instrumentalism is model-invariant is just that its core is uncontroversial. Other theories have to accommodate the fact that conforming to the epistemic standards is robustly useful. As that is all embedded epistemic instrumentalism requires of the world, it is likely that other theories will be consistent with embedded epistemic instrumentalism. But whatever the explanation, the usefulness of conforming to epistemic standards is uncontroversial and explains a lot.

4.2. Embedded Epistemic Instrumentalism's Immunity

According to the picture from section 4.1, rejecting the instrumentalist's story amounts to motivating an addition to the story. There are many possible motivations to add to it. For example, it might be that certain semantic facts demand some type of non-cognitivist semantics. Similarly, it might be that our practices of assigning praise and blame can only be fully explained by adding a reactive-attitudes account of holding people responsible. I am open to the instrumentalist's story needing to be supplemented in these ways, in ways that do not require positing extra metaphysics, in ways that causally interact with various other bits of the world. The world is a large and complex beast. The epistemic standards touch a lot of it. So it would be surprising if there were absolutely

⁶⁰ It can also explain differences between these phenomena (for example, why evolution conforms to the epistemic standards worse than human groups) but I take it most theories can explain differences, so I did not focus on this feature.

no additions needed for the instrumentalist's story to account for the normativity of the epistemic standards, at least around the edges of the phenomenon.

However, in this section, I want to reject one type of motivation for adding to the instrumentalist's story, **intuition-based criticisms**. I identify two types of intuition-based criticisms: intuitive counterexamples and intuitive desiderata. Both of these things function to push philosophers to add to the instrumentalist's story. My claim is that the instrumentalist has a master response to all intuitive counterexamples and desiderata. To argue for this claim, I will first clarify what I mean by intuitive counterexamples and intuitive desiderata. I will then offer the master response. Finally, I will offer a specific response generated by the master response. Once the master response is on the table, there is not much motivation left, if any, to add to the core instrumentalist story. This is the sense in which instrumentalism explains all the relevant phenomena associated with epistemic normativity. It undermines the typical motivation to add additional metaphysics to instrumentalism to account for epistemic normativity.

4.2.1. Intuitive counterexamples and intuitive desiderata

Lots of philosophy involves appealing to intuitions as evidence (Climenhaga, 2018; Ashton & Mizrahi, 2018). It is controversial what this intuition-based method actually is (cf. Bealer, 1996; Jackson, 1998; Williamson, 2007). But being purely descriptive, philosophers often describe a case, report a judgement in reaction to that case, and without identifying any additional evidence, treat the content of the judgement as true. Similarly, philosophers describe a proposition, judge it is intuitive, and without identifying any additional evidence, claim that it is desirable if philosophical theories are consistent with the truth of the proposition. What it is to 'have a judgement' I am leaving vague. The judgement might be a judgement that we know the proposition. Or it might be a feeling that the proposition is true. Or it might be a basic 'plausible' feeling, something *sui generis*. But the important point is that philosophers are not identifying other evidence for the proposition.

For an example of an intuitive desideratum, consider what I am calling **the universalist judgement**. Universalism is the position that the epistemic standards give rise to epistemic normativity, epistemic reasons, always and everywhere. The universalist judgement is the judgement that universalism is true. Côté-Bouchard asserts that, arguably, the motivations for universal epistemic instrumentalism outweigh the motivations for any other type of epistemic

instrumentalism (2015, p. 343).⁶¹ I read Côté-Bouchard as suggesting that universalism is so intuitive that if we can vindicate the universalist judgement, we should. Hilary Kornblith is also keen to square epistemic instrumentalism with universalism in (1993), more or less anticipating the next three decades of instrumentalist literature. So at least some philosophers hold that universalism is intuitive enough that our theories of epistemic normativity should be consistent with it. For this thesis, I am assuming that the proposition: ‘epistemic standards produce epistemic reasons always and everywhere’ feels true to the majority of people who can entertain it. This is made into an intuitive desideratum if, without identifying any additional evidence, philosophers claim that theories of epistemic normativity should be consistent with the truth of that proposition.

For an example of an intuitive counterexample recall the **too-few-reasons objection**. As a reminder, the strategy of the too-few-reasons objection is to find a case where there are no instrumental reasons to conform to the epistemic standards, yet where philosophers judge that there are still epistemic reasons in the case; therefore, the objector concludes, epistemic reasons cannot be instrumental reasons. Again, a case where, supposedly, there are no instrumental reasons but where, supposedly, we judge that there are epistemic reasons, is **a too-few-reasons case**. Recall the representative too-few-reasons case from Côté-Bouchard that I have been using throughout this thesis:

Spoiler Alert Max missed the finale of his favourite television series, which aired last night and revealed whether character X ends up dying or surviving. Given his passion for the show, Max really wants to avoid learning the answer to that question before he watches the rerun tonight after work. But to his dismay, one of his colleagues – who is convinced that Max did watch the show last night – comes up to him in the morning shouting: ‘I knew it! I told you character X would die!’ (Côté-Bouchard, 2015, p. 340).

Again, the thought is that Max has no instrumental reasons to conform to the epistemic standards (he wants to not believe the spoiler) yet we judge that the epistemic standards are still compelling. Epistemically speaking, Max should believe his spoiler-uttering colleague. So, it seems, usefulness cannot be the whole story about epistemic normativity.

⁶¹ I note Côté-Bouchard (2015) uses different terms to me. He classifies different types of embedded epistemic instrumentalism with how they respond to the too-few-reasons objection (I discuss this below). He calls universal embedded epistemic instrumentalism ‘the optimistic response’. Near-universal embedded epistemic instrumentalism is a version of the ‘pseudo-reasons response’. Why did I change terms? To a previous paper, an anonymous referee objected to Côté-Bouchard’s terminology as not as informative as possible, and I was not going to dig my heels in, risking a rejection, over a name. Now, I wish to be consistent with my own terminology.

As I have argued above, a deficiency with the too-few-reasons objection is that it is not so hard to find relevant instrumental reasons in Spoiler Alert, or any case like it. It is straightforwardly useful for Max to practice conforming. If Max has the typical range of human goals, he will want to be smart, clever and reasonable. Not believing reliable testimony is not very smart, clever or reasonable. If Max has the standard range of social goals, he will not want to offend his colleague by refusing to believe their testimony or be viewed as *weird* for not believing reliable testimony. And if Max wants to promote the goals of his community, it will be useful to predictably conform to the epistemic standards so in the future the people around Max will find him easier to work with (more predictable, more trustworthy). So it is overwhelmingly useful to conform to the epistemic standards in Spoiler Alert, and standard too-few-reasons cases in general.

But, again following our idealization strategy, we can modify standard too-few-reasons cases to exclude all possible instrumental reasons:

An Extreme Too-Few-Reasons Case Suppose Shelly lives in a world with an all-powerful god. This god will intervene to make it the case that conforming to or flouting the epistemic standards will not impact any of Shelly's interests or goals. Shelly believes this with certainty. Shelly has no other goals in life other than to survive. Shelly is thirsty and she believes that there is a glass of water in front of her. Then she receives a lot of evidence that the apparent glass of water is actually poison. (Of course, if Shelly drinks the poison, the all-powerful god will intervene, changing the poison to water.)

Does Shelly have epistemic reasons to believe there is poison in front of her? I do not judge so. I agree she has evidence that the glass of water is in fact full of poison. I agree that maintaining a belief 'the glass is full of water' is against the epistemic standards. I just hold that in a world with an all-powerful caretaker god, believing against the epistemic standards has no normative sting. Of course, I may just be infected with theory; my judgements may be warped from defending instrumentalism for so many years.

Suppose you judge that Shelly does have epistemic reasons.⁶² Let us call a judgement that an agent has epistemic reasons in a case a **too-few-reasons judgement**. In this case, Shelly does not have any goal or end that is promoted or hindered by not conforming to the epistemic standards. A god will intervene to ensure that the poison does not hinder Shelly's survival. She cannot have

⁶² I am granting my interlocutors that philosophers do judge this way: see n. 15.

subjective instrumental reasons to believe because she believes with certainty that a God will intervene and she is right. Extreme too-few-reasons cases are successful because they combine two elements: 1. the too-few-reasons judgement that an agent has epistemic reasons to believe and 2. a situation where there are no relevant identifiable instrumental reasons. This is an intuitive counterexample.

My claim is that such a style of argument, both appealing to an intuitive desideratum or an intuitive counterexample, provides no significant evidence that an addition to the instrumentalist's story is needed. The master response I offer below, a schema for generating specific responses, can account for the intuitiveness of any proposition about epistemic normativity. For the instrumentalist, the availability of this master response undermines the evidential value of intuitive criticisms against embedded epistemic instrumentalism.

4.2.2. The master response

In this sub-section, I aim to convince you that this master response is true:

The Master Response Intuition-based criticisms lack any significant evidential force against embedded epistemic instrumentalism. Intuition-based criticisms assume a particular explanation of judgements. Instrumentalists can offer an at-least-as-likely alternative explanation of those judgements. The availability of such an explanation implies that intuition-based criticisms have no significant evidential value against embedded epistemic instrumentalism. So intuition-based criticisms fail to identify significant evidence against embedded epistemic instrumentalism.⁶³

To defend this master response, I first characterise the explanation that intuition-based criticisms must assume (4.2.2.1). I contrast the assumed explanation with an example of an alternative explanation and argue this alternative implies the relevant criticism has no significant evidential weight against embedded epistemic instrumentalism (4.2.2.2). Finally, I offer a recipe for generating alternative explanations in all cases (4.2.2.3). The upshot is that the instrumentalist will

⁶³ Why 'significant evidence' as opposed to 'any evidence'? Evidence can be very promiscuous and cheap; finding non-black shoes in your wardrobe might count as evidence that all ravens are black (Crupi, 2021, section 1.2). My argument does not rule out the possibility that intuition-based criticisms produce vanishingly small amounts of evidence against embedded epistemic instrumentalism. But I do argue that no intuition-based criticism could provide enough evidence to rationally change someone's mind about endorsing embedded epistemic instrumentalism. This is what I mean by 'significant evidence', evidence that would actually move a debate forward.

always be able to generate an alternative explanation, so intuition-based criticisms have no significant force against embedded epistemic instrumentalism.

4.2.2.1. The assumed explanation

Suppose there is a judgement that X, where X is some proposition that is putatively evidence against embedded epistemic instrumentalism. For example, the X-judgement in question might be the judgement ‘evidence always produces epistemic reasons’ or ‘epistemic reasons are normative’ and ‘usefulness is descriptive’. One might offer arguments against embedded epistemic instrumentalism based on these judgements and many judgements like them.

Intuition-based criticisms assume two facts best explain why we make the X-judgement:

1. X is true.
2. Humans have the ability to track that truth.

Both facts need to be true for an intuitive counterexample or intuitive desideratum to count as significant evidence against a theory. If X is false, there is no need for a theory to be consistent with the truth of X. If we cannot track X, the fact that we have the X-judgement does not provide much evidence for X. This is what ‘saving the appearances’ amounts to: assuming that we latched onto the X-judgement because X is true and we have an ability to track that truth.

Consider An Extreme Too-Few-Reasons Case. Assume that the vast majority of philosophers have the too-few-reasons judgement, that there are epistemic reasons for Shelly to believe there is poison in front of her. Therefore, the objection supposes, instrumental reasons fail to properly account for epistemic reasons. When offered as an objection, we assume or infer that the too-few-reasons judgement is veridical, that there are epistemic reasons in the case, and we have the ability to track that truth. This is why interlocutors are supposed to take reports ‘there are epistemic reasons for Shelly to believe’ seriously.

4.2.2.2. Alternative explanations

There are many possible explanations about why we make X-judgements. And in many of the possible explanations, X is false. For example, we might make X-judgements because X is close to the truth in important ways and being wrong has little cost. This is a circumstance where we shouldn’t be surprised that creatures evolved to make such a judgement. Indeed, I used this type

of explanation to reply directly to too-few-reasons cases in section 2.3. Conforming to the epistemic standards is highly useful in almost all cases. And our world is such that the expected cost of misidentifying when the epistemic standards have normative force is very low. So I propose that we developed a heuristic. We make too-few-reasons judgements because they were and remain a useful heuristic. We judge as if the epistemic standards give rise to epistemic reasons always and everywhere because those judgements are rarely false, and even when they are, there is little cost for being wrong.

Now compare the two possible explanations:

Intuitive explanation: We make a too-few-reasons judgement because epistemic standards have normative force always and everywhere and this is a fact we track.

Heuristic explanation: We make a too-few-reasons judgement because too-few-reasons judgements are almost always veridical and errors are almost costless, so we evolved a heuristic.

If the heuristic explanation is correct, then the fact many humans make too-few-reasons judgements is not evidence against embedded epistemic instrumentalism. Indeed, it confirms embedded epistemic instrumentalism because developing useful heuristics is exactly what one would expect to happen if the theory were true. So the objector appealing to the fact that humans make the X-judgement, without supporting the intuitive explanation, does little to rationally compel the embedded epistemic instrumentalists to her position.

I believe the heuristic explanation is a much better explanation. But it is important to note that the instrumentalist does not need it to be better. Intuitive counterexamples and intuitive desiderata are supposed to function as criticisms of embedded epistemic instrumentalism. They are supposed to move the instrumentalist to be an instrumentalist-plus. But if this alternative explanation is at least as likely true as the intuitive explanation, then the X-judgements are not significant evidence against the theory. Moreover, the instrumentalist does not even need a particular alternative explanation to be equally likely. They just need it to be equally likely that *some* alternative explanation that does not imply the truth of X is correct. This is the dialectical fact that the master response takes advantage of.

Offering the instrumentalist intuition-based criticism is similar to offering a ghost-skeptic testimony of people experiencing ghosts. If the best explanation for why people testify to having ghost experiences was the existence of ghosts, then perhaps this testimony might move the ghost-skeptic. But the ghost-skeptic holds there are many other explanations for why people report ghost experiences. So the testimony alone does little to move the ghost-skeptic.

4.2.2.3. A recipe for generating alternative explanations

The goal of the master response is to establish that some alternative explanation, which does not imply the truth of X, is at least as likely as the intuitive explanation. If the instrumentalist can establish this, then the instrumentalist can costlessly reject the intuitive-based criticism.

The relevant possible explanations depend on the facts about the X-judgement in question. Sometimes the circumstances are such that X being true looks like a fortunate coincidence (just like in evolutionary debunking arguments) so it is at least equally likely that X is false. Sometimes the circumstances are such that more is needed to justify an available alternative explanation. We can determine the relevant response by answering two questions about the target X-judgement.

Question 1: is the judgement widespread?

By ‘widespread’ I mean a large proportion of the people who entertain the relevant proposition or case make the X-judgement. If the judgement is not widespread, then it looks quite lucky that the few people that have the judgement happen to be correct. So it is at least equally likely that the few people having the judgement are just wrong as it is likely that the intuitive explanation is correct. If the appeal of the X-proposition is idiosyncratic, it should not motivate adding to the instrumentalist’s story. In general, we should not modify a theory because only a handful of people find a certain proposition intuitive. It is uncontroversial that appealing to intuitions is very noisy; being so sensitive to idiosyncratic judgements would take us wildly off-course. So it is not a cost for the instrumentalist to reject the intuitive explanation about an idiosyncratic X-judgement.

If the judgement is widespread, we must continue on. Recall that we are in a situation where we cannot identify any explicit arguments that compel us to endorse the X-proposition. We are relying on some appearance or feeling or judgement. If we were not compelled to make the judgement as a result of some explicit argument, the only plausible cause of our widespread judgement is some

evolutionary process, just as philosophers hold in the evolutionary debunking literature. The next question we have to ask ourselves is how did the judgement make it through the process?

Question 2: was the judgement useful?

If the judgement was not useful, then the evolutionary process could not track the truth of the content of that judgement. The only way evolution learns about the world, the only way it can teach us, is via usefulness. If the X-judgement is not useful to have, then at best evolution can randomly produce it, and the content happens to be true; the judgement could be a side consequence (Gould & Lewontin, 1979) or part of a random walk, or be produced in some other epistemically disconnected way. In this case, the truth of X again looks like a fortunate coincidence. So again that X is false is at least as likely as the intuitive explanation being correct.

If the judgement was useful, then the situation is more complicated. Usefulness is itself an epistemic property. A judgement might be useful because it is true. As I claim above, accurate representations of the world are very useful. This is the force of the No-Miracles Argument (Putnam, 1975, p. 73). Scientific theories allow us to predict and intervene on the world in a variety of impressive ways. The only explanation that does not make this usefulness of scientific theories a miracle, is *some* connection to the truth, or so holds the No-Miracles Argument.

But usefulness does not imply strict truth. As I hinted at above, scientific realists, in the face of what is known as the pessimistic induction, tend to explain the usefulness of scientific theories with approximate truth rather than strict truth (Lange, 2002; Laudan, 1981; Lewis, 2001; Lyons, 2017; Niiniluoto, 2017; Psillos, 1999; Saatsi, 2005; Vickers, 2013; Wray, 2015). This is the gist of the pessimistic induction: when we look at the history of science, we see many abandoned but useful theories. Indeed, some useful theories look wildly wrong, such as the Ptolemaic model of the cosmos. So strict truth cannot be the only explanation of usefulness because there are many theories that were useful but not true.⁶⁴ The standard realist response is to claim that past theories were approximately true. What ‘approximately true’ means is controversial and vague. But the important point for our purposes is that even the impressive usefulness of scientific theories does not imply the strict truth of those theories. This brings us to one candidate alternative explanation to the intuitive explanation. Being ‘partially accurate’ is a relevant alternative explanation for the

⁶⁴ A note on the dialectic: if the realist is wrong, if we can appeal to anything other than truth to explain the usefulness of a theory, then the instrumentalist can also appeal to that other thing. Any explanation that does not imply the truth of X, implies that the X-judgement has little evidential value against embedded epistemic instrumentalism.

usefulness of making the X-judgement. And this is a circumstance where, apart from the X-judgement itself, we have no identifiable evidence that X is true. So the explanation that the X-judgement is somewhat accurate is at least as likely as the intuitive explanation.

In addition, there are non-accuracy based alternative explanations for the usefulness of the judgement, too. A judgement might be useful to make because it makes humans behave in a certain way and that behaviour is itself useful. I take it this is the thrust of many evolutionary explanations of religious belief (Schloss & Murray, 2009). It might be that certain X-judgements were useful to have, not because they are strictly true or even approximately true but because they encourage useful behaviour. Therefore, in a circumstance where we can identify a potentially useful behaviour promoted by having the X-judgment, it is at least as equally likely that the judgement was useful and strictly false as it is likely that the intuitive explanation is correct. So, again, in this circumstance it is no cost for the instrumentalist to reject the intuitive explanation.

There may be more strategies for identifying an at-least-as-likely alternative explanation. But I do not think the instrumentalist should worry about a poverty of relevant alternative explanations. Just-so stories about why certain judgements evolved are not hard to tell. And all the instrumentalist needs to establish is that *some* alternative explanation is at least as likely as the intuitive explanation. So I bet that there will be many possible alternative explanations that utilise the approximate truth or the useful behaviour strategy, more than enough to make rejecting the intuitive explanation costless.

In summary, the master response:

When offering an intuitive counterexample or desideratum, the instrumentalist can point to an alternative explanation to the X-judgement being true and humans having the ability to track that truth.

- 1) If the X-judgement is not widespread, it is at least as likely that the handful of people who make it are wrong.
- 2) If the widespread X-judgement was not useful, then again it is at least as likely that the judgement is wrong.

- 3) If the widespread X-judgement was useful, then it is at least as likely that X is strictly false but useful by being close to the truth or by causing a useful behaviour.⁶⁵

So in no case is it a cost for the instrumentalist to deny that X is true and that we can track X. More evidence needs to be identified to rationally move the instrumentalist from their austere theory.

4.2.3. An example response

Let's see the master response in action. Recall one such instance, in response to An Extreme Too-Few-Reasons Case. The instrumentalist can appeal to the heuristic explanation, that a judgement that Shelly has epistemic reasons (even if she has no instrumental reasons) is a useful judgement to make even if it is strictly false in this particular case. You can imagine a similar response to the universalist judgement. The explanation about why it feels intuitive that the epistemic standards have normative force always and everywhere is that the judgement is a useful heuristic. This response is the key difference in content between universal and non-universal instrumentalists, like embedded epistemic instrumentalists. But, if I am right, the master response handles a much larger range of intuitive criticisms.

Let's introduce a new intuitive desideratum for a theory of epistemic normativity. Suppose you judge that epistemic reasons are decisive when considering what one ought to believe. So you offer this argument:

P1) Epistemic reasons settle what one ought to believe.

P2) Facts about usefulness do not settle what one ought to believe.

Therefore,

C1) Epistemic reasons cannot reduce to facts about usefulness.

⁶⁵ I note that if, like Jackson (1998), you hold that judgements in response to cases are judgements about the meanings of words then the master response has an extra step. I am open to the idea that our judgements are evidence for the meanings of words. But then the question becomes: are the meanings of X-words evidence for the truth of X? In response, the instrumentalist can adapt the master response by substituting the 'meaning of X-words' for 'X-judgements'. For example, suppose you judge that the meaning of 'epistemic normativity' precludes a reduction to instrumental reasons. Is that significant evidence against the instrumentalist? No. Why? Because there are at-least-as-likely alternative explanations about why we would develop a term with that meaning, that do not require the meaning of the term to perfectly track metaphysical reality. Adopting this response implies that embedded epistemic instrumentalism is not an *analytic* reduction but, as the positive case for embedded epistemic instrumentalism is not based on semantic or conceptual relations, this is not a problem for the theory.

Suppose you claim that P1 is intuitive; you judge P1 true.⁶⁶ So, you further assert, it is a desideratum for any account of epistemic normativity that it is consistent with the truth of P1. We agree P2 is true. So you reject embedded epistemic instrumentalism because it is not consistent with P1. This is something like an epistemic open question argument (Greco, 2015). I take it this type of judgement is a major source of dissatisfaction with the instrumentalist's proposed reduction of epistemic reasons. The judgement might not be about whether epistemic reason settle what one ought to believe. It might be about the ineffability of normativity or the necessity of normativity or some other property normativity intuitively has. "Genuine normativity," someone pressing this line of argument will say, "is just something completely different to the physical circumstances that constitute facts about usefulness! So we should reject embedded epistemic instrumentalism."

I will now offer a response on behalf of the embedded epistemic instrumentalist. If the P1-judgement is idiosyncratic or making this judgement was never useful, then it looks extremely lucky if the judgement is correct. So it is at least equally likely that the judgement is wrong, that epistemic reasons do not settle what one ought to believe. In this circumstance, it is no cost to the instrumentalist to reject the P1-judgement.

But suppose the P1-judgement is widespread and we find out that having the P1-judgement was very useful for humans at a certain point in time. Then the instrumentalist will point to the usefulness, explaining that it helped humans at a certain point in time to make this judgement and that is why it was selected for. In this particular case, I would use a useful behaviour alternative explanation. A judgement that epistemic reasons settle what you ought to believe might help simplify epistemic evaluations. Sometimes it helps to make assessments divorced from a broader context. For example, when playing basketball, all-things-considered evaluations are awkward to calculate because they are extremely complex. Think about all the ways various actions taken in a basketball game could impact all the other things in the world. It is much easier to just evaluate

⁶⁶ Spencer Case (2020) believes P1 is true:

The essence of epistemic standards is that they are genuinely authoritative over how we should think. They are not merely epistemically authoritative over what we ought to think, which would be trivial. Rather, they tell us how we Just Plain Ought to think—at least in ordinary circumstances. That is what distinguishes epistemic standards from arbitrary standards of thought, and what makes them a worthy subject of philosophical investigation (Case, 2020, p. 97) [Footnotes removed].

The only things resembling explicit arguments in favour of the assertion that epistemic standards genuinely authoritative that Case refers to, is an assertion that the epistemic standards are like moral standards, and perhaps the claimed explanatory power from the passage above. Nevertheless, as I read the passage, the intuitive force of the assertion is supposed to do some work. I return to Case's position in chapter 6, specifically sections, 6.2.2.4, 6.3.3 and 6.5.

the actions from a basketball perspective, or even from a simpler perspective, say getting the very next basket. Returning to the P1-judgement, it might be that epistemic evaluations are so common and so useful, that nature has endowed us with a feeling that epistemic reasons settle, so we just make the epistemic evaluation rather than the all-things-considered evaluation. This could be nature's way of getting us to focus on the epistemic equivalent of the very next basket.

Importantly, I am not claiming this explanation is correct! I lack a serious amount of evidence for that. Rather, I am just claiming this alternative explanation is at least as likely as the intuitive explanation that our judgements about the essence of epistemic normativity are true and we have the ability to track this truth about essences. Indeed, I am inclined to believe it is much more likely that we would evolve a P1-judgement because it helped us simplify our evaluations of the world than we evolved the ability to accurately track the essential metaphysics of epistemic normativity via a judgement—but that is a stronger claim than I need.

From this example, we can see that applying the master response is not mechanical. For any particular judgement, there is a story that needs to be told about how that judgement could be useful but not strictly true. For the universal judgement, the story involved approximate truth. For the P1-judgement, the story was about a useful behaviour, simplifying a decision procedure. The master response, while pointing you in the right direction, does not give you all the details. But, as I said above, telling a plausible just-so story about why a judgement evolved is not hard. Indeed just-so stories are so unconstrained it is difficult to settle on any particular one. This lack of constraints suggests that there will be an available alternative explanation no matter the judgement, be it about the wrong type of reasons, the way epistemic reasons trade-off against non-epistemic reasons or whatever intuitive criticism future philosophers will generate (cf. Berker, 2013a; Berker, 2013b, 2018; Way, 2012).

4.3. The Not-All-Judgments Objection

In this section, I respond to an objection that my master response is too general. To see why one might worry about the generality of my master response, note that the first wave of experimental philosophy, the so-called 'negative program', was extremely critical of the use of intuitions or judgements in philosophy (Knobe & Nichols, 2017, section 2.1; Machery, 2017). A response to this negative program was to point out that using judgements is necessary and extremely widespread when doing philosophy (Williamson, 2004). So you cannot undermine the use of judgements wholesale without undermining your own arguments. I anticipate some readers will

have a similar response here. I use judgements in this thesis. For example, I use judgements about the likelihood of explanations being correct. So you might worry that my argument applies to all judgments, undermining the very judgements I use to establish the conclusion.

I concede that I do not know the exact scope of the master response. But there are five factors that separate judgements at large from the target judgements of my argument. Firstly, in this thesis, I am not denying that judgements can be evidence for the content of the judgement. Maybe if we are conducting first philosophy, maybe if we are **superbabies**,⁶⁷ all we can do is be guided by our judgements alone. But we know a lot about the world, a lot about how we got here and about how our judgements developed. There is a lot more evidence available to us than our brute judgements about epistemic normativity.

Secondly, by stipulation the target judgements lack independent confirmation. Consider too-few-reasons judgements. Apart from the judgement itself, there is no way to verify that the epistemic standards have genuine normative force in any particular case. This is not like judgements based on what you can see, or judgements about whether an argument will be convincing, where we get a lot of independent confirmation. Even judgements about usefulness have independent lines of confirmation. A useful heuristic for playing poker will eventually help you win; mountains of poker chips in front of you provides an independent way to confirm usefulness. This is unlike the target judgements. They have to shoulder all the epistemic burden themselves. And they can only do that if we trust them completely.

Thirdly, particularly when it comes to the universalist judgement or the too-few-reasons judgement, the judgements must be strictly true. It is not enough for the judgements to merely be tracking some truths closely. It must be the case that epistemic standards have normative force always and everywhere, in every too-few-reasons case, for the universalist judgement or the too-few-reasons judgement to count against embedded epistemic instrumentalism. This is highly demanding for judgements that were produced in a tumultuous evolutionary process. Even when we try our hardest to describe our world, being as careful as possible with the evidence, attempting to ignore anything not epistemically relevant, our best descriptions (scientific theories) are at most approximately true rather than strictly true. So it is a large burden for our judgements to shoulder,

⁶⁷ The term ‘superbabies’ was coined by David Lewis. It refers to a rational agent who has absolutely no evidence (Hájek, 2012, p. 418). Hájek first heard the term during an undergraduate lecture by David Lewis on time travel (Hájek, 2012, p. 418, n. 6). I believe I first heard the term during an undergraduate lecture on David Lewis (given either by Hájek or Daniel Nolan).

which in the best epistemic case, were forged to help us survive rather than only concerned with getting at the truth.

Fourthly, ‘epistemic normativity’ is not a common English term (see section 1.3.4). At the time of writing,⁶⁸ searching the *Corpus of Contemporary American English* for the term returned zero results. I think we partially latch onto the phenomenon with certain uses of other terms, like ‘reasons’, ‘rational’, ‘reasonable’, ‘evidence’ and most importantly ‘justification’. But all these terms are noisy and, I suspect, are sensitive to other types of normativity. So it is difficult to even identify what the common-sense metaphysical judgements about epistemic normativity of native English speakers actually are, if there are any.

Fifthly, the judgements are supposed to function as criticisms. This is the major dialectical difference between the master response and evolutionary debunking arguments. Evolutionary debunking arguments are used as a sword, as a criticism that points out that a fortunate coincidence needs to take place for the content of certain judgements to be true, so philosophers should not accept the veridicality of those judgements. But in this thesis, similar evolutionary facts are being used as a shield. Interlocutors use the judgements as evidence that embedded epistemic instrumentalism is false. But instrumentalists can point out that other explanations are at least equally likely. The instrumentalist is pointing out that from the instrumentalist’s perspective, it is costless to reject the strict truth of certain judgements so the judgements do not function as criticisms. This is especially important because, if my argument from section 4.1 is correct, everyone should already accept that usefulness plays an important part in the story of epistemic normativity. These judgements are supposed to rationally motivate adding to the instrumentalist’s story, supposed to identify some phenomena that the instrumentalist lacks the tools to explain, but they ultimately fail to do that.

Combined, these five factors make criticism via intuitive counterexamples or intuitive desiderata impotent against embedded epistemic instrumentalism. Not all judgements are subject to the same issues. So my master response will not generalise to all judgements.

⁶⁸ The search was conducted on 29/06/2022. ‘Epistemic reasons’ only returned one result, a blog post about theology. Interestingly, the *Corpus of Contemporary American English* includes some academic sources and some philosophy specific sources, such as entries from the Stanford Encyclopedia of Philosophy. So the term ‘epistemic normativity’ may even be relatively rare in philosophy and academia at large, as well as in English.

4.4. The Third-Factor Response Objection

In this section, I explain how third-factor responses to evolutionary debunking arguments impact the dialectic. A third-factor response is one in which non-debunkers claim that our capacity to identify the relevant reasons is a side-consequence of some other capacity. Jessica Isserow characterises such a response (Isserow, 2019, pp. 1033-1036). We evolved a capacity for recognizing consequences of actions, *a capacity to judge certain things as more or less desirable* (make note of this for later), and a capacity to choose between alternative courses of action. I take it as uncontroversial that these capacities could evolve. A third-factor response then claims that our capacity for normative judgement is a by-product of evolving other capacities, such as those I have listed. So it was not fluky that we evolved a capacity to make normative judgements; rather they are just a product of having other capacities, capacities that unproblematically integrate with our knowledge of evolution. Isserow does not find anything especially problematic about asserting this type of evolutionary history (Isserow, 2019, p. 1036).

When I have presented this material, philosophers have responded that my master response does not account for third-factor responses to evolutionary debunking—that it is possible that we evolved some capacity to identify epistemic normativity via some other capacity that evolved, so it is possible that the relevant judgements are veridical. I have two responses.

Firstly, I still find asserting this type of third-factor evolutionary history fluky, depending on your particular metaphysics of normativity. Suppose you hold that normativity is metaphysically-basic, just a brute fact about our world. Now consider the third-factor history I just put forward. It looks like a fortunate coincidence that our *capacity to judge certain things as more or less desirable* in any way aligns with the metaphysically-basic normative properties. If the *capacity to judge certain things as more or less desirable* is some instrumental capacity, an ability to judge what we want, what others want, what might help achieve a goal, etc. then it is not lucky that such a capacity would evolve but it is lucky that the orderings of desirability match the objectively correct orderings determine by basic metaphysical properties. On the other hand, if the *capacity to judge certain things as more or less desirable* is supposed to be something non-instrumental, an ability to determine the objective desirability of something say, then this is the very capacity we need some non-fluky evolutionary history to explain. In general, claiming both that judgements evolved (not because having veridical judgements was useful but) as a side consequence of evolving some other useful capacity and, at the same time, claiming that nonetheless the judgements are still veridical, is a suspiciously lucky history. And it is not a history that the instrumentalist needs to offer more evidence to reject.

Secondly, the dialectical situation is different in the master response compared with evolutionary debunking arguments. With the latter, the debunkers are the people offering a criticism and the non-debunkers are the people defending. So the fact that non-debunkers can potentially offer a particular evolutionary history that does not involve a fluke is enough to dissolve the debunking argument. Contrast that dialectical context with the present one. The objector offers an intuitive criticism of embedded epistemic instrumentalism with an associated judgement. The instrumentalist responds with the claim that there are plenty of explanations about why we would make the relevant judgement, which do not require the content of the judgement to be strictly true. So, the instrumentalist maintains, the judgement has no dialectical force against the instrumentalist. The response that it is possible the judgement is veridical does not change the dialectical situation because it is also possible the judgement is not veridical. What the objector needs is explicit evidence that the judgement is veridical but, by stipulation, all intuitive criticism lacks that extra evidence.

The instrumentalist is not leaning on any particular evolutionary history to argue for their position. I do note that embedded epistemic instrumentalism is not subject to the debunking worry. And that, even with our limited evidence, the usefulness of conforming to the epistemic standards will likely be an important part of our cognitive evolutionary history. I take it these are reasons to accept embedded epistemic instrumentalism. And I note the wide variety of plausible possibilities where a judgement is not veridical, to reject intuitive criticism. But none of this requires a particular evolutionary history. Indeed, considering that evolutionary processes are tumultuous, asserting that such a process produced a veridical judgement without other evidence is an extremely undermotivated claim. So, the instrumentalist should not be moved by the possibility of a third-factor response until more explicit evidence is offered in support. That is all the master response demands of interlocutors.

5. Two (Relevant) Interpretations of Embedded Epistemic Instrumentalism

In the previous chapters, I have characterised and defended embedded epistemic instrumentalism. However, there is an important detail I have left out. I have not stated what type of theory embedded epistemic instrumentalism is. In this chapter and the next chapter, I want to defend the various ways embedded epistemic instrumentalism might be interpreted. My position is that it does not really matter how one interprets the theory. But to argue for this, I will characterise and defend the different interpretations.

The plan: in section 5.1, I summarise Parfit's taxonomy of meta-normative theories to give a sense of the options. In section 5.2, I add some additional distinctions to the taxonomy. In section 5.3, I introduce another way to classify the same theories. In section 5.4, I note another distinction, a distinction between types of error theories. In section 5.5, I settle on two relevant interpretations, non-analytic natural realism or error theory.

5.1. A Taxonomy of Normative Theories

I will use Parfit's taxonomy of theories about the metaphysics of normativity to give a sense of the options (Parfit, 2011, p. 263). I take this taxonomy to be fairly representative of how philosophers carve up the options. However, in some cases I will use different labels, as they better reflect the specific literature I am engaging with in this thesis.⁶⁹ Interestingly, embedded epistemic

⁶⁹ Here is a translation between Parfit's terms and my own. I use the label 'cognitivism' where Parfit uses the label 'semi-cognitivism'. I use the label 'error theory' where Parfit uses the label 'nihilism'. I use the label 'realism' where Parfit uses the label 'cognitivism'. I use the label 'non-natural realism' where Parfit uses the label 'non-natural cognitivism'. I use the label 'natural realism' where Parfit has no label for this. I use the label 'non-analytic natural realism' where Parfit uses the label 'non-analytic naturalism'. I use the label 'analytic natural realism' where Parfit uses the label 'analytic naturalism'.

Why the change of terms? It is because of a complicated balance of reasons. The theories that Parfit refers to with the term 'semi-cognitivism' are usually referred to with the label 'cognitivism' (cf. van Roojen, 2018). I suspect Parfit uses the term 'semi-cognitivism' because he reserves the term 'cognitivism' for another group of theories. However, every member of that group is a type of realism, so I use the term 'realism' for that group. I suspect Parfit did not want to use the term 'realism' because its usage is contested. Sometimes 'realism' in metaethics is reserved for theories that are non-reductive and mind-independent. I hold that this is a vice. Chairs are real even if you hold that they reduce to particles arranged chair-wise. Universities are real even if you hold they are partly constituted by the attitudes of people. But this view of realism is controversial. I use 'error theory' rather than 'nihilism' because 'error theory' is better for the discussion in chapter 6, even though it is slightly different to the way I used the term 'error theory' in chapter 2. But the two uses are closely related. I refer to my heuristic explanation of too-few-reasons judgments as 'an error theory' because it holds that some too-few-reasons judgments are in error, strictly speaking. Error theory is the position that all of our normative judgments are in error, strictly speaking.

instrumentalism is not neatly classified by this taxonomy. Nevertheless, articulating why it is difficult to classify will help clarify how it is related to other theories.

Parfit classifies theories with a series of four binary choices. The first choice-point: are normative claims intended or believed to state normative truths? If a theory holds that normative claims are not intended or believed to state normative truths, if normative claims are not truth-apt, then the theory is a type of **non-cognitivism**.⁷⁰ This type of theory holds that normative claims are similar to utterances like: “Boo Collingwood football club.” This sentence is not true or false. Non-cognitivist normative theories hold that normative claims are similarly not true or false.

According to embedded epistemic instrumentalism, whether epistemic normativity supports any particular belief is a fact, capable of being true (or a purported fact, capable of being false). For example, in cases involving Zen Boltzmann brains (see section 2.3) it holds that in fact epistemic normativity does not compel any beliefs. This *suggests* that it is not a non-cognitivist theory. Nonetheless, strictly speaking, embedded epistemic instrumentalism is ecumenical on the semantics of normative language (including the intentions or beliefs of people who use normative language). So one can imagine a non-cognitivist semantics combined with the metaphysics of embedded epistemic instrumentalism. Parfit’s taxonomy does not anticipate such a combination of metaphysics and semantics. I suspect this is because one of the key motivations to offer a non-cognitivist semantics is a pessimism about a successful ontological reduction of epistemic normativity, exactly the type of thing that embedded epistemic instrumentalism purports to do. I would interpret this combination as *something like* an error theory, where there is a slight mismatch between the semantics and the metaphysics (see below, section 5.4 and chapter 6). With a non-cognitivist semantics, embedded epistemic instrumentalism suggests that our epistemic practices, including our language, are tracking facts about usefulness—just not via truth-apt claims. Importantly, even with such a semantics, embedded epistemic instrumentalism still posits usefulness as the central explanans of epistemic normativity. Because it posits usefulness as so central, and usefulness is a fact, I do not consider embedded epistemic instrumentalism as a type of non-cognitivism. Even though a non-cognitivist interpretation is a theoretical possibility, embedded epistemic instrumentalism is too different *in spirit* for it to be a relevant interpretation.

⁷⁰ Parfit uses the language of ‘intentions’ and ‘beliefs’ (Parfit, 2011, p. 263). This implies that whether a sentence is truth-apt is settled by facts about intentions or beliefs. But you might hold that the truth-aptness of a sentence is not settled by such facts. If that is the case, then you ought to substitute the relevant facts that do settle whether a sentence is truth-apt.

If a theory holds that normative claims are intended or believed to state normative truths, then the theory is a type of **cognitivism**. Cognitivist theories are further divided. The second choice-point: are there any normative truths? If a theory holds that there are no normative truths, then that theory is a type of error theory. An **error theory** holds that normative claims are truth-apt but their truth value is always false. This is a possible interpretation of embedded epistemic instrumentalism, and this interpretation is the focus of chapter 6 (see section 5.4 for some distinctions between error theories).

If a theory holds that there are some normative truths, then that theory is a type of **realism**. Realist theories are further divided. The third choice-point is: are these truths irreducibly normative? While some comments below complicate the picture (see section 5.2), on first pass, if a theory holds that these truths are irreducibly normative then that theory is a type of **non-natural realism**. Non-natural realism holds that normativity is somehow fundamental. Consider the original question: “You ought not believe ‘The Earth is flat’; why?” A non-reductive answer holds that the fact that you ought not believe this is a fundamental fact about the world—in some sense there is no further explanation possible. One way we can characterise the divide between a reductive and a non-reductive metaphysics is to point to what we consider the same. A reductive metaphysics holds that normative properties are fundamentally the same as some, perhaps complex, collection of non-normative properties, similar to how a rabbit is fundamentally the same as a collection of body parts arranged in a certain way. For another example: a population of rabbits is fundamentally the same as a collection of rabbits with certain relations. A non-reductive metaphysics holds that at least some normative properties are fundamentally different to non-normative properties. Some normative properties have no non-normative parts, so they are unlike rabbits and populations of rabbits (Moore, 1903, pp. 7-8). As embedded epistemic instrumentalism offers a reduction, it is not a type of **non-natural realism**.

If a theory holds that these truths are not irreducibly normative, then that theory is a type of **natural realism** (see section 5.1.1 for a note on why these theories are called ‘natural’; see section 5.2 for some complications about the term ‘reductive’). Natural realist theories are further divided. The division turns on the nature of the supposed reduction: whether it is conceptual or ontological. You might hold that the reduction of the normative to the non-normative is conceptual. The usual example of a conceptual reduction is ‘bachelor’ to ‘unmarried’ and ‘man’. Arguably, English speakers could reflect on the concept of ‘bachelor’ and reasonably arrive at the conclusion that the concept reduces to the concepts of ‘unmarried’ and ‘man’. This is contrasted with an ontological

reduction. The usual example is ‘water’. English speakers could not reflect on the concept alone and reasonably arrive at the conclusion that ‘water’ reduces to the chemical kind H_2O . But by examining water we have discovered that it in fact (in some sense) reduces to H_2O . So while the concept ‘water’ may not be reductive, the thing water reduces to a non-water chemical kind H_2O , at least on first pass.⁷¹ Natural realists are all reductive but they are divided by how the normative reduces to the non-normative.

So the fourth choice-point is: are normative *concepts* irreducibly normative? If a theory holds that normative truths reduce to non-normative truths but that normative *concepts* are irreducibly normative, then that theory is a type of **non-analytic natural realism**. A non-analytic natural realist holds that normative concepts are not reducible to non-normative concepts in the way that the concept ‘water’ is not reducible to chemical kind concepts even though water reduces to H_2O (in some sense). This is a possible interpretation of embedded epistemic instrumentalism. This interpretation of my account holds that epistemic normativity is like water, conceptually irreducible but ontologically reducible.

If a natural realist theory holds that normative *concepts* are not irreducibly normative then that theory is a type of **analytic natural realism**. Analytic natural realists hold that normative concepts are reducible to non-normative concepts the way ‘bachelor’ is reducible to ‘unmarried’ and ‘man’. I doubt anyone will be motivated to offer embedded epistemic instrumentalism as an analytic reduction. While I may be wrong, my sense is that ‘usefulness’ is not a plausible conceptual reduction of ‘epistemic reasons’.

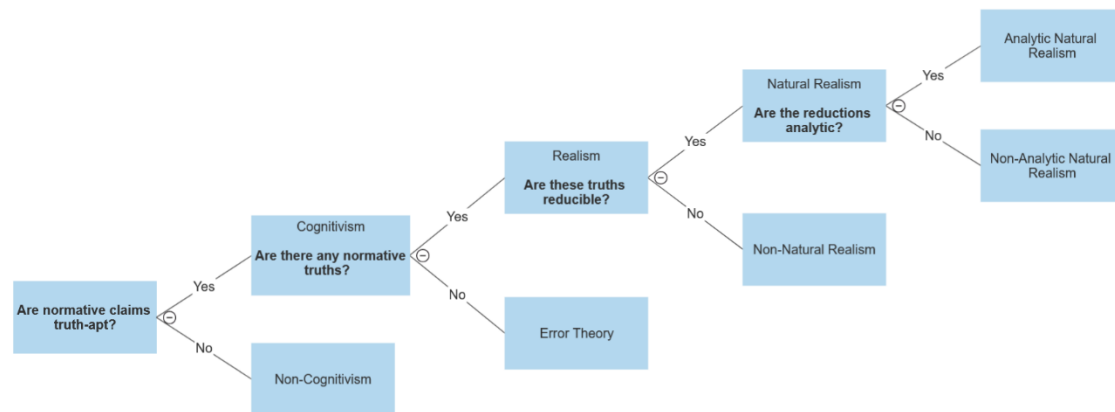
There are no further divisions in Parfit’s taxonomy. To summarise, a theory is either non-cognitivist, an error theory, non-natural realist, non-analytic natural realist or analytic natural realist.

⁷¹ Some philosophers deny, or at least complicate, the reduction of water to the chemical kind H_2O (Weisberg, 2006). Philosophers offering a reduction of normativity should welcome the complications and even the denials that water reduces to H_2O . Philosophers have a reasonable understanding of the metaphysical relationship between the stuff that we drink, the stuff that makes up the oceans, rivers, etc. and the chemical kind H_2O . But spelling out the semantics, spelling out how the English term ‘water’ relates to the chemical kind H_2O is not so straight forward. In the case of normativity, some philosophers, such as non-naturalists, point to the difficulty of relating normative terms to natural facts as evidence that normativity is different to natural phenomena like water (Hare, 1964; Horgan & Timmons, 1991, 1992a, 1992b; Moore, 1903; Parfit, 2011; Rosati, 1995; Shafer-Landau, 2003; Veluwenkamp, 2017). But the denials and complications of the reduction of water to H_2O suggests that normativity is just like other natural phenomena!

See section 5.2 for some more complications about reductions. The upshot of that section is that a better model for the reductive relationship of normative facts to non-normative facts might be the relationship between clocks and the things that make-up those clocks.

Of these, embedded epistemic instrumentalism is plausibly interpreted as non-analytic natural realism or an error theory.

Decision tree for classifying meta-normative theories.



5.1.1. On the term ‘natural’

Why call theories that posit a conceptual or ontological reduction ‘natural’? The answer is that most (but not all) theories that fall into the category ‘natural realism’ share a constraint. They only posit natural entities and properties. Of course, the term ‘natural’ is itself contested and subject to many controversies in philosophy. (This is why I avoided it in chapter 4!) So to explain the constraint, I will discuss two imperfect strategies for distinguishing the natural from the non-natural.

One strategy to understand ‘natural’ and its opposite ‘non-natural’ is to list natural and non-natural properties. Typical natural things include the property of being a chair, being useful, tigers and an electron. There are no non-controversial non-natural things. Perhaps the least controversial is a *sui generis* normative property. Nonetheless other supposed non-natural things are abstract objects like the number 2 or the possible worlds in which I was not writing this sentence. A God or a ghost or magic are clearly not natural things either, however non-naturalists might call these supernatural rather than non-natural; non-naturalists tend to not want their metaphysical beliefs to be lumped in with the beliefs of the religious and beliefs in psychics.

A problem with this listing strategy is that it lacks precision. ‘Natural’ and ‘non-natural’ are semi-technical notions in philosophy (just like ‘epistemic’: see section 1.3.4). While we can develop a sense of which things will be grouped as natural and non-natural, it is hard to pin down exactly

why we are grouping these things together. Moreover, the metaphysics of abstract objects is also a matter of controversy. So, the issue of finding what unifies the non-natural properties is made doubly hard. Merely listing natural properties and non-natural things is not incredibly effective at settling the lines of the dialectic.

Another strategy is to pick out the properties that have a certain relationship with the natural sciences (Shafer-Landau, 2005, pp. 58-62). For example, we might hold that the natural properties are the ones that the scientists currently use in their theories. A problem with this strategy is that almost certainly the scientists will change what properties they use in the future. Another relation with the natural sciences might be: the natural properties are the properties that a final science will use. Here the thought is that when science is complete, the properties it refers to will be the natural properties. There are many problems with this approach. I will offer two. Firstly, there is no guarantee that science will ever finish. Secondly, even if science will finish, it is not finished yet, so it is unclear how this strategy can help us identify the distinction between properties today.

I do not know the solution to this problem of distinguishing between the natural properties and the non-natural properties. Nevertheless, armed with the imperfect strategies I can talk about what most ‘natural realists’ share. There is a serious worry at the heart of the natural and non-natural divide. Naturalists typically hold that our greatest epistemic achievements have come from the natural sciences and mathematics. Moreover, they typically hold that no significant epistemic achievement has come from mysticism. To the extent that our investigation into the normative character of the world looks more like mysticism rather than the natural sciences or mathematics, we should be very worried about how effective such an investigation is. I feel the force of such a worry.

There is a strong motivation in philosophy to create *natural* theories. In the latest Philpapers survey by Bourget and Chalmers, ‘naturalism’ is the most endorsed metaphilosophy out of ‘naturalism’, ‘non-naturalism’ and ‘other’.⁷² My sense is that naturalistic philosophers are motivated to construct theories that are continuous with the sciences because they think that these are the type of theories that are most likely true, or at least most likely closer to the truth than other types of theories. Natural theories are constrained; they only posit entities and properties that we would list as ‘natural’. For example, embedded epistemic instrumentalism posits the property of usefulness.

⁷² An important caveat: even though philosophers could endorse all options, only just over half of respondents endorsed naturalism as a metaphilosophy.

This is a natural property. This property is amenable to empirical enquiry. Indeed, engineering and medicine are two of many disciplines that study the usefulness of materials and actions empirically. So embedded epistemic instrumentalism is a natural theory. Most of the theories that fall into the category ‘natural realism’ only posit natural entities and properties. The philosophers that offer natural realist theories are aiming to create something that is continuous with the sciences. This is why we might call the relevant category ‘natural realism’.

However, not all the theories that fall into the category ‘natural realism’ share these motivations and constraints. Divine command theory holds that what a god wills is what ought to be done (M. Murphy, 2019).⁷³ Divine command theory is reductive. There are analytic and non-analytic versions of divine command theory. Some versions of divine command theory hold that concepts like ‘ought to be done’ reduce to concepts like ‘what a god wills’. Other versions hold that normative properties ontologically reduce to commands of a god. On Parfit’s taxonomy, both are a type of natural realism. Of course, philosophers motivated by naturalistic concerns will reject divine command theory because it is not continuous with the sciences and posits entities that are not natural. Gods are not amenable to empirical enquiry. This categorisation of divine command theory as ‘natural realist’ is infelicitous. But as is a common theme in the thesis, this is just one way to keep the books. So, for this thesis, nothing of substance hangs on calling a theory ‘natural realist’. A rose by any other name...

5.2. Some Additions to Parfit’s’ Taxonomy

Parfit’s taxonomy gives us a good sense of the range of options and how the philosophical community divides the space of possibilities. However, there are some distinctions and conceptual tools I wish to add to Parfit’s taxonomy to enrich our understanding of the possibilities, and ultimately the interpretations of embedded epistemic instrumentalism.

Something that is not strictly true but is a useful heuristic: non-natural theories of normativity are associated with a non-reductive metaphysics and natural theories with some type of reductive metaphysics. This is not strictly true because there are two distinctions that complicate the picture. The first distinction is between **type** non-reductionism and **token** non-reductionism. The second distinction is between ontologically committed non-naturalism and ontologically uncommitted non-naturalism. I will discuss these two distinctions in the next two sub-sections.

⁷³ Divine command theory is also known as ‘theological voluntarism’.

5.2.1. Token non-reductivism and type non-reductivism

The term ‘non-reductive’ has its own baggage. On some taxonomies, not all non-reductive realist theories are non-natural theories. For example, Cornell realism is considered natural and by some, non-reductive (Boyd, 1988; Brink, 1989; Sayre-McCord, 1988; Sturgeon, 1984). The reason Cornell realists hold a qualified version of the reductive claim is that they hold a distinction between token-reductions and type-reductions. Cornell realists hold that *type* normative properties are not reducible to *type* non-normative properties.⁷⁴ Type normative properties do not reduce to any single type of non-normative thing. Clocks are not reducible in this sense. Clocks can be made out of stone (a sundial), sand and glass (a hour glass), can be electronic (the clock on your computer), can be hugely complicated (an atomic clock) or even naturally occurring (in some sense, the solar system is a clock and humans have used it as such for millennia). Indeed even you are a clock. The *type clock* is not reducible to any single type of non-clock thing. Indeed, as the list above indicates, clocks can be made out of many different non-clock things. This is to say clocks are **multiply realisable**. Cornell realists hold that normative properties are similarly multiply realisable and therefore are not reducible (in some sense of reducible). Indeed, as usefulness itself is multiply realisable, embedded epistemic instrumentalism also offers this type of metaphysics.

Neil Sinhababu argues that the *non-reductivism* of Cornell realists is primarily motivated by a concern with multiple realisability of normative properties (Sinhababu, 2018, pp. 35-41).⁷⁵ I do not believe it matters whether we call this type of multiply realisable metaphysics reductive or not. But for the purpose of keeping the books in this thesis I do call this a reductive metaphysics. For this thesis, a normative metaphysics is reductive if all token normative properties can be reduced to token non-normative properties. In this qualified sense, all natural theories, all non-cognitivist theories and all error theories offer a reductive metaphysics. Indeed, as I will argue below, even some non-natural realist theories offer a reductive metaphysics of this kind.

⁷⁴ To remind the reader, a **type** is a general thing and a **token** is a specific instance of that thing (Wetzel, 2018). Consider the sentence: ‘apples are red or green’. In this sentence, ‘apples’ refers to the type *apples*, the general thing. Consider this different sentence: ‘this apple was delicious.’ In this sentence, ‘this apple’ refers to a token apple, a specific instance.

⁷⁵ As well as a concern about ‘natural kinds’ (Sinhababu, 2018, pp. 35-41). This additional concern is not important for our purposes.

5.2.2. A distinction between non-natural realist theories

Non-natural realists divide into **ontologically committed** non-natural realists (Enoch, 2011; Huemer, 2005; Moore, 1903) and **ontologically uncommitted** non-natural realists (Parfit, 2011; Scanlon, 2014; Shafer-Landau, 2003). Ontologically committed non-natural realists are making much more ambitious assertions about reality. Not only is normativity fundamental, but it requires extra ontological commitments. Something exists that is very different to the stuff the natural sciences investigate, the normative stuff! According to the ontologically committed non-naturalist, for normative claims to be true, an **odd objectively normative property**⁷⁶ needs to be instantiated. Error theorists tend to agree with ontologically committed non-naturalist that, for normative claims to be true, an odd objectively normative property needs to be instantiated (see sections 5.3 and 6.2.1) but error theorists hold that no such property is instantiated.

However, it is not even clear to me that ontologically uncommitted non-naturalists are offering rival theories to natural realist theories. As I argued in chapter 1, when interpreted as offering a different notion of truth, non-natural realism is consistent with natural realism. Here, I want to make a bolder claim. When interpreted as offering a different notion of truth, ontologically uncommitted non-natural realism is an extremely similar theory to a version of error theory known as **fictionalism** (Joyce, 2001). Fictionalism is the theory that normative claims are literally false, like claims in fiction. But also like claims in fiction, they can be considered *true* with a suitable conception of truth. One such conception is an epistemic conception of truth.

An epistemic conception of truth treats the truth of a sentence as determined by the conditions under which it is established as assertable, believable or in some other way acceptable (Dummett, 1958; Prawitz, 2012, p. 12). So as the facts about when a sentence is assertable etc. change, the truth of the sentence changes. One of the motivations to posit an epistemic conception of truth is to make sense of sentences and terms that do not refer in a correspondence sense, the scientific-like sense that most of this thesis has assumed, but where there is still some reasonable sense of truth (Prawitz, 2012). For example, we might say: “In *Lord of the Rings*, Frodo was reluctant to drop the ring into Mt Doom when he had the chance.” This sentence seems true, in a sense. It is false to say: “Frodo was keen to drop the ring into Mt Doom when he had the chance.” One way to

⁷⁶ Describing what an ‘odd objectively authoritative property’ amounts to is controversial. It is important that the property is *odd* because this property is supposed to set the golden glow of normativity apart from mere descriptive properties, like the descriptive property of being useful or having an electric charge. Saying any more provokes a chorus of objections. I direct the curious reader to J. L. Mackie’s classic discussion of the queerness objection to objective values to get a sense of what an odd objectively authoritative property would be and why one might doubt the instantiation of such a property (Mackie, 1977, pp. 38-42).

make sense of the seeming truth of the former sentence is to adopt an epistemic theory of truth. The former sentence is assertable in a way the second sentence is not. This is what makes the sentence “In *Lord of the Rings*, Frodo was reluctant to drop the ring into Mt Doom when he had the chance.” true. Importantly, this is compatible with a reductive metaphysics of fiction. It does not require us to postulate some extra stuff that makes it true that ‘Frodo was reluctant...’.

Being a fictionalist about normative language is one way to be an error theorist (Joyce, 2017). The error theorist might say that normative properties do not exist in a correspondence sense, agreeing with Parfit, but nonetheless hold normative talk is an important fiction. And one way to be a fictionalist is to adopt an epistemic theory of truth about normative discourse. So when one is engaging in discussions *about* normative discourse, one adopts a correspondence theory of truth. And when one is engaging *in* normative discourse, one adopts an epistemic theory of truth. For example, this thesis, and meta-epistemology more generally, engages in discussions *about* normative discourse, so is conducted with a correspondence notion of truth. But if we were to start discussing what specific beliefs we ought to have, then we would be engaging *in* normative discourse so we would adopt an epistemic notion of truth, according to this picture.

As I claimed in chapter 1, Veluwenkamp (2017) interprets both Parfit and Scanlon as adopting some kind of truth pluralism to maintain realism about reasons. This may represent an unexpected convergence of non-natural realists and error theorists qua fictionalists, where once we fix the appropriate theory of truth, the difference between these views is at least greatly diminished, if not completely evaporated. Both agree that normative claims are not true in the correspondence sense. And both agree that there may be another sense of truth, for example an epistemic notion of truth, according to which normative claims are true.

Indeed, for another example of an unexpected convergence between non-natural realists and natural realists, Russ Shafer-Landau only asserts non-naturalism as an epistemic property. He holds that normative claims should be studied with a different method to scientific claims, but all token normative properties reduce to a combination of non-normative properties, in a similar way that populations of rabbits reduce to sets of and relations between rabbits, and rabbits themselves reduce to body parts (Shafer-Landau, 2003, p. 72). Of course, the relationship between these biological facts and non-biological facts is exactly the type of relationship that non-analytic natural realists would also point to as an example of a naturalistic metaphysics (Boyd, 1988; Brink, 1989,

2001). So it is not clear that the metaphysics posited by Shafer-Landau in anyway rivals non-analytic natural realism. Indeed, it might be substantively the same type of metaphysics.⁷⁷

That being said, as I interpret non-natural realists, none of them want to be associated with naturalistic reductions or error theories. So the conciliatory interpretations I have put forward in this section do not satisfy the intentions of the authors. For that reason, you might hold these conciliatory interpretations are bad interpretations. Fair enough. But the issue is that once you assert that your theory is ontologically uncommitted, that your claims have no ontological implications, then we should expect your theory to be consistent with all ontological claims. Indeed, if your supposed ontologically uncommitted theory is inconsistent with an ontological claim, then it is not actually ontologically uncommitted. In so far as I am trying to characterise the ontology of epistemic normativity, it should be consistent with any ontologically uncommitted theory. If the authors intend their theories to be ontologically uncommitted and also intend their theories to be rivals to other ontologies, then that is a tension in their intentions.

I consider embedded epistemic instrumentalism to be a reductive theory. It holds that epistemic reasons are instrumental reasons. But this kind of reduction, a token reduction, is a similar metaphysics to the majority of other theories of normativity. In the same way we can combine a fictionalist semantics with error theory, we can combine a fictionalist semantics with embedded epistemic instrumentalism too. This combination is consistent with ontologically uncommitted non-naturalism, at least as it is interpreted by Veluwenkamp (2017). So most theories of the metaphysics of normativity are reductive. The exception is *ontologically committed* non-natural realist theories.

⁷⁷ As an aside, I reject Shafer-Landau's epistemology. He holds that the relationship between normative and non-normative properties motivates a method of enquiry that looks very peculiar; for example, the study of normativity is not amenable to empirical enquiry on his picture. However, the examples of similar relationships to the normative and the non-normative do not require a peculiar method of enquiry. The examples of relationships similar to the normative and non-normative include: biological facts to physical facts, chess facts to physical facts, and historical facts to physical facts (Shafer-Landau, 2003, p. 72). But the study of biology, chess and history are all amenable to empirical enquiry. They are all continuous with the sciences. If physics holds that perpetual motion is impossible, historians have evidence that no one has actually invented a perpetual motion machine. So I hold that his metaphysics in no way motivates his epistemology, indeed it tells against it. But it takes us too far off course to develop this thought in detail here.

5.3. Two-theses Classification of Theories

I now want to relate Parfit's classifications to another way of classifying types of normative theories. This new way is less informative about the way the classifications relate to each other (which is why I started with Parfit's). But it will be useful in the following sections and chapters.

Non-cognitivism, error theory, non-natural realism, non-analytic natural realism and analytic natural realism can all be classified with two theses each. I call the first thesis the 'what-it-takes thesis'. I call the second thesis 'the metaphysical thesis'. The what-it-takes thesis is a thesis about what is required for the relevant sense of normativity to exist. As we are talking about epistemic normativity, we can consider the what-it-takes thesis to be a thesis about what it would take, a necessary condition if not sufficient condition, for epistemic reasons to exist. The metaphysical thesis is about whether such a condition is satisfied.

The two theses that classify an epistemic version of non-cognitivism:

What-it-takes thesis: For epistemic reasons to exist, positive and negative attitudes need to exist—similar to 'boo Collingwood football club' but directed at beliefs in various contexts rather than a football club, for example 'boo your belief that non-natural realism is true'.

Metaphysical thesis: such attitudes exist.

For error theory (see section 6.2.1):

The What-It-Takes Thesis: For epistemic reasons to exist, odd objectively authoritative properties need to be instantiated.

The Metaphysical Thesis: Odd objectively authoritative properties are not instantiated.

For ontologically committed non-natural realism:

The What-It-Takes Thesis: For epistemic reasons to exist, odd objectively authoritative properties need to be instantiated.

The Metaphysical Thesis: Odd objectively authoritative properties are instantiated.

For ontologically uncommitted non-natural realism, as interpreted by Veluwenkamp:

The What-It-Takes Thesis: For epistemic reasons to exist, normative sentences need to be true, according to an epistemic notion of truth.

The Metaphysical Thesis: some normative sentences are true in this way.

For non-analytic natural realism:

The What-It-Takes Thesis: For epistemic reasons to exist, some collection of natural facts exists such that they give rise to the phenomena we are pointing to with the term ‘epistemic reasons’.

The Metaphysical Thesis: such a collection of natural facts exists.

For analytic natural realism:

The What-It-Takes Thesis: For epistemic reasons to exist, whatever the conceptual reduction of ‘epistemic reasons’ refers to needs to exist.

The Metaphysical Thesis: the referent of such a conceptual reduction exists.

Note that I have not specified what facts would settle which what-it-takes thesis is correct. This is related to the metaphilosophical controversies about the use of intuitions in philosophy, discussed in section 4.2.1. It is unclear whether the relevant facts are semantic, about the content of concepts, about facts in the world, about intentions or beliefs, or some complex combination of these facts. I am happy for my interlocutors to set which facts are relevant and I am happy for my interlocutors to set what the relevant facts are. Depending on how they do that determines which interpretation of embedded epistemic instrumentalism is appropriate.

If they hold, like Kornblith (2007) and Williamson (2007), that it is facts about what we are pointing at that settle the what-it-takes thesis, then non-analytic natural realism is the appropriate interpretation. When we talk about cases of believing and what forces are at play that push us to believe in certain ways, then the instrumental reasons are it. If like Jackson (1998) they hold it is the platitudes about the term in question of a relevant population (in this case the platitudes of philosophers about the term ‘epistemic reasons’ or ‘epistemic normativity’), then it depends on what those platitudes are. Perhaps philosophers will agree that ‘epistemic reasons’ are whatever it is that compels us to conform to the epistemic standards. In this case, non-analytic natural realism

is the appropriate interpretation. Perhaps philosophers will have more demanding platitudes, that ‘epistemic reasons’ are something fundamentally different from non-normative facts. In this case, error theory is the appropriate interpretation. On this interpretation ‘epistemic reasons’ does not actually refer, but there is something that plays the role of epistemic reasons, the special set of instrumental reasons that I have identified in this thesis. In general, if you think that embedded epistemic instrumentalism requires some revision of the concept or the term or our core beliefs about ‘epistemic reasons’ then the error theory interpretation is appropriate. If you do not hold that it demands revision, then non-analytic natural realism is appropriate. Either way, the case for embedded epistemic instrumentalism is strong.

5.4. A Distinction Between Error Theories

Here I follow an argument made by Kim Sterelny and Ben Fraser (2017). Parfit’s taxonomy also does not distinguish between two important types of error theories. To understand this distinction, note that the two-thesis way of classifying theories (as well as Parfit’s taxonomy) is not particularly discerning.

To see this, consider how such an analogous two-thesis scheme classifies scientific theories and folk tales. Firstly, let us consider Newtonian physics. For Newtonian physics to be true, it would take, among other things, Mercury’s orbit to conform to Newton’s laws of motion. It does not. This means that we are ‘error theorists’ about Newtonian physics according to the scheme. Secondly, let us consider what it would take for *The Little Mermaid* to be true. It would take, among other things, for mermaids to exist. They do not. So we are ‘error theorists’ about *The Little Mermaid*, too. The issue is that this classification scheme groups our attitudes about Newtonian physics with our attitudes about *The Little Mermaid*. Both are error theories. But Newtonian physics is much more accurate, it is vindicated to a much larger extent by the facts of the world, than *The Little Mermaid*. We might distinguish these types of error theories. We are **partial realists** about Newtonian metaphysics, holding that it is a fairly accurate representation within certain bounds. And we are **complete error theorists** about *The Little Mermaid*, holding that it is not a very accurate representation of reality.

Sterelny and Fraser are motivated by the thought that even if there is some mismatch between the moral platitudes that constitutes our folk moral theory and reality, it is partially vindicated in the way Newtonian physics is partially vindicated by reality. While I am not committed to a theory about moral normativity, I note that embedded epistemic instrumentalism, even when interpreted

as an error theory, is extremely vindicatory of our epistemic practices. There is a species of reasons that can be studied relatively independently of other reasons, that are very important for human life, that have had a large influence on biological and cultural development. These are the special group of instrumental reasons I identified in this thesis. This vindicates our practice of talking and theorising about epistemic reasons, even if there is some mismatch between our prior concepts (etc.) of ‘epistemic reasons’ and reality.

5.5. The Two Relevant Interpretations

With all the distinctions and different ways of classifying theories on the table, I will now offer the interpretations of embedded epistemic instrumentalism. The three relevant interpretations from Parfit’s taxonomy are non-cognitivism, error theory and non-analytic natural realism. As embedded epistemic instrumentalism holds that there are true and false claims about what epistemic reasons support, it is not easily a non-cognitivist theory. But, strictly speaking, it is compatible with a non-cognitivist semantics. It holds that these semantics do not give us much insight into the metaphysics of epistemic normativity. Embedded epistemic instrumentalism is not a non-naturalist realist theory because it does not hold that normativity is fundamental. It may be consistent with ontologically uncommitted non-natural realism, if not the motivations for such a theory. But I interpret the authors of such theories to be hostile to theories like mine. So there is some exegesis needed to identify exactly what is inconsistent. I have also rejected the idea that there is an analytic reduction. Strictly speaking if an interlocutor wanted to assert that ‘epistemic reasons’ just means ‘the sub-set of instrumental reasons that I identify in this thesis’ then embedded epistemic instrumentalism could be a version of analytic natural realism, but I doubt any philosopher would find this plausible. This leaves error theory and non-analytic natural realism.

5.5.1. Embedded Epistemic Instrumentalism as non-analytic natural realism

The non-analytic natural realist interpretation of embedded epistemic instrumentalism is appropriate if:

Non-analytic natural realist condition

Our terms, or concepts, or platitudes, or however it is we coordinate on the topic of ‘epistemic normativity’, do not preclude a reductive metaphysics, nor the reduction of epistemic reasons to instrumental reasons.

There are many ways this non-analytic natural realist condition might obtain. For example, it might be that terms like ‘epistemic normativity’ just point at some part of the world, just point to cases of humans being compelled to believe in certain ways, criticizing others for not believing in certain ways, expending resources and effort to attempt to believe in certain ways. So whatever they are pointing at, whatever gives rise to those things, constitutes the referent of ‘epistemic normativity’. Another way the condition might obtain is if the public concept ‘epistemic normativity’ is incredibly vague. It could be that philosophers have idiosyncratic concepts of ‘epistemic normativity’ but there is very little overlap in content. Some may have concepts that preclude a natural realist theory. Some may not. But on this picture, there would be no substantial agreement. In this case the concept is just vague. Embedded epistemic instrumentalism is consistent with one way we might make that vague concept more precise. Broadly speaking, the non-analytic natural realist interpretation is appropriate if one holds that the *what-it-takes thesis* for ‘epistemic reasons’ is not particularly metaphysically demanding.

On this interpretation, embedded epistemic instrumentalism is a description of epistemic normativity. The model is that epistemic normativity is like friction or heat, a force in the aggregate. It holds that all token epistemic reasons are reducible to some set of physical properties. But it does not hold that there will be a type reduction of epistemic reasons to any particular set of physical things. The metaphysics of epistemic normativity is similar to the metaphysics of a clock. Clocks are multiply realisable. Many different collections of physical stuff constitute a clock. Ditto for epistemic reasons. These collections are what we are pointing at with the term or concept ‘epistemic normativity’, or what we intend to refer to, or what it is we are coordinating on (however it is we coordinate on philosophical topics).

5.5.2. Embedded Epistemic Instrumentalism as an Error Theory

The error theory interpretation of embedded epistemic instrumentalism is appropriate if:

Error theory condition

Our terms, or concepts, or platitudes, or however it is we coordinate on the topic of ‘epistemic normativity’, preclude a reductive metaphysics, or the reduction of epistemic reasons to instrumental reasons.

There are many ways the error theory condition may obtain. It would obtain if the term ‘epistemic normativity’ were shorthand for ‘the odd objectively authoritative property that makes it the case

that, all things considered, that you ought to believe in a way that conforms to the epistemic standards' then that would rule out a non-analytic natural realist account of epistemic normativity. Similarly, if the platitudes about 'epistemic normativity' and 'epistemic reasons' that philosophers agree on precluded a reductive metaphysics and a reduction of epistemic reasons to instrumental reasons, say by referring to an odd objectively authoritative property, then the error theory condition would also obtain. In general, the error theory interpretation is appropriate if the *what-it-takes thesis* for 'epistemic reasons' is metaphysically demanding, requiring some difficult-to-define odd objectively authoritative property.

The error theory interpretation holds that the term 'epistemic reasons' fails to refer to anything. For example, it might be that the term 'epistemic reasons' refers to an odd objectively authoritative property but no such property is instantiated. Nevertheless, the error theory interpretation also holds there is a part of the world that plays the epistemic-reasons-role, that compels us to believe in certain ways. That part of the world is the special sub-set of instrumental reasons that I identified in this thesis. The extension of the term or the concept or our collection of beliefs about 'epistemic reasons' is largely accurate. When we think there are epistemic reasons, we are fairly accurate at tracking the instrumental reasons that play the epistemic-reason role. There is a little mismatch in extreme cases but that is all. In particular, this error theory holds the same attitude to 'epistemic normativity' as we might hold to Newtonian physics, a largely accurate and useful model that is wrong in extreme cases. This is a form of partial realism. On this interpretation, embedded epistemic instrumentalism is a refinement on the conceptual scheme involving 'epistemic reasons'.

There is a remaining problem, however. Some philosophers hold that an epistemic error theory, such as the error theory interpretation of embedded epistemic instrumentalism, is self-undermining. I address this concern in the next chapter.

6. Epistemic Error Theory and the Self-Undermining Objection

6.1. Introduction To the Self-Undermining Objection

As I argued in chapter 5, embedded epistemic instrumentalism may be interpreted as a type of epistemic error theory.⁷⁸ epistemic error theory holds that there are no epistemic reasons, strictly speaking. Many interlocutors have charged this theory and related positions with being self-undermining (Case, 2020; Copp, 1995, pp. 46-48; Cuneo, 2007, chapter 4; Husi, 2013; Joyce, 2001, pp. 49-52; Nagel, 1997, pp. 19-26; Streumer, 2013, 2017). The thought is that certain actions, such as arguing for epistemic error theory, will in some sense presuppose that there are epistemic reasons. So performing these actions in relation to the theory, such as arguing for it, will be self-undermining. Let us call this objection the **self-undermining objection**.

In this chapter, I argue that the self-undermining objection fails. Epistemic error theory doesn't undermine itself. However, this is not obvious. The central theme of the objection—that arguing for epistemic error theory somehow suggests the existence of epistemic reasons—can be given in many different ways. Rather than focusing on arguing (Cuneo, 2007, chapter 4), the objector might focus on believing (Streumer, 2013), accepting norms (Husi, 2013), reasons (Case, 2020) and many other actions. There are a lot of different actual and possible versions of this objection; Terence Cuneo offers three versions himself (2007, chapter 4). So it is very easy to talk past interlocutors or have merely verbal disputes, which I take it no one in this debate wants.

My methodological reaction to this situation, a theme of this entire thesis, is to never insist on an exclusive definition of a term to make an argument; while I stipulate the meaning of some terms, they are never supposed to replace someone else's definition (I am always stipulating the meaning of a new term). So my definitions are always a matter of book-keeping rather than a matter of substance. And I will grant my interlocutors any term they wish. This requires a lot of focus on terms, and a lot of exegesis. But the upshot is that by the end of this chapter, we will have a way of responding to all possible versions of the self-undermining objection. I aim for a robust refutation, not merely another epicycle in the dialectic.

⁷⁸ Epistemic error theory goes by many names (see section 6.2).

The dialectic has several moves. epistemic error theorists offer their theory. Self-undermining objectors then offer their objection. epistemic error theorists then offer something I am calling **the not-so-bad reply**. And Spencer Case (2020) has responded to this reply. In the next four sections of this chapter (sections 6.2-5), I aim to build a general model of this complicated dialectic. Even if philosophers are unpersuaded by my position, part of the aim of this chapter is to make a murky debate more transparent. Then in section 6.6, I go on the offensive. I offer an extended analogy. I compare the epistemic error theorist to a spirit error theorist, someone denying the existence of spirits in a fictitious community where spirit-talk dominates epistemic practice. From our perspective we can see that the spirit error theorist is not self-undermining and by analogy, neither is the epistemic error theorist. In section 6.7, I offer a conjecture about a deep flaw in the self-undermining objection.

Importantly, everything I say here is compatible with epistemic error theory being false. In this chapter, I only aim to reject the self-undermining objection. If you are partial to an error theory interpretation, the positive case for embedded epistemic instrumentalism is supposed provide the evidence for epistemic error theory.

6.2. Epistemic Error Theory

The dialectic about epistemic error theory and the self-undermining objection is one that requires a lot of book-keeping—a lot of concepts and subtle distinctions need to be defined. Different authors keep their books differently (cf. Case, 2020; Copp, 1995, pp. 46-48; Husi, 2013; Joyce, 2001, pp. 49-52; Nagel, 1997, pp. 19-26; Streumer, 2013). In this section I will define epistemic error theory, comparing my formulation to the various other ways it has been defined in the literature, along with some related concepts and theories. The aim of this section is to coordinate the various ways the books are kept so we can avoid talking past any interlocutors.

6.2.1. My formulation

Epistemic error theory denies that epistemic reasons exist. Following the two-theses classification from section 5.3, we can break this claim into two parts: The What-It-Takes Thesis and The Metaphysical Thesis.

The What-It-Takes Thesis: For epistemic reasons to exist, odd objectively authoritative properties need to be instantiated.

The Metaphysical Thesis: Odd objectively authoritative properties are not instantiated.

The What-It-Takes Thesis posits a necessary condition for ‘epistemic reasons’ to exist, the existence of ‘odd objectively authoritative properties’.⁷⁹ The Metaphysical Thesis denies such a property is instantiated, whatever it is.

Strictly speaking, we need not define epistemic error theory with this two-part structure. Someone might hold an alternative what-it-takes thesis and still deny that epistemic reasons exist; they might hold that epistemic reasons are propositions that are absolutely certain, and no proposition is absolutely certain. We might define epistemic error theory in such a way to capture this possibility. However, there are some motivations not to do this. One motivation is that the literature usually defines error theories with reference to J. L. Mackie (1977) and Mackie’s position has this two-part structure. Another motivation is that while some philosophers deny The What-it-takes Thesis and some deny The Metaphysical Thesis, I know of no one that denies both. So I consider this two-part theory to be the theory under discussion, rather than the bare denial that epistemic reasons exist.

That being said, in this discussion, it does not matter what an odd objectively authoritative property is; I treat it as a variable with no specific content, just a label. So my rejection of the self-undermining objection will also work with alternative formulations of epistemic error theory, even though I have followed the convention of defining it in two parts.

Importantly, I am not a committed epistemic error theorist, because I do not endorse The What-It-Takes Thesis unconditionally. As I stated in section 5.3, and as we will see below, sometimes The What-It-Takes Thesis is interpreted as a claim about the meaning of a term, that the meaning of the term ‘epistemic reasons’ implies that odd objectively authoritative properties are necessary for epistemic reasons to exist (Cuneo, 2007, p. 115; see section 6.2.2.1). Sometimes The What-It-Takes Thesis is interpreted as a claim about the content of a belief; when people have beliefs about epistemic reasons, the content of those beliefs imply that odd objectively authoritative properties are instantiated (Streumer, 2013, p. 194; see section 6.2.2.2). Indeed, there is a general controversy about the content of theses like The What-It-Takes Thesis, about what such theses are actually claiming (cf. Bealer, 1998; Jackson, 1998; Kornblith, 2007; Williamson, 2007), so there are many possible interpretations of The What-It-Takes Thesis. One aim of my formulation is to remain as

⁷⁹ For this discussion, the difference between properties and relations does not matter.

uncommitted as possible about these meta-philosophical controversies, to remain uncommitted about how philosophers actually coordinate on what it would take for a thing to exist.

Whether I endorse 'The What-It-Takes Thesis' depends on the type of claim it is. It depends on the specifics of the philosophical task we are engaging in. I doubt the instantiation of odd objectively authoritative properties is implied by the meaning of the term 'epistemic reasons'. However, I am open to persuasion on this point. The type of evidence that would convince me would come from a combination of corpus analysis and survey methods of English speakers, in particular philosophers as I doubt 'epistemic reasons' is a term in standard English (see section 1.3.3-4 and 4.3). However, if we interpret 'The What-It-Takes Thesis' as a claim about the content of a belief, and the content of a belief can be very idiosyncratic, then I think it is much more likely that 'The What-It-Takes Thesis' is true for certain people at certain times. Indeed, I suspect it has been true about the content of my beliefs in the past.

Of crucial importance, and a central motivation for this chapter, is my doubting that 'The What-It-Takes Thesis' provides any significant evidence for 'The Metaphysical Thesis'. One of the ways the 'What-It-Takes Thesis' could potentially provide evidence for 'The Metaphysical Thesis' is via the self-undermining objection. Supposedly, one cannot hold that *the meaning of a term or the content of a belief implies that X is necessary for epistemic reasons to exist and X doesn't exist* without undermining oneself. If that is true, then whether 'The What-It-Takes Thesis' is true becomes very important to settle. Indeed, if the self-undermining objection is successful, settling what it actually takes for epistemic reasons to exist is *incredibly* informative. If God is required for epistemic reasons to exist,⁸⁰ then via the self-undermining objection, we have just learnt that one cannot deny the existence of God without undermining oneself. And it should come as no surprise that philosophers have offered similar arguments for the existence of God, in spirit if not exactly the same (Evans, 2018). So even though I do not endorse epistemic error theory, I cannot abide the self-undermining objection because, methodologically speaking, it sends philosophers up the garden path.

6.2.2. Other formulations

I will now compare my formulation of the target thesis with various others found in the literature.

⁸⁰ Why might one hold this?—epistemic reasons are normative and normativity requires the commands of a divine being, or so say divine command theorists.

6.2.2.1. Cuneo's

Cuneo analyses the self-undermining objection as it applies to 'Epistemic Nihilism' (2007, chapter 4). He defines 'Epistemic Nihilism' as holding three theses:

The Epistemic Realist's Speech Act Thesis: Some epistemic discourse is assertoric (Cuneo, 2007, p. 115).

The Epistemic Nihilist's Alethic Thesis: If the content of a predicative epistemic claim is true, then it is true in the realist sense, and no such content is true (Cuneo, 2007, p. 116).

The Epistemic Nihilists Ontic Thesis: If there are any particular epistemic facts, then they are as the realist says, and there are no such facts (Cuneo, 2007, p. 116).

While this formulation may seem rather different to mine, both Cuneo and I are pointing at the same idea. Cuneo is explicit that Epistemic Nihilism is the epistemic analogue of the more general error theory, specifically referencing Mackie (Cuneo, 2007, p. 115). I might have followed Cuneo and called our target theory 'Epistemic Nihilism'. However, I chose 'epistemic error theory' because it is easier to see how this theory is related to slightly different theories that are discussed in this literature and in this chapter

While we point to the same idea, Cuneo's formulation of the target theory is more specific than my formulation. The Epistemic Realist's Speech Act Thesis and The epistemic error theorist's Alethic Thesis together play the same functional role as The What-It-Takes Thesis from my formulation—that is together the two theses specify what it would take for epistemic reasons to exist. I will explain each in a little more detail, and some of the motivation for formulating the theory like Cuneo does.

The Epistemic Realist's Speech Act Thesis is a claim about semantics or meta-semantics (depending on how you distinguish those topics). It claims that epistemic-talk asserts something about the world. It claims that epistemic-talk is truth apt, like the sentence 'There is an apple in front of you.' and unlike the sentence 'Hooray!'. This claim rules out non-cognitivist theories of epistemic normativity.⁸¹

The Epistemic Nihilist's Alethic Thesis is a claim about truth and a claim about content. As I interpret it, the claim about truth is supposed to rule out ontologically uncommitted ways a sentence might be true. For example, a version of an epistemic notion of truth holds that a sentence

⁸¹ For an introduction to non-cognitivism about moral normativity, see: Roojen (2018); also see section 5.1.

is true just in case it is acceptable to assert the sentence (Prawitz, 2012). So, on an epistemic notion of truth, the sentence ‘odd objectively authoritative properties are instantiated’ is true just because it is acceptable to assert the sentence, even if there is no corresponding property in the world. I take Cuneo to be ruling out this type of theory of truth.

Unfortunately, Cuneo attempts to do this by pointing to a realist notion of truth. However, as I say above, realists no longer have an agreed notion of truth. Some hold that normative talk is not metaphysically loaded, and one interpretation of this rejection of metaphysical import is that the concept of truth differs between normative talk and other types of talk, such as scientific discourse or every day discourse (Veluwenkamp, 2017). In ordinary everyday discourse, if I say ‘there are apples in the basket’ I am committed metaphysically to, among other things, the existence of apples; I would have uttered a falsehood if, to my surprise, apples did not exist in the way tables, water and fire exists. Not so for epistemic discourse with a pluralist notion of truth. We might say ‘there are epistemic reasons not to believe the earth is flat’ and not utter a falsehood even if epistemic reasons do not exist like tables, fire and water. Other realists hold that normative talk is metaphysically loaded, implying no difference in the notion of truth between types of talk (Enoch, 2011; Huemer, 2005). epistemic error theorists hold that talk about epistemic reasons is metaphysically loaded. As is standard for this thesis, when I use phrases like ‘odd objectively authoritative properties are instantiated’ interpret them as metaphysically loaded claims, on a par with claiming that electrons exist.

The claim about the content of epistemic talk is supposed to identify a part of reality that would need to exist for epistemic reasons to exist. Cuneo refers to this thing by referring to ‘realists’ and ‘epistemic discourse’. Interpret ‘epistemic discourse’ as only the epistemic-talk that is normatively valenced. So talk about the fact that certain animals are sensitive to changes in their environment will not count as epistemic discourse. Indeed, even knowledge-talk may not count, depending on whether you hold that talk of knowledge commits you to the existence of epistemic reasons or not.

Cuneo refers to ‘realist content’ to specify what is required for epistemic reasons to exist. A problem with this approach is that not all realists agree. For example, when it comes to moral normativity, some realists do not require any odd objectively authoritative properties (Brink, 1989) while others do (Enoch, 2011). epistemic error theorists agree with ontologically committed non-natural realists that odd objectively authoritative properties are necessary for epistemic reasons to

exist, like David Enoch (2011) and Michael Huemer (2005), and arguably G. E. Moore (1903). I take it Cuneo is trying to point to the content that ontologically committed non-natural realists point to. But for this thesis, it does not actually matter.⁸² So when I use the term ‘odd objectively authoritative properties’ interpret that phrase as referring to the same thing that Cuneo’s realists refer to when they specify what is required for epistemic reasons to exist.

Recall that sometimes the What-It-Takes Thesis is interpreted as a claim about the meanings of words. With this semantic interpretation, The Epistemic Realist’s Speech Act Thesis and The Epistemic Nihilist’s Alethic Thesis together amount to a very similar claim to the What-It-Takes Thesis. But the What-It-Takes Thesis is open to the other ways, non-semantic ways, we might identify what it would take for epistemic reasons to exist.

Finally, I note that The Epistemic Nihilist’s Ontic Thesis and The Metaphysical Thesis are functionally the same; they both claim that the odd objectively authoritative properties—the things that realists require—are not instantiated. So despite appearances, both Cuneo and I are trying to discuss the same idea. We are just characterizing it in different ways.

6.2.2.2. Streumer’s

Bart Streumer analyses the self-undermining objection in regard to ‘error theory’ (2013).⁸³ Error theory holds that “normative judgements are beliefs that ascribe normative properties, even though such properties do not exist” [footnote removed] (Streumer, 2013, p. 194). Note the two-part definition of the target theory in this quote. The analogous element to The What-It-Takes Thesis is the clause: ‘normative judgements are beliefs that ascribe normative properties’. This clause is a more committed version of The What-It-Takes Thesis. This clause holds that we have a thing called a ‘normative judgement’. This thing is a belief. And the content of this belief in some sense presupposes that ‘normative properties exist’. Not everyone is going to accept this formulation of how we come to recognise the relevant facts about what it would take for there to be normative properties (cf. Bealer, 1998). I aim for my formulation to be less committed about how humans come to recognise the relevant facts. The analogous element to The Metaphysical

⁸² Because I treat ‘odd objectively authoritative properties’ as just a label in this thesis, it does not matter that realists do not agree. The Self-Undermining Objection is about as strong no matter how we fill in the required property for epistemic reasons to exist. The fact that the objection is not sensitive to whether epistemic reasons require little green monsters to exist or merely atoms in the void is a hint that something is wrong with it.

⁸³ Also see Streumer (2017).

Thesis is the clause ‘such properties don’t exist’. There is no important difference between this clause and my formulation.

epistemic error theory is a similar claim to error theory, but specifically about epistemic normative properties. The ‘epistemic normative properties’ we have labelled ‘epistemic reasons’. I focus on the epistemic case for largely idiosyncratic motivations; it is what I happen to be interested in. However this does affect the argument I will make. My position is closest (although not that close) to Christopher Cowie’s (2016). Cowie is primarily interested in moral error theory. Moral error theory is a similar claim to epistemic error theory, but specifically about moral reasons. Cowie is interested in the epistemic dependence between moral error theory and epistemic error theory, whether we should hold only one true, hold both true or hold neither true. Strictly speaking, he is agnostic about the self-undermining objection. He holds that either the self-undermining objection is false or moral error theory and epistemic error theory are not similar to each other (the former is true but the latter is false) (Cowie, 2016, pp. 126-127). However, I aim to show that the self-undermining objection is false, so I cannot adopt Cowie’s agnosticism.

6.2.2.3. Husi’s

Stan Husi analyses the self-undermining objection in regard to ‘radical meta-normative skepticism’ (2013).⁸⁴ Radical meta-normative skepticism denies that any standard has “objective authority or normative force” (Husi, 2013, p. 424). This is error theory. Again, epistemic error theory is a similar but more specific theory. It focuses on *epistemic* standards rather than *any* standard. As a reminder (see section 1.3.1), I note that talk of standards is useful because one can assert there is a standard without asserting there is any part of the world providing objective authority to conform to that standard. For example, we can invent the standard: always hop on one foot. You are not conforming to this standard. But there is no need for atonement. The hop-on-one-foot standard isn’t normative. We might say that there are no genuine reasons to conform to it. In a similar way, we can talk of epistemic standards. These are the standards that govern our epistemic practices. What exactly these standards are is controversial. Perhaps the least controversial: believe in proportion with your evidence. Calling this a ‘standard’ leaves open whether there are reasons to conform to it or not. We can characterise epistemic error theory as the claim that the epistemic standards, such as ‘believe in proportion with your evidence’, do not produce epistemic reasons.

⁸⁴ Husi uses the term ‘Reasons Skepticism’ in the title but not in the body of the paper.

6.2.2.4. Case's

Case defends the self-undermining objection in relation to error theory, too (2020). Case defines error theory by citing Streumer. But for Case, Mackie is not an error theorist (2020, p. 93). This is because Mackie leaves room for instrumental normativity. Mackie believes that there is a sense of *should*, a sense of 'reasons', that does not commit us to the existence of odd objectively authoritative properties. For example, if you want a bridge to support the weight of motor vehicles, you *should* not make it out of cooked pasta. We might say one has instrumental reasons not to build a bridge out of cooked pasta. This can be analysed in purely descriptive terms so the instantiation of odd objectively authoritative properties is not necessary for instrumental reasons to exist. Because Mackie holds this, he is not an error theorist (in Case's sense). Indeed, Case is open to the claim that no one is an error theorist in his sense (Case, 2020, p. 93).⁸⁵ Because he defines Error Theory in this way, it is not obvious he is engaging in this particular debate. Nonetheless I will address Case's view in section 6.5. And moving forward, an **error theorist (in Case's sense)** is someone who thinks that any talk of reasons, even supposedly instrumental reasons, commits one to the existence of odd objectively authoritative properties and no such properties are instantiated.

6.2.2.5. Nagel's

Thomas Nagel discusses a self-undermining objection in regard to 'Epistemological Skepticism' (1997, pp. 19-26). While this discussion has some similarities with the other discussions I have referenced—indeed Husi references Nagel's discussion as an example of the self-undermining objection (Husi, 2013, p. 425)—I take Nagel's discussion to be a slightly different topic to the one we are engaged with. Nagel is analysing what he calls 'Epistemological Skepticism', the type of radical doubt that René Descartes engaged in during the first couple of meditations (2017). Nagel holds that epistemological skepticism that is the product of an argument cannot be total; you cannot doubt everything (cf. Empiricus, 2000) because in order to make an argument you must accept some proposition. I do not intend to take a stance on this point in this thesis. But I want to point out the difference between epistemological skepticism and epistemic error theory, between this point and the self-undermining objection, because understanding the difference is key to responding to one of Cuneo's versions of the self-undermining objection.

Skeptical theories, like external world skepticism or induction skepticism, are claims about what a certain set of standards do or do not endorse. External world skepticism is a claim that our

⁸⁵ Although, I suspect Streumer is an error theorist in this sense (2017, pp. 111-114)

epistemic standards do not endorse the proposition ‘there is an external world’ over the proposition ‘there is no external world’. Similarly, induction skepticism claims that our epistemic standards do not support standard inductions over arbitrarily formed beliefs. epistemic error theory, on the other hand, can hold that our standards support ‘there is an external world’ and standard inductions over alternatives. epistemic error theorists just hold that our epistemic standards lack something, say normative force, something that we usually hold the standards have; the epistemic standards do not produce epistemic reasons. This is a claim about the metaphysics of the epistemic standards. So it is a different kind of claim to ‘Epistemological Skepticism’ (in Nagel’s sense). This is why I did not use the term ‘Skepticism’ when labelling the target theory. epistemic error theory is importantly distinct from other skeptical theses.

6.2.3. Summary of terms

There are a few terms to keep track of in this chapter. Here is a summary:

Epistemic error theory denies that epistemic reasons exist. We interpret this claim as having two parts. **The What-It-Takes Thesis**: for something to be an epistemic reason, that thing must have an odd objectively authoritative property. **The Metaphysical Thesis**: no odd objectively authoritative property is instantiated.

The aim of this formulation is to capture a target theory of interest, while remaining uncommitted on many controversies.

To relate epistemic error theory to other terms:

Epistemic reasons: the name for epistemically normative parts of the world. Unfortunately, some epistemic error theorists that draw a distinction between reasons that are genuinely normative and reasons that are not. I call the reasons that are not genuinely normative **non-authoritative reasons** (this is the same distinction as in section 3.4). You have non-authoritative epistemic reasons to believe just in case your belief conforms to the epistemic standards.⁸⁶

⁸⁶ Talk of non-authoritative reasons is awkward because ‘reasons’ is the term I am using to refer to the supposedly non-descriptive stuff. A problem with discussions about the metaphysics of normativity is that whatever term one uses to identify *normativity*, philosophers will want to make a distinction between types of normativity that are purely descriptive, say the *normativity* associated with manners, and the type of normativity that is genuinely normative, say the *normativity* of morality (Finlay, in press). The term ‘epistemic reasons’ has suffered the same fate (Côté-Bouchard & Littlejohn, 2018; Maguire & Woods, 2020; Cowie, 2019 cited in A. Murphy, 2021). I might have invented a new

Epistemic standards: refers to the norms or rules of epistemic practice but not to any objective authority that those norms or rules have. The term ‘epistemic standards’ is a way of referring to the norms or rules of epistemic practice without asserting that they produce epistemic reasons.

Error theory and **radical meta-normative skepticism** are like epistemic error theory but more general. They hold that all standards lack an odd objectively authoritative property. epistemic error theory only holds that epistemic standards lack that property. Error theory denies that there are any reasons whereas epistemic error theory only denies there are epistemic reasons.

Error theory (in Case’s sense) is the position that any talk of reasons, even instrumental reasons, commits one to odd objectively authoritative properties and no such properties are instantiated.

Epistemological skepticism (in Nagel’s sense) is a distinct theory. Epistemic error theory is a metaphysical claim about the nature of epistemic standards. Epistemic Skepticism (in Nagel’s sense) is a claim about what those standards do or do not support.

Importantly, this is just one way to keep the books. I intend nothing of substance to depend on these choices.

6.3. The Self-Undermining Objection

The self-undermining objection is the claim that in performing some action involving epistemic error theory, the epistemic error theorist is in some sense in tension with themselves. What the exact problem is varies between authors (cf. Case, 2020; Cuneo, 2007, chapter 4; Streumer, 2013). But they are all driving at the same point. So I will offer a general version of the objection first, and then I will relate it to specific versions found in the literature.

term, say *epistemic snorms*, but the new term would be less intuitively informative, and I suspect exactly the same thing will happen again. So I have just accepted this infelicity and reserved the term ‘epistemic reasons’ in this chapter for the genuinely epistemically normative part of the world (*normative* in the super-duper golden-glow sense...).

6.3.1. The general version

The general version of the self-undermining objection:

Self-Undermining Objection

P1) Epistemic error theory denies there are epistemic reasons.

P2) [Epistemic action] presupposes the existence of epistemic reasons.

C1) To be consistent, anyone holding epistemic error theory must accept the conjunction: *epistemic error theory and I don't [epistemic action] epistemic error theory.*

P3) Accepting the second conjunct undermines any warrant to accept the first conjunct.

C2) Epistemic error theory is self-undermining.

I will now explain these propositions.

P1 is just stating what epistemic error theory claims (see section 6.2).

P2 is a claim that when we perform some specific action, we somehow presuppose the existence of epistemic reasons. This is something like claiming that by praying, one presupposes a god exists. The action in question varies between authors. One version of P2 focuses on arguing. By arguing, an objector might claim, one is committed to the existence of epistemic reasons because successful arguing just is identifying epistemic reasons to believe something. But different versions may focus on different actions, such as believing or conforming to norms. '[Epistemic action]' is a placeholder for the specific action in question.⁸⁷

C1 is a claim that an epistemic error theorist must accept a certain conjunction to be consistent. Accepting this conjunction is, supposedly, problematic so let us call the conjunction from C1, **the problematic conjunction**. The first conjunct is just epistemic error theory, so they have to accept that to qualify as an epistemic error theorist. The second conjunct is a consequence of P1 and P2. The thought is that if the action in question presupposes the existence of epistemic reasons, then epistemic error theorists cannot consistently take themselves to be performing that action in relation to epistemic error theory.⁸⁸ For example, if arguing commits one to the existence of

⁸⁷ I am using the term 'action' very loosely. Sometimes the 'action' in question is 'having epistemic reasons' which is probably not an action. Indeed, you might hold that believing is not an action but more of a reflex. However, none of this matters for the argument and I cannot think of a better placeholder name.

⁸⁸ This assumes that epistemic error theorists must accept the implications of P1 and P2. Actual agents might not accept the implications and remain consistent by not contemplating the implications at all. But this kind of ignorance is fragile. Once the putative implications are pointed out to an error theorist, they are no longer ignorant. So I am happy to accept the assumption, which is helpful for my interlocutors, and offer a less fragile reply to the argument. For an argument for this assumption see: (Streumer, 2013, pp. 195-199).

epistemic reasons, one cannot consistently be an epistemic error theorist and hold that one argues for epistemic error theory. So an epistemic error theorist must accept that they do not argue for epistemic error theory. There is an analogous conjunct for any [epistemic action].

P3 is the claim that accepting the conjunction from C1 undermines the warrant one has to accept epistemic error theory. The conjunction from C1 is something like a Moorean Paradox sentence: p but I do not believe p . It might be that various Moorean Paradox-like sentences are acceptable and that philosophers might already accept them. For example, it might be fine to assert: ‘one plus one equals two and I have no argument for that’ because basic arithmetic is epistemically fundamental. And it might be fine to assert some potential conjunctions that involve [epistemic actions]. Indeed, at least on first pass, the reply that epistemic error theorists offer is to claim that the putative problematic conjunctions are not actually problematic to endorse.

Nonetheless, Cuneo explains the supposed problem with the conjunctions from C1:

... [the problematic conjunction appears] to be absurd because the person who accepts it is committed to accepting both its conjuncts. However, the acceptance of the proposition expressed in its second conjunct undercuts the justification or warrant we might have for accepting the proposition expressed in its first conjunct. And the type of justification or warrant in question seems to be epistemic in character—justification or warrant from the perspective of gaining truth. So, [epistemic error theorists]...appear committed to accepting a proposition, the mere acceptance of which guarantees that it is unjustified or unwarranted (2007, p. 118)⁸⁹ [my footnote].

So, P3 asserts that the specific conjunction identified in C1 undercuts the warrant for accepting epistemic error theory. At least on first pass, P3 is the controversial premise. However, as I argue in section 6.7, epistemic error theorists can be interpreted as rejecting P2 as well.

C2 is the ultimate conclusion of the argument. Because an epistemic error theorist is forced to accept the conjunction from C1, epistemic error theory is supposedly in tension with itself. So arguing, asserting or believing etc. Epistemic error theory undermines the warrant we have to argue, assert or believe etc. the theory in the first place. This is the general form of the self-undermining objection.

⁸⁹ In the original text, this last sentence is a conditional. I have not kept that form for ease of reading. I do not believe anything of substance has changed.

6.3.2. Cuneo's versions

As I understand the literature, Cuneo has offered the most influential version of the self-undermining objection. Cuneo holds that there are at least three undesirable implications that follow from arguing for epistemic error theory (2007, chapter 4). They all have a disjunctive form. The first disjunct is that we have epistemic reasons to believe epistemic error theory, which would be a straightforward contradiction. The second disjunct is something else Cuneo judges undesirable. So all three undesirable implications for epistemic error theory are in the form of a dilemma.

The first dilemma is that either we have epistemic reasons to believe epistemic error theory or we have no epistemic reasons to believe epistemic error theory (Cuneo, 2007, p. 118). Cuneo holds that either disjunct is unattractive but that the second is much less severe than the first. The first is straightforwardly inconsistent. The second has the flavour of a Moorean paradox. To relate this version of the self-undermining objection to our general form: the [epistemic action] we would substitute into P2 is 'having epistemic reasons to believe'; the problematic conjunction that we would substitute into C1 is *epistemic error theory and I have no epistemic reasons to believe epistemic error theory*.

Cuneo argues that holding 'there are no epistemic reasons' leads to further undesirable results. If one holds there are no epistemic reasons, then one also must hold that there are no epistemic reasons for any proposition. So, Cuneo claims, one must hold that no propositions can have epistemic merit or demerit (2007, p. 119). Cuneo calls this position 'Epistemological Skepticism' (2007, p. 119). Importantly, this notion of epistemological skepticism is different from the position Nagel was talking about. **Epistemological skepticism (in Cuneo's sense)** is compatible with denying epistemological skepticism (in Nagel sense) because the former is about the metaphysics of the epistemic standards and the latter is about what those standards endorse. The second dilemma is that either we have epistemic reasons to believe epistemic error theory or epistemological skepticism (in Cuneo's sense) is true. So the [epistemic action] we would substitute into P2 is again 'having epistemic reasons to believe'. The problematic conjunction to substitute into C1: *epistemic error theory and I have no epistemic reasons to believe, and therefore I accept epistemological skepticism (in Cuneo's sense)*.

Pushing further, Cuneo claims that without epistemic reasons there cannot be arguments. "A statement being offered as evidential support for a conclusion, however, is just a matter of its being

offered as a reason for accepting that conclusion. And, when all goes well, premises are reasons to accept a conclusion” (Cuneo, 2007, p. 121). So the third undesirable result is that either we have epistemic reasons to believe epistemic error theory or there can be no successful arguments for anything. So again the [epistemic action] we would substitute into P2 is ‘having epistemic reasons to believe’. The problematic conjunction: *epistemic error theory and I have no epistemic reasons to believe, and therefore, there are no successful arguments, even for epistemic error theory.*

Cuneo holds that having to accept either that we have epistemic reasons to believe epistemic error theory or one of the second disjuncts is a serious problem for a theory.

6.3.3. Other versions

Streumer’s version focuses on beliefs (2013). Using our terminology, we can summarise the position: to believe something in some sense presupposes that there are epistemic reasons to believe that thing. So you cannot really believe ‘there are no epistemic reasons’ because as soon as you contemplate that proposition you realise that there cannot be any epistemic reasons to believe it. The relevant [epistemic action]: believing. The problematic conjunction: *epistemic error theory and I don’t believe epistemic error theory.*

One of Husi’s versions of the self-undermining objection, which he attributes to David Enoch, focuses on norms and deliberation (Husi, 2013). Using our terminology, we can summarise this position: to engage in deliberation in some sense presupposes that you accept certain norms; to accept certain norms in some sense presupposes there are epistemic reasons; so denying there are epistemic reasons is inconsistent with engaging in deliberation. The relevant [epistemic action]: engaging in deliberation. The problematic conjunction: *epistemic error theory and I haven’t deliberated about epistemic error theory.*

Indeed, you can imagine the space of possible self-undermining objections. A potential objector just needs to hold that some action in some sense presupposes there are epistemic reasons. And then the objector can force the epistemic error theorist to accept a proposition that the epistemic error theorist himself does not perform that action. If the specific action is common, like believing or asserting, then the resulting proposition is going to sound counter intuitive. This is where the force of the objection comes from, making the epistemic error theorist accept some absurd-sounding proposition. Let us call holding that some action in some sense presupposes the existence of epistemic reasons, being an **epistemic normativist** about that action. There is a wide

range of actions that a potential objector might be an epistemic normativist about. So there is a wide range of potential self-undermining objections.

6.4. The Not-So-Bad Reply

Epistemic error theorists reply that the problematic conjunction from C1 is not actually problematic to accept—that is they reject P3. They argue for this by pointing out that there is a conjunction in the vicinity of the problematic conjunction that they can accept. That they can accept this second conjunction shows that it is not a problem to accept the supposedly problematic conjunction from C1.

The reply starts with a distinction. The distinction depends on the version of the Self-Undermining Argument that the epistemic error theorist is replying to. Streumer replies to his own argument with a distinction between believing that in some sense presupposes there are epistemic reasons and believing that does not, between fully, non-compulsively believing and other types of believing (Streumer, 2013). Husi distinguishes between deliberating in such a way that in some sense presupposes there are epistemic reasons and deliberating in such a way that does not (Husi, 2013). In our terminology, we can interpret Cowie as distinguishing between having epistemic reasons for a proposition and having evidence for a proposition (Cowie, 2016).

These responses amount to the same strategy. At least on first pass, the not-so-bad reply is that there are two senses of [epistemic action]: one sense that suggests that there are epistemic reasons and another sense that does not. Let us label the two senses [epistemic-reasons-suggesting action] and [non-epistemic-reasons-suggesting action]. The epistemic error theorist then claims that P3 is false because when properly understood, accepting the conjunction from C1 is not undermining. The epistemic error theorist does accept:

Epistemic error theory and I don't [epistemic-reasons-suggesting action] epistemic error theory.

But they also accept another conjunction:

Epistemic error theory and [non-epistemic-reasons-suggesting action] epistemic error theory.

Accepting the second conjunction recontextualises what it meant to accept the first conjunction, removing the seeming undermining quality that accepting the first conjunction had, at least for the

epistemic error theorist. So P3 is false. This is easier to understand when we substitute in some epistemic actions.

Cuneo had three conjunctions, each of which was, supposedly, self-undermining to accept. The first conjunction the epistemic error theorist must accept:

Epistemic error theory and there are no epistemic reasons to believe epistemic error theory.

In response, the epistemic error theorist concedes that they accept that conjunction. But they can also accept:

Epistemic error theory and there is lots of evidence for epistemic error theory (Cowie, 2016).

The fact they can consistently accept that there is a lot of evidence for epistemic error theory removes the undermining quality that accepting the first conjunction had, at least for the Epistemic Error theorist.

Cuneo's second conjunction, that the epistemic error theorist must accept:

Epistemic error theory and...I accept epistemological skepticism (in Cuneo's sense).

But they can also accept:

Epistemic error theory and I reject epistemological skepticism (in Nagel's sense).

Cuneo's third is:

Epistemic error theory and...there are no successful arguments for epistemic error theory.

But the epistemic error theorist can also accept:

Epistemic error theory and there is a lot of evidence that we can point to for epistemic error theory.

In each case, accepting the second of the pair of conjunctions removes the undermining quality that accepting the first seems to have had, at least for the epistemic error theorist. So in all cases, they deny P3.

We can characterise Streumer's response to his own version of the self-undermining objection (2013) in a similar way. The problematic conjunction is:

Epistemic error theory and I don't fully, non-compulsively believe epistemic error theory.

But he also accepts:

Epistemic error theory and I partially believe epistemic error theory.

Accepting the second removes the undermining quality of accepting the first.

We can characterise Husi's response to one version of the self-undermining objection, which he attributes to David Enoch (Husi, 2013). The problematic conjunction is:

Epistemic error theory and I haven't deliberated about whether epistemic error theory is true or not.

But he also can accept:

Epistemic error theory and the epistemic standards strongly support believing epistemic error theory.

And so on. I suspect the reader will be able to anticipate the types of distinctions that can be made and the types of couples of conjunctions the epistemic error theorist can accept. When the objector asserts that it is problematic to accept some conjunction, the epistemic error theorist will deny it is problematic, because they accept something in the neighbourhood that makes them realise there is nothing undermining about accepting the first conjunction.

Here is a way of thinking about the not-so-bad reply: suppose someone claims, 'time doesn't exist'. That claim seems absurd such that I might dismiss their position without hearing any more. But they continue, and secondly claim: "I think that for time to exist, all observers must agree on the order in which events occur; all observers must agree that event A happen before B, which

happened before C. According to our best understanding of the physics of our world, if matter is distributed in a certain way, not all observers will agree on the order of events. This possibility means that time doesn't exist. Of course, I believe there is a time-like part of spacetime that functions a lot like time. But 'time' properly so-called doesn't exist" (cf. Gödel, 1949). That second claim, the one with the explanation, changes my mind about the first claim. The second claim may be false, but it is not absurd. And that makes me realise that, understood in a very specific way, the first claim is not absurd. On first pass, one can think of the not-so-bad reply as doing something similar. It is attempting to convince you that accepting the supposedly problematic conjunction is not actually undermining because there is a conjunction in the vicinity that the epistemic error theorist can accept. They accept the analogue of *something like time exists*, even though they also accept the analogue of the intuitively absurd *time doesn't exist*.

6.5. Case's Retorts

Spencer Case rejects the not-so-bad reply in (2020). However, recall that error theory (in Case's sense) does not include Mackie's position because an error theorist (in Case's sense) must hold that *any* talk of reasons commits one to the existence of odd objectively authoritative properties. So it is consistent with Case's argument that the epistemic version of Mackie's position is not self-undermining, which is the position of interest for this chapter.

Nevertheless, Case offers two criticisms of the not-so-bad reply. While I do not think these criticisms are successful, they are illustrative of some common problems with the objector's position. So it will be useful to have them on the table. I only critique the first criticism in this subsection. I respond to the second criticism below.

6.5.1. There is no non-reasons-suggesting alternative

Case's first criticism of the not-so-bad reply is that there is no relevant alternative non-reasons-suggesting substitute for the [epistemic action]. Case understands the not-so-bad reply to be a distinction between reasons. We can call one side of the distinction 'normative reasons', the reasons that imply the existence of odd objectively authoritative properties and the other side 'non-authoritative reasons', reasons that do not imply the existence of odd objectively authoritative properties. The epistemic error theorist is saying it is not self-undermining to accept *there are no normative reasons to believe 'there are no epistemic reasons'* because they also accept that *there are non-*

authoritative reasons to believe 'there are no epistemic reasons.' This is the version of the not-so-bad reply that Case engages with.

Case maintains this reply cannot work:

This response relies on the assumption that it is possible to reject the authority of the epistemic domain while continuing to accept that there are identifiable *epistemic* reasons. ... That does seem to be true of the rules of games and the rules of grammar. We cannot credibly cite grammatical rules to criticise a poet who deliberately bends or breaks them for effect, or someone who is speaking a different language in which they do not apply. The same goes for legality. An anarchist can recognise that some actions are illegal without thinking that this must affect his practical deliberation. He is certainly not contradicting himself when he says that some actions are illegal but that we may disregard the law's recommendations.

Other norms seem to be defined in part by their claim to authority. Morality is a case in point. So, *contra* Husi, to reject the authority of moral norms is to reject the moral norms themselves.

The same is true, I think, of epistemic reasons, requirements, and permissions. The essence of epistemic norms is that they are genuinely authoritative over how we should think. They are not merely *epistemically* authoritative over what we ought to think, which would be trivial. Rather, they tell us how we Just Plain Ought to think—at least in ordinary circumstances (Case, 2020, pp. 96-97).

I will offer two problems with this criticism. The first problem with this criticism is that it is hard to make sense of the error theorist (in Case's sense) who offers the not-so-bad reply about reasons. Recall that the error theorist (in Case's sense) denies that there is any sense of normativity that does not commit us to the odd objectively authoritative properties that error theorists are so keen to deny. This is why Mackie does not count as an error theorist (in Case's sense). The not-so-bad reply about reasons wants to make a distinction between the reasons that commit us to the existence of odd objectively authoritative properties and reasons that are just like instrumental reasons, the reasons we have not to make a bridge out of cooked pasta, that do not imply any odd objectively authoritative properties. So, on first pass, the not-so-bad reply just looks like a denial of error theory (in Case's sense).

One way we can make sense of an error theorist (in Case's sense) who offers the not-so-bad reply about reasons, is for them to claim that up until *now*, all talk of reasons has committed us to the existence of objectively authoritative properties. But *now* they are introducing a completely new term:

I will now define a new term, ‘non-authoritative reasons’. You have non-authoritative reasons when some action would promote an end. For example, if you want a bridge to support motor vehicles, you have non-authoritative reasons to not build that bridge out of cooked pasta. Similarly, if you want to believe the truth, then you have many non-authoritative reasons to believe epistemic error theory (in Case’s sense).

A problem with this interpretation is that Case’s objection is beside the point. An error theorist (in Case’s sense) agrees that talk of reasons *has* committed one to the existence of objectively authoritative properties. But they can simply introduce a new term. Perhaps one could read Case’s first criticism as denying the possibility of defining a new term. But the error theorist (in Case’s sense) can disagree; maybe they are wrong, but the position is not self-undermining. Perhaps one could read Case’s first criticism as suggesting that the term ‘non-authoritative reasons’ does not refer. But the error theorist (in Case’s sense) can disagree; maybe they are wrong, but the position is not self-undermining. So the first problem with this criticism is that it is difficult to interpret an error theorist (in Case’s sense) offering the not-so-bad reply about reasons while interpreting Case’s first criticism as relevant and supporting the self-undermining objection. I cannot find an interpretation that satisfies all these constraints.

The second problem is that this criticism is very fragile, in the sense that one must view the world in a very specific way for the criticism to have purchase, so it hits relatively few targets (cf. section 6.7). Note that Mackie, who is not an error theorist (in Case’s sense) would not agree with Case. Mackie thinks there are instrumental reasons and these lack odd objectively authoritative properties. Indeed, many philosophers can be read as disagreeing with Case (Baker, 2018; Côté-Bouchard & Littlejohn, 2018; Foot, 1972; Maguire & Woods, 2020; Cowie, 2019 cited in A. Murphy, 2021). So it is not only error theorists (in Case’s sense) that want to identify things that are like epistemic reasons but that do not imply the existence of odd objectively authoritative properties. Indeed, Cowie (2016, pp. 20-21) and Stephen Finlay (in press) point out that this kind of distinction is extremely common in philosophy.⁹⁰ So, the central claim of this criticism, that all senses of ‘epistemic reasons’ imply the existence of objectively authoritative properties, is controversial. In general, there are few philosophical positions that are not self-undermining if one is allowed to cherry pick a controversial thesis to add to the position. So this criticism of the not-

⁹⁰ Regarding legal normativity, the same kind of distinction about legal reasons is debated by Legal Positivists and Natural Law Theorists (Fuller, 1958; Hart, 1958; Nadler, 2008). As we saw in the quote above, Case holds the Legal Positivist’s position, that the essence of legal reasons does not include objective authority (2020, pp. 96-97). Considering there are many similarities between the debates about legal reasons and about epistemic reasons, this different treatment needs more motivation.

so-bad reply does not compel the epistemic error theorist to concede their position is self-undermining; they can deny the controversial thesis about ‘epistemic reasons’.

6.5.2. The non-reasons implying alternative doesn’t explain the practice

The second criticism Case offers does not depend on the first:

Finally, even if we concede that epistemic reasons can be like this [not objectively authoritative] it is not clear that they could help to defuse the self-defeat objection.⁹¹

Compare the following:

Self-Defeat 1. Error theory is true, but there is no reason for anyone to believe that.

Self-Defeat 2. Error theory is true, but there is no reason—of a kind that anyone need take the least bit seriously, all things considered—for anyone to believe that.

The error theorist who takes the response that we are now considering would reject Self-Defeat 1 but would remain committed to Self-Defeat 2. ... I think that if Self-Defeat 1 is problematic or paradoxical, then Self-Defeat 2 is, as well. The error theorist’s having a reason to accept error theory only resolves the paradox if that reason is in some way binding on what he ought to think (Case, 2020, p. 99).

There is a way of interpreting this criticism as just another version of the self-undermining objection that is open to the not-so-bad reply. The error theorist might accept *Error theory is true, but there is no reason—of a kind that anyone need take the least bit seriously, all things considered—for anyone to believe that*. But they also accept *error theory is true, and if you want to believe the truth there are overwhelming non-authoritative reasons that indicate it is true*.

But I don’t think this is what the second criticism amounts to. Case makes this comment (2020, p. 98):

⁹¹ The ‘self-defeat objection’ is Case’s name for the Self-Undermining Objection.

The issue was never that of explaining how belief in error theory is a psychological possibility, however. It is, rather, to explain how error theory can *seem reasonable from the first-person perspective*. [My emphasis]

Husi, while disagreeing with Case about the answer, agrees that something like this is the important question (2013, p. 445):

Whether or not deliberation can be made sense of without objectively normative reasons constitutes the philosophical test case for what meta-normative theory to accept.

So I interpret the second criticism as not offering the self-undermining objection again but rather pointing out that denying the existence of epistemic reasons removes any objective authority to engage in epistemic work. It is like praying to God when you do not believe there is one—arguing (etc.) just looks pointless once you deny there are epistemic reasons. The challenge for the epistemic error theorist is to explain their project rather than to merely assert that their claims are logically possible.⁹²

6.6. Denying Spirits Exist

This is how I understand the current state of the dialectic. The epistemic error theorist claims that there are no epistemic reasons. The Self-Undermining Objector notes that the epistemic error theorist must accept the conjunction *epistemic error theory and there are no epistemic reasons to believe epistemic error theory*; the objector claims this conjunction is self-undermining. The epistemic error theorist denies it is self-undermining; they point to another conjunction, *epistemic error theory and there are many non-authoritative reasons to believe it*, to show that the first conjunction is not self-undermining, properly understood. Case responds that ‘non-authoritative reasons’ do not make it reasonable to engage in epistemic practice, so it is not reasonable to argue for epistemic error theory.

My response to this dialectic is to shift from a debate about specific terms and specific conjunctions to a model of what the epistemic error theorist thinks is happening. The model is an analogy to a situation where it is clear that the position in question is not self-undermining despite the fact that there are many absurd-sounding commitments.

⁹² Indeed, I suspect this thought is behind Bernard Williams’ dismissal of a purely instrumental conception of the value of sincerity. Such a conception, supposedly, lacks a proper human self-understanding (whatever that may mean) (Williams, 2002, pp. 57-62). But this is only a suspicion; I find the passage rather opaque, despite many re-readings.

Suppose you live in a world where people have no term for belief. Rather they use the term ‘Spirit gifts’. ‘Spirit gifts’ is defined thus: a representation of the world given to us by the spirits used to guide behaviour. Everyone agrees this is what the term ‘spirit gifts’ means and this is the definition found in dictionaries, etc. They use the term ‘spirit gifts’ in much the same way we use the term ‘beliefs’. To express that they think a certain sentence is true, they say something like *I spirit gift that sentence*. Their epistemic practice is permeated with talk of spirits. For example, they say the point of arguing is to *test which is the strongest spirit*. When they assert a position, they say *The Spirit told me that...* They are in the grasp of a spirit-paradigm.

Now suppose a spirit error theorist asserts ‘there are no spirits’. A philosopher in this world may very well say: “that is self-undermining because you must endorse the Moorean-like proposition *there are no spirits and I don’t spirit-gift ‘there are no spirits’*. To people in this spirit community, this will sound very odd, perhaps even absurd. Additionally, a philosopher in this community may even say: “what is the point of arguing if you don’t think there are Spirits? To argue is just to find out the strongest Spirit?” Arguing will just look pointless to them, if one denies there are spirits. So to many people in the Spirit community, this position may sound absurd and self-undermining.

But from the outside we can see that the spirit error theorist is not undermining their position. While they don’t spirit-gift ‘there are no spirits’, they do believe it. And using our conceptual machinery, it does not even sound odd to us to deny the existence of spirits. The spirit error theorist may even put forward our conceptual machinery to argue for their position. The spirit error theorist might say “There is a non-spirit-suggesting alternative for spirit gifts: ‘beliefs’. These things do not suggest that spirits exist but play a similar role otherwise. I believe that ‘spirits don’t exist’ even though I don’t spirit gift that ‘spirits don’t exist’.” Additionally, they might say “I am not interested in what is the strongest spirit gift because I don’t believe there are any. But I am interested in what is true. I don’t believe spirits help us find the truth. But I think we find it out in some other way. And that is what I am doing.” The spirit error theorist’s position is not self-undermining. But note that for all I have said, the spirit error theorist may be wrong. In this community, spirits might exist, and they might be intimately involved with epistemic practice. But that is some fact the needs to be discovered.

As with the spirit error theorist, so too with the epistemic error theorist. To be consistent, they will have to accept *there are no epistemic reasons and there are no epistemic reasons to believe ‘there are no epistemic reasons’*. This may sound odd, even absurd to English-speaking philosophers. But epistemic

error theorists are claiming that we are in a situation like the community accustomed to spirit-gift talk. Epistemic error theorists can attempt to introduce new conceptual machinery to understand the position, just like the spirit error theorist did. Maybe there are no epistemic reasons but there are things like them. Maybe epistemic error theorists cannot *believe* the theory but they can do *something like believe* the theory.

Here is an example story the epistemic error theorist can tell:

The Error Theorist's Meta-Story

When I do a thing that looks like an [epistemic action], I am doing something else, something that I don't think requires epistemic reasons to be successful. When it looks like I am arguing (granting that 'arguing' requires the identification of epistemic reasons) I am doing something else: identifying evidence that guides us to understand the way the world is. Of course, we may fall prey to the same dialectic: evidence itself in some sense presupposes epistemic reasons. I am happy to grant the objector all the terms they want. So I am happy to concede that 'evidence' in some sense presupposes that epistemic reasons exist. I am doing something else then. Let's call it 'identifying signs', looking at parts of the world and thinking about what they tell us about other parts of the world. I don't 'believe' (or Ersatz *believe*) that I need to identify epistemic reasons to successfully do the thing I am doing. I believe (or Ersatz *believe*) I can track the truth. I can't track it perfectly, of course. I may believe (or Ersatz *believe*) many falsehoods, near truths along with radical inaccuracies. But importantly, I can track the truth even though there are no odd objectively authoritative properties. That is what I am trying to do.

And this is only one understanding of an enterprise that the epistemic error theorist can engage in that does not require epistemic reasons to exist. Perhaps they think they are just trying to make their beliefs overall coherent. Perhaps they think they are just trying to make their beliefs empirically adequate. Perhaps they think they are just playing some game, like chess, but more complicated and less fun. All they need to think is that there are enterprises that do not require 'epistemic reasons' to exist.

A way to think about this model: suppose you come across an atheist praying. You then accuse them of self-undermining their atheism because, by praying, they are suggesting there is a god. They reply that they are just praying because it makes them feel relaxed—taking a moment to think

about what they are grateful for, their hopes and their fears. Replying that the essence of ‘praying’ is communicating with a God is not relevant. They can simply concede they were never really ‘praying’ in the first place. They were simply doing something like praying. Ditto the epistemic error theorist when an objector accuses them of self-undermining by accepting, asserting, arguing, believing, deliberating, holding, etc. They do not think of themselves as doing those things in such a way that commits them to the existence of epistemic reasons. If that means they are not *really* accepting, asserting, arguing, believing, deliberating, holding, etc. so be it. They are doing something else.

6.7. A Deep Flaw in the Self-Undermining Objection

Recall that we defined an epistemic normativist in this chapter about some action as someone who holds that performing that action in some sense presupposes there are epistemic reasons. So one is an epistemic normativist about ‘successfully arguing’ if one holds that the existence of successful arguments implies the existence of epistemic reasons. If one is a realist about ‘successful arguments’, one holds that sometimes people successfully argue. Being an epistemic normativist and a realist about an action, like successfully arguing, is in tension with being an epistemic error theorist. I concede that one cannot consistently accept that people successfully argue and that successfully arguing requires epistemic reasons to exist and that there are no epistemic reasons. But a deep problem with the self-undermining objection is that it *requires* the epistemic error theorist to be an epistemic normativist and a realist about the relevant action in question, things they are not compelled to be.

Being an epistemic normativist about any action is extremely controversial. To get a sense of how controversial,⁹³ note that most metaphysical accounts of *moral* normativity are consistent, if not outright agree, with The Metaphysical Thesis; most positions are consistent with there being no odd objectively authoritative properties or agree that there are no such properties. Any naturalistic reduction agrees. Any non-cognitivist account agrees. Arguably, any constitutive account agrees. Along with moral error theorists, this accounts for about 70% of philosophers according to the latest Philpapers survey by Bourget and Chalmers (forthcoming, p. 16).⁹⁴ The other about 30% of philosophers are non-naturalists. This is the upper limit on philosophers who outright disagree

⁹³ A better method would be to directly survey philosophers. I would bet that the majority of philosophers hold that believing, etc., does not imply the existence of the odd objectively authoritative properties that Mackie was keen to deny. Nonetheless, it would be best to actually find out.

⁹⁴ About 70% of philosophers who felt sufficiently familiar with the topic to answer the question and could only pick one answer to accept or lean towards.

with The Metaphysical Thesis. But, as I have argued in section 5.2.2, some extremely influential non-naturalist positions are also consistent with The Metaphysical Thesis. Scanlon (2014) and Parfit (2011) as interpreted by Veluwenkamp (2017), and Shafer-Landau (2003, p. 76) are all consistent. It is only ontologically committed non-natural realists like Enoch (2011), Huemer (2005) and arguably G. E. Moore (1903) that deny the moral equivalent of The Metaphysical Thesis. So while 30% is the upper limit, the percentage of philosophers who outright disagree with The Metaphysical Thesis might be much lower. We can infer from this that it is *extremely* controversial that something like arguing or believing, even arguing successfully or believing successfully, would in some sense presuppose The Metaphysical Thesis is false because most philosophers do not even think the existence of morality implies The Metaphysical Thesis is false. So the self-undermining objection is built upon an extremely controversial foundation.

But the situation is worse than merely having a criticism that requires one to accept a controversial premise. The self-undermining objection requires that the epistemic error theorist themselves hold the controversial premise, otherwise the objection cannot establish the charge of *self*-undermining. The objection requires that the epistemic error theorist is an epistemic normativist about the action in question, say about arguing, and that they are a realist about the action in question, that many people actually argue. But the epistemic error theorist does not need to agree. They can either be a non-epistemic normativist about the action (rejecting P2 from the general argument) or they can be a non-realist about the action (rejecting P3 from the general argument). Indeed, this is a useful way of interpreting the not-so-bad reply: the epistemic error theorist is either claiming that the action in question, arguing successfully say, does not require the existence of epistemic reasons or that no one *successfully argues* exactly, rather they do something functionally similar to *successfully argue*, so it is not undermining to claim they do not *successfully argue* for epistemic error theory. And this is a situation where we can infer that most philosophers agree with them; most philosophers probably do not think successfully arguing does commit one to the existence of odd objectively authoritative properties because most philosophers do not even think moral normativity commits one to the existence of odd objectively authoritative properties.

A deep problem with the self-undermining objection is that it requires the epistemic error theorist to have some pretty controversial epistemic normativist commitments, commitments that they are unlikely to have, indeed commitments that the majority of philosophers are unlikely to have. So while epistemic error theory may be false, it is not self-undermining.

7. Normative Metaphysics is Unsatisfying

This thesis summarised: epistemic normativity is the robust and potent usefulness of conforming to the epistemic standards. Accurate representations are broadly useful. Humans have developed to take advantage of that fact. They take advantage of that fact by conforming to the epistemic standards. They have developed psychologies and societies that promote conforming to the epistemic standards. This creates a complex web of relationships between conforming to the epistemic standards and the promotion of various ends. The robust and pervasive usefulness of conforming is a fact about this web of relationships, that conforming in fact promotes many ends in many ways. Epistemic normativity is usefulness.

I hold that there is a lot of evidence that this theory is true or close to the truth. Indeed, I take it that the usefulness of conforming is uncontroversial. And I take it that this usefulness influenced the development of humans is also uncontroversial (even if we lack evidence about the specifics of our evolutionary history). I note that the account does not depend on any particular approach to epistemology. The account explains the way many phenomena interact with epistemic standards, not just humans and not just beliefs. And I note there is a ready-made recipe for responding to intuition-based criticisms, removing the evidential force of those criticisms. The result is that the evidence strongly tells in favour of the truth or approximate truth of the theory.

I hold that the semantics gives us very little evidence about the ontology of epistemic normativity. An interlocutor may hold that ‘epistemic normativity’ cannot possibly refer to usefulness. I am open to the possibility that the English term ‘epistemic normativity’ does not refer to usefulness. Nevertheless, the robust and potent usefulness of conforming to the epistemic standards is what has shaped our development and is what causes us to make certain judgements about what we should believe. So the usefulness is the thing that we are tracking with our talk of ‘epistemic normativity’, even if we are tracking it imperfectly. However, I am also open to the possibility that ‘epistemic normativity’ does refer to such usefulness and the possibility that the term is so vague that it does not refer to much at all (see section 1.3.2 and 5.5).

7.1. A Difference Intuition

I imagine some philosophers will agree with almost everything I have said but remain unsatisfied. I imagine them claiming that I have failed to identify any *normativity* in my theory of normativity.

Objections like this are often leveled against natural realist theories. Matthew Lutz and James Lenman (2021) describe such an objection:

Although Moore's original version of the Open Question Argument today has few defenders, there have been a number of recent attempts to refashion it into a more compelling form. One popular version of the Open Question Argument, called the Normativity Objection (Scanlon 2014; Parfit 2011), sidesteps questions about the cognitive significance of moral and descriptive terminology and appeals to considerations regarding the natures of natural and normative facts. Moral facts tell us what is *good* in the world and what we have *reasons* or *obligations* to do. Natural facts—the kinds of facts that scientists study—are facts about the innate physical structure of the universe and the causal principles that govern the interaction of matter. Those are obviously just two different kinds of facts. Moral facts, because they are facts about goodness, reasons, obligations, and the like, are *normative* facts. But natural facts are not normative. In trying to give a naturalistic account of morality, naturalists forgot the most important thing: that moral facts aren't purely facts about the way the world is; they are facts about *what matters*. (Mathew Lutz & Lenman, 2021, section 2.2)

While this objection is about moral facts, it can be applied to epistemically normative facts. One can imagine an interlocutor claiming that facts about usefulness are just too different to facts about epistemic normativity for my account to offer a satisfying metaphysics.

As the quote suggests, this objection is related to the open question argument (Moore, 1903, para [1]-[14]). I also note that David Copp (2011) discusses five versions of this objection, all of which are subtly different to Lutz and Lenman's characterization. And we might even add moral twin-earth arguments to the list of related arguments (Horgan & Timmons, 1991, 1992b). What unifies these arguments is that they all leverage an intuition that normative facts are different to natural facts to argue against natural realism. I call an intuition that normative facts are of a different kind to some other set of facts **a difference intuition**.

In this chapter, I am not interested in the details of the objections that leverage a difference intuition because I have already implicitly responded to these objections. I take the master response from chapter 4 to address the evidential value of these objections. In addition, my openness to error theory discussed in chapter 5 also addresses these objections; it is consistent with the error theory interpretation of embedded epistemic instrumentalism that 'normativity' does not reduce to natural facts but that nonetheless the only relevant facts are natural facts and we are tracking those natural facts with our normative concepts. So in this chapter I am not concerned with defending my theory from these objections.

Rather, I want to explain myself. I want to explain myself because when I present material from this thesis—after Q and A, at the bar or at dinner—this is the type of objection philosophers are most keen to discuss. I suspect a difference intuition motivates a lot of philosophers to reject accounts like mine. In this chapter, I want to explain why I am not motivated to offer anything but a natural realist or error theory account, even in the face of a difference intuition. In this chapter I argue that at least for me, a difference intuition is not a difference-maker.

7.2. There is No Satisfying Metaphysics of Normativity (for me)

I experience a difference intuition about all accounts of the metaphysics of normativity (that I have come across). The mere fact I feel as if normativity is completely different from natural facts (at times: see section 7.2.3) does not distinguish natural realism from any other metaphysics, for me. No metaphysics of normativity feels satisfying, *feels* as if it has captured normativity, for me.

To help convey my difference intuition, perhaps even provoke a similar intuition in you, let me characterise some metaphysical accounts of normativity and offer my reactions to them. Afterwards I will offer a conjecture.

7.2.1. Semantic, conceptual or constitutive-driven metaphysical accounts

Facts about normativity do not feel like facts about semantics, about concepts or about constitutive relationships, for me. Facts about semantics, concepts, and constitutive relationships feel too arbitrary.

I suspect others feel as I do. Hilary Kornblith (1993, section 1) criticises Alvin Goldman's work on justification. The criticism is that in some sense a focus on the English term 'justification' is an arbitrary choice. Kornblith's observation is one specific instance of how an approach to the metaphysics of epistemic normativity may feel too arbitrary. I will first explain Kornblith's worry. Then I will generalise it to many other accounts. While I do not know the proportion of accounts that this worry applies to, I suspect it is most of them. I conclude that any semantic-driven, concept-driven or constitutive-driven metaphysics will feel as different from normative facts as the natural facts do, for me.

Kornblith notes:

In what Goldman conceives of as the foundational part of epistemology, we engage in autonomous inquiry into the meaning of various epistemic terms. We do this by way of testing proposed analyses against our intuitions, and we thereby attempt to capture our ordinary concept of, for example, justification or knowledge.... To be justified is simply to be something good; this is a matter of the meaning of the term. Similarly, and to a first approximation, for a belief to be justified just is for it to be the product of a truth-conducive process; this, too, is a matter of the meaning of the term (Kornblith, 1993, pp. 360-361).

Kornblith claims that this type of answer is unsatisfying because we might agree about the semantic facts but nevertheless hold that they are not important:

Now there is something terribly unsatisfying about this. Imagine someone, such as Stephen Stich, who claims that conduciveness to truth is of no value at all; that a belief which is the product of a truth conducive process is not, in virtue of that very fact, a good belief to have. On Stich's view, it is not merely that truth by itself does not make a proposition worthy of belief; rather, the truth of a belief, or the truth-conduciveness of the process which produced it, does not even count in its favor. Now if Stich were to confront Goldman and ask why he should favor true beliefs, or beliefs produced by truth conducive processes, it seems that Goldman should offer him the semantic argument above. It also seems that this should not convince Stich at all.

Indeed, this is exactly what Stich argues. Stich is concerned that even if Goldman has given the proper conceptual analysis of our epistemic terms, the fact that he has uncovered the proper conceptual analysis carries no normative force. In particular, as Stich points out, suppose there were another culture whose epistemic terms embodied a different set of standards. Suppose that in this other culture beliefs meeting different conditions were approved of. What reason is there to approve of beliefs meeting our conditions rather than those meeting the conditions of this other culture? It will hardly do to point out that this is what our terms mean (Kornblith, 1993, pp. 361-363).

We can paraphrase Kornblith's argument against a semantic-driven investigation into the metaphysics of epistemic normativity. Suppose as a matter of semantics, the term 'epistemic standards' implies that you have 'epistemic reasons' to conform to the epistemic standards. That is just a semantic fact about the term. Now consider two new terms: 'sh-epistemic standards' and 'sh-epistemic reasons'. The new terms have similar semantics to the old terms. The new terms, like the old, offer prescriptions about what you ought to believe. But the new terms offer competing prescriptions (the sh-epistemic standards tell you to believe against the evidence, for example). We can ask a question: why should our beliefs be guided by the old terms of 'epistemic standards' and 'epistemic reasons' as opposed to the new terms of 'sh-epistemic standards' and 'sh-epistemic reasons'? The semantic-driven philosopher might reply that the old terms and the new terms are not on a par, because there is some further semantic fact, say that the term 'epistemic reasons' has a certain semantic relationship to 'ought' which 'sh-epistemic reasons' does not have. But again,

why should we be guided by what we *ought* to do than what we *sh-ought* to do? Schematically, any semantic answer of the form: <the term '*f*' has a certain semantic relationship to the term '*q*'> is open to the challenge, why care about *f* rather than *sh-f* (or for that matter, why care about *q* rather than *sh-q*)? Only referring to semantic facts will never answer the challenge. Only referring to semantic facts forces us to arbitrarily choose to be interested on one set of semantic relations as opposed to another set.

A merely semantic answer misses the point of a metaphysical enquiry, for me. I am not trying to solve a mystery about the meanings of particular words. Rather, I am trying to solve a mystery about particular features of the world. I want to understand the properties of the epistemic standards that urge us to believe in a certain way, that make it better to believe in a certain way, that make certain beliefs right and others wrong. An answer that is only about the meanings of particular words will never be a satisfying answer for me. A semantics-driven investigation may tell us something interesting about the semantics of particular terms. However this approach does not have any tools to identify why those specific terms are important as opposed to other terms. A focus on specific terms (as opposed to other terms) is arbitrary unless we appeal to non-semantic facts. So, for me, an account of normativity that is driven by the semantics of specific terms will provoke a difference-intuition. Normative facts feel different to semantic facts. One way I can characterise this difference: semantic facts feel arbitrary in a way normative facts do not, for me.

Extending Kornblith's point, this arbitrary choice to focus on a specific term (as opposed to other terms) has analogues in many other theories. Concept-driven accounts face a similar problem.⁹⁵ Christos Kyriacou extends one such account, the moral fixed-points framework (Cuneo & Shafer-Landau, 2014), to epistemic normativity (Kyriacou, 2018). The central idea is that epistemic normativity is to be explained by a series of necessary relations between concepts. For example, according to this approach it is a necessary truth that: "It is pro tanto epistemically rational to believe what is best epistemically justified, given relevant evidence at one's possession" (Kyriacou, 2018, p. 78).⁹⁶ Supposedly, this is a necessary truth given by the essences of the concepts 'rational',

⁹⁵ You might think that a semantic-driven account and a concept-driven account are really the same. Some defenders of concept-driven accounts hold that they are different from semantic-driven accounts (Cuneo & Shafer-Landau, 2014, pp. 410-411; Kyriacou, 2018, n. 2). I am agnostic on this issue.

⁹⁶ As an aside, I doubt that this is a necessary truth. Suppose you roll a six-sided die. You know the die has a slight bias towards 4. Once the die has landed but before you know the result, should you believe 'the die landed on 4'? As far as I can tell nothing about the concepts of 'belief', 'rational' or 'best epistemically justified' rules out the possibility that you ought not believe 'the dice landed on 4' even though it is the best epistemically justified proposition according to your evidence. Indeed, we can vary the case by varying how many sides the dice has and how slight the bias is. I would have thought given enough sides and a small enough bias, it is possible that 'the dice landed on 4' is the best epistemically justified of the options, but not justified enough that the belief is rational. That is, I would have thought

‘justified’ and ‘evidence’. But, as we did above, we can invent a rival conceptual scheme involving ‘sh-rational’, ‘sh-justified’ and ‘sh-evidence’. Why should we care about the specific concepts like ‘rational’ rather than other concepts? Kyriacou responds to such a problem by expanding the conceptual scheme (Kyriacou, 2019); for him, it is necessarily true that concepts like ‘rational’ offer ‘normative guidance’. But again, we can expand our rival conceptual scheme; it is also necessarily true that sh-rational concepts offer sh-normative guidance. Why care about normative guidance as opposed to other kinds of guidance? Concept-driven accounts face a similar problem to semantic-driven accounts.

Constitutive accounts also face a similar problem. Constitutive accounts explain epistemic normativity as being constitutive of what it is to believe or what it is to be a part of an epistemic community (Engel, 2013; Greenberg, 2020; Shah, 2003; Shah & Velleman, 2005; Swindlehurst, 2020; Wedgwood, 2002, 2013; Whiting, 2010). For example, you might hold that for something to be a belief it must have the property that it is good when it conforms with the evidence. The classic problem with constitutive accounts, first offered against a constitutive account of agents, is known as the *shmagency* objection (Enoch, 2006). Suppose you hold that it is constitutive of reasonable agents that they conform to the epistemic standards and that is what explains epistemic normativity. In a move that should feel extremely familiar, we can also hold that it is constitutive of sh-reasonable agents that they flout the epistemic standards. Then we must ask why we should be a reasonable agent as opposed to a sh-reasonable agent?

While I do not do them justice in these comments, I know of two types of responses to this objection. One is to claim that it is inescapable to be an agent (for epistemic normativity, we might say it is inescapable to have beliefs) (Ferrero, 2009). A problem with this response is that we can also define sh-agency or sh-beliefs in such a way that makes them inescapable, too.⁹⁷ I can stipulate the concept of a ‘sh-agent’:

that *it is possible* that believing the best epistemically justified proposition is sometimes irrational when no relevant proposition is justified to any great extent. Of course, perhaps my understanding of the concepts is deficient in some way. Some empirical evidence about the way people actually use these concepts could help settle matters. But we can set aside this issue for the purposes of this chapter.

⁹⁷ On this point, I was inspired by how some misogynists talk about ‘beta males’. As I understand it, part of the concept of ‘beta male’ is that a beta male is a bad male. Further, it is inescapable whether you are a beta male or not. I can imagine a clever misogynist generating a whole conceptual framework that relates being a good male to not being a beta male and that whether you are a beta male or not is inescapable. Nevertheless, despite its inescapability, I remain hostile to using such a framework to guide my behaviour. The fact that my behaviours may inescapably be assigned positive or negative valence by a conceptual framework does not compel me to modify my behaviours. And it certainly does not imply considerations of being a beta male are genuinely normative.

A) All humans are sh-agents.

B) A good sh-agent must conform to sh-epistemic standards.

A and B are just a matter of the concept of 'sh-agent'. This is just what the concept is. Yet the concept is not made any more important by defining it in such a way to make it inescapable that you are a sh-agent. So inescapability is not rare, nor does inescapability suddenly make facts about constitutive relations feel like normative facts, for me. It still feels arbitrary to focus on some constitutive relations as opposed to others, for me.

The other response is to deny that constitutive accounts need to have an answer to this question (Paakkunainen, 2018); note a similar response is available to the analogous problems with semantic-driven and concept-driven accounts. But for me, the issue remains. Granted if one cares about the semantic or conceptual or constitutive relations between things, then perhaps one does not need to provide an answer about why we should focus on X rather than sh-X. But for me, the very question I am asking is 'what is it about the epistemic standards that make them so significant.' The very question I am asking is the very question these accounts cannot offer an answer to.

If philosophers are interested in the metaphysics of epistemic normativity, about how the world is constructed such that humans tend to aim to believe in proportion with their evidence, that humans criticise other humans for not believing in proportion with their evidence, that our psychologies and societies are permeated with mechanisms to reenforce the significance of believing in proportion with our evidence, then why X as opposed to sh-X is an extremely important question to answer for semantic-driven, concept-driven or constitutive-driven accounts of epistemic normativity. And it is a question that these accounts lack the resources to give a satisfying answer to, because they are confined to semantic, conceptual, or constitutive properties.

Semantic-driven, concept-driven or constitutive-driven metaphysics are in some sense arbitrary. Even in a world where such metaphysics is instantiated, this arbitrariness prevents such a metaphysics from being a satisfactory account of normativity. To me, normative facts are not facts about 'normative'. Even if it were true that the mere concepts necessitate that *'the earth is flat' ought not to be believed*, that feels like playing with words, to me. A semantic-driven, concept-driven or constitutive-driven theory, however it is developed, simply does not identify a metaphysics that satisfies me.

7.2.2. Metaphysically extravagant theories

Perhaps you agree that semantic-driven, conceptual-driven and constitutive-driven accounts also give rise to a difference intuition. But, you might hold, that is because they are metaphysically light. We need a metaphysically-loaded theory. There are two types of metaphysically-loaded theories. There are natural and non-natural theories. We are granting that natural theories elicit a difference-intuition. So we must look to non-natural realism to provide a metaphysics that does not elicit a difference intuition (Enoch, 2011; Huemer, 2005; Moore, 1903). Ontologically committed non-natural realism holds that normativity is a fundamental part of reality, ontologically extra from non-normative properties. And by postulating a fundamental normative part of reality, so someone pushing this line of thinking holds, we can secure a metaphysics that does not elicit a difference-intuition.

Not so, for me. Suppose there is some mysterious property that floats on top of the causal mess that is the world. Normative facts do not feel like facts about that property, for me. Indeed, I suspect many philosophers feel as I do. Here is how Frank Jackson puts this point:

[I]t is hard to see how the further properties could be of any ethical significance. Are we supposed to take seriously someone who says, “I see that this action will kill many and save no-one, but that is not enough to justify my not doing it; what really matters is that the action has an extra property that only ethical terms are suited to pick out”? In short, the extra properties would be ethical “idlers.” (Jackson, 1998, p. 127)

Here I believe Jackson is driving at a similar idea. The existence of non-natural properties, no matter how robust, ontologically committed or non-committed, just does not make a difference to what is *good*. For me, a world is as normative with them as it is without them. Normative facts feel different to facts about non-natural properties.

In response, Richard Yetter Chappell claims:

... this objection assumes that in order to be normatively significant, purely normative properties must themselves feature in (correct) justifications. Otherwise, Jackson claims, they are “ethical idlers.” But this is too quick. It neglects the possibility that their role may instead be to serve as truthmakers for (or as reflections of) the normative truths about which natural properties are normatively significant in which ways. One such normative truth is that the natural property of *killing many and saving none* is to be avoided—in contrast, say, to the natural property of *saving many and harming none*.

Those two natural properties differ drastically in normative valence. This normative difference is reflected, metaphysically, in their having different purely normative

properties: One is to be avoided, the other is to be pursued. Insofar as we consider these normative properties to be robustly objective properties, and irreducible to any natural properties, we thus find an important theoretical role for non-natural, purely normative properties. Nevertheless, it is not these normative properties themselves that are the things to be avoided or pursued (Chappell, 2019, pp. 130-131).

Chappell is positing that the non-natural normative truths are like signs that point us to the natural facts that are normatively valenced. However, these non-natural facts do not feel like normativity, for me. The problem, for me, is that a sign does not make something more than it is. A city is not made a city by putting up a sign. And it does not stop being a city if you take down a sign. A sign points you towards a city; it does not make it a city. Similarly, non-natural facts cannot make natural facts any more normatively significant than they are by themselves, for me. Rolling a rock up a hill does not become more normatively significant if there is a large, ghost-like arrow, pointing at the activity, declaring that it is normatively significant.

Non-natural facts are still just some further facts, like the fact that a term in a language implies that some natural facts are 'good'. To understand why I treat non-natural facts like this, consider divine command theory. Divine command theory holds that what god wills is what ought to be done (M. Murphy, 2019).⁹⁸ For me, normative facts do not feel like facts about the will of a God. To explain this, I offer the Euthyphro dilemma to characterize why the will of a god does not feel right. Is something good because a god wills it or does a god will it because it is good? If whatever a god wills is good, then the good is just the will of a powerful being. As I do not find facts about the commands of a powerful king or alien to feel similar to normative facts, I do not find facts about the will of an all-powerful being to feel similar to normative facts. Alternatively, if a god wills something because that thing is good, then a god's will does not make a thing good and I continue my search for facts that do not feel different to normative facts.

To return to non-natural facts, a world that comes with the normative valence of natural facts necessarily inscribed in the non-natural facts is similar to a world that comes with a god's will. We can put a Euthyphro-style dilemma to the non-natural realist. Is some natural fact normative just because there is a non-natural fact that indicates it as such or does the non-natural fact indicate it as such because the natural fact is normative? If a natural fact is normative just because there is a non-natural fact that indicates it as such, then normativity is just a matter of what is inscribed in the non-natural facts. But that is just like a god telling you what to do. It is just some part of the

⁹⁸ Divine command theory is also known as 'theological voluntarism'.

world telling you that some natural fact is normatively significant. Why ought we listen to the non-natural facts? For me, normative facts feel different to facts about such non-natural properties. The non-natural properties feel irrelevant. Of course, if the non-natural facts indicate a natural fact as normative because the natural fact is normative itself, then the non-natural facts are otiose, just like a god's will on the second horn of the Euthyphro dilemma.

Another way of characterising my difference intuition in response to non-natural metaphysics is by invoking an open question. Moore holds that there is a sensible open question for any natural realist theory (Moore, 1903, para [1]-[14]). Say a theory holds that what is good is what will maximise happiness. Moore holds that we can sensibly ask: an action may maximise happiness, but is it good? Similarly, we can sensibly ask: having this belief is useful, but should I really have this belief? But consider the same type of question put to non-natural realists. Non-natural realists hold that some non-natural facts indicate what is good. We can sensibly ask: an action may be indicated as good by the non-natural facts, but is it *really* good? Non-natural facts feel as distant from normativity as natural facts, to me.

7.2.3. The No-Metaphysics-Feels-Right Conjecture

I cannot imagine a metaphysics that feels the same as normativity. No combination of natural, non-natural, semantic, conceptual or constitutive facts that I have considered feels the same, to me. Indeed, I suspect a significant proportion of philosophers have a similar psychology to me. Let me offer this conjecture: whatever metaphysics of normativity one offers, there will be a significant percentage of the philosophical community that will genuinely feel as if normative facts are different from the facts indicated by that metaphysics. In response, advocates of such a metaphysics can only dig in their heels, pointing out that their account *just is* what normativity is. Non-naturalists might call it a brute fact. Constitutivists might maintain there is no need to say any more. Semantic-focused philosophers might claim that all we can do is offer an analysis of the relevant terms. Naturalists might assert that this is all that actually exists. But one way or another, philosophers must concede, there is nothing they can say to *compel* other philosophers to be satisfied by their metaphysics of normativity. If this conjecture is true, then a difference intuition about natural facts is not a mark against natural realism.

Speaking for myself, every metaphysics I have considered provokes a difference intuition. Normativity does not feel like facts about the acceptability of normative sentences. Normativity does not feel like a scheme of conceptual relationships. Normativity does not feel like some non-

natural fact that sits on top of the natural world. So that normativity does not feel like a set of natural facts does not distinguish natural realism from any other metaphysics, for me. No normative metaphysics is satisfying.

Although feelings change. Over the years, as I worked on embedded epistemic instrumentalism, it became more familiar. There are times when epistemic normativity does feel like facts about usefulness, to me. The normative facts begin to feel like facts about usefulness when I consider how facts about usefulness are objective, motivating and important.

Embedded epistemic instrumentalism describes a world that tells in favour of humans conforming to the epistemic standards. That the world tells in favour of conforming is a fact. This fact is a conjunction of many other facts. First and foremost it is a fact that accurate representations promote the achievement of a wide variety of ends. It is a fact that conforming to the epistemic standards helps us get accurate representations. Added to that, our biology, our psychology and cultures have developed to reinforce the usefulness of conforming. For humans, conforming to the standards does not merely produce accurate representations but also is intimately involved in coordinating with others, being respected by yourself and others, involved with being smart, clever and reasonable. It is a fact that there is a constellation of mechanisms that scaffold and promote the usefulness of conforming to the epistemic standards. So the feature of the world that embedded epistemic instrumentalism identifies as epistemic normativity is objective. It exists regardless of what any individual may think about it, even if it is partly constituted by our psychologies.

Embedded epistemic instrumentalism explains the motivational force produced by epistemic normativity. For almost every human, there is pressure to conform. The pressure comes from many sources. Raw usefulness exerts a pressure by piggybacking on any motivation to achieve other ends. If you are motivated to achieve a certain end, then you have a type of derivative motivation to do what achieves that end. Added to this, as a result of some evolutionary processes, for many of us, our psychologies exert a pressure to conform to epistemic standards; a desire to be smart for its own sake. Added to this, many societies police epistemic practices, offering rebukes to flouters and esteem to conformers.⁹⁹ It is a fact that there are many pressures, from usefulness,

⁹⁹ It is an interesting question just how similar the methods of policing epistemic practices are for the different cultures around the world. I suspect that there is always some method of policing but the specific methods vary. However, I may be wrong about both of these suspicions.

from evolutionary processes, from our psychologies, from our communities, all urging us to conform. This is the motivational force of the epistemic standards.

Embedded epistemic instrumentalism identifies an incredibly important feature of the world. Most humans need to be enculturated into conforming to the epistemic standards to both participate in their community and to achieve things that they want. Conforming to the epistemic standards is like learning to communicate with language. For most humans, it is a fundamental ability that forms part of the foundation for living. But it is not only humans that are affected by epistemic normativity. It affects evolutionary processes, non-human animals and even the machines and algorithms that humans create. The way the world tells in favour of having accurate representations impacts huge swaths of the rest of the world. This feature of the world is important no matter our terms, no matter our concepts, no matter the constitutive relationships between things.

Embedded epistemic instrumentalism portrays epistemic normativity as objective, motivating and important. It identifies epistemic normativity as the robust and potent usefulness of conforming to the epistemic standards. This metaphysics may not posit a *sui generis* property; this metaphysics may not be fundamental. But this metaphysics posits a complex fact that is itself fundamental to human existence, and more broadly affects much of the world we are interested in.

When I think of embedded epistemic instrumentalism like this, it starts to feel right. However, I do not think that this is a difference-maker, either. I suspect that the more a philosopher advocates for a theory, the more that they experience the world through the theory's lens, the more integrated and fundamental the theory becomes in their web of belief, the more the theory will feel right. So I hold that we should just not be concerned with what facts have the right feeling. Rather we should just concern ourselves with what the evidence indicates is more likely true. The evidence favours embedded epistemic instrumentalism.

References

- Ashton, Z., & Mizrahi, M. (2018). Intuition Talk is Not Methodologically Cheap: Empirically Testing the “Received Wisdom” About Armchair Philosophy. *Erkenntnis*, 83(3), 595-612.
- Baker, D. (2018). Skepticism About Ought Simpliciter. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics 13*: Oxford University Press.
- Bealer, G. (1996). "A Priori" Knowledge and the Scope of Philosophy. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 81(2/3), 121-142.
- Bealer, G. (1998). Intuition and the Autonomy of Philosophy. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 201-240): Rowman & Littlefield.
- Berker, S. (2013a). Epistemic Teleology and the Separateness of Propositions. *The Philosophical Review*, 122(3), 337-393.
- Berker, S. (2013b). The Rejection of Epistemic Consequentialism. *Philosophical Issues*, 23(1), 363-387.
- Berker, S. (2018). A Combinatorial Argument against Practical Reasons for Belief. *Analytic Philosophy*, 59(4), 427-470.
- Bogardus, T. (2016). Only All Naturalists Should Worry About Only One Evolutionary Debunking Argument. *Ethics*, 126(3), 636-661.
- Bourget, D., & Chalmers, D. J. (forthcoming). Philosophers on Philosophy: The 2020 PhilPapers Survey. *Philosophers' Imprint*.
- Boyd, R. (1988). How to Be a Moral Realist. In G. Sayre-McCord (Ed.), *Essays on Moral Realism*. Cambridge: Cambridge University Press.
- Bradie, M., & Harms, W. (2020). Evolutionary Epistemology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition ed.).
- Brink, D. O. (1989). *Moral Realism and the Foundations of Ethics*: Cambridge University Press.
- Brink, D. O. (2001). Realism, Naturalism, and Moral Semantics. *Social Philosophy and Policy*, 18(2), 154-176.
- Case, S. (2020). The Normative Error Theorist Cannot Avoid Self-Defeat. *Australasian Journal of Philosophy*, 98(1).
- Chappell, R. Y. (2019). Why Care About Non-Natural Reasons? *American Philosophical Quarterly*, 56(2), 125-134.
- Clifford, W. K. (1877 [1999]). The ethics of belief. In T. Madigan (Ed.), *The ethics of belief and other essays* (pp. 70–96). Amherst, MA: Prometheus.
- Climenhaga, N. (2018). Intuitions are Used as Evidence in Philosophy. *Mind*, 127(505), 69-104.
- Cohen, S. (2016). Theorizing about the epistemic. *Inquiry*, 59(7-8), 839-857.
- Conee, E., & Feldman, R. (2004). *Evidentialism: Essays in Epistemology* (Vol. 48): Oxford University Press.
- Copp, D. (1995). *Morality, Normativity, and Society* (Vol. 58): Oxford University Press.
- Copp, D. (2011). Normativity and reasons: five arguments from Parfit against normative naturalism. In G. Seay & S. Nuccetelli (Eds.), *Ethical Naturalism: Current Debates* (pp. 24-57). Cambridge: Cambridge University Press.
- Côté-Bouchard, C. (2015). Epistemic Instrumentalism and the Too Few Reasons Objection. *Philosophical Studies*, 23(3).
- Côté-Bouchard, C. (2019). Two types of epistemic instrumentalism. *Synthese*.
- Côté-Bouchard, C., & Littlejohn, C. (2018). Knowledge, Reasons, and Errors About Error Theory. In M. R. Kyriacou C. (Ed.), *Metaepistemology*. Cham: Palgrave Macmillan.
- Cottingham, J., Stoothoff, R., & Murdoch, D. (1985). *The Philosophical Writings of Descartes: Volume 2*: Cambridge University Press.
- Cowie, C. (2014). In defence of instrumentalism about epistemic normativity. *Synthese*, 191(16).

- Cowie, C. (2016). Good News for Moral Error Theorists: A Master Argument Against Companions in Guilt Strategies. *Australasian Journal of Philosophy*, 94(1), 115-130.
- Crupi, V. (2021). Confirmation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition).
- Cuneo, T. (2007). *The Normative Web: An Argument for Moral Realism* (Vol. 69): Oxford University Press.
- Cuneo, T., & Shafer-Landau, R. (2014). The moral fixed points: new directions for moral nonnaturalism. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 171(3), 399-443.
- Descartes, R. (2017). *René Descartes: Meditations on First Philosophy: With Selections From the Objections and Replies* (J. Cottingham, Trans. J. Cottingham Ed. 2nd Edition ed.). Cambridge: Cambridge University Press.
- Dummett, M. (1958). Truth. *Proceedings of the Aristotelian Society*, 59, 141-162.
- Dyke, M. M. (2021). Could our epistemic reasons be collective practical reasons? *Noûs*, 55, 842-862.
- Empiricus, S. (2000). *Outlines of Scepticism* (J. Annas & J. Barnes Eds.). Cambridge: Cambridge University Press.
- Engel, P. (2013). In Defence of Normativism about the Aim of Belief. In T. Chan (Ed.), *The Aim of Belief* (pp. 31-62): Oxford University Press.
- Enoch, D. (2006). Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action. *The Philosophical Review*, 115(2), 169-198.
- Enoch, D. (2011). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford University Press UK.
- Ereshefsky, M. (2022). Species. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition).=
- Eriksson, L., & Hájek, A. (2007). What are degrees of belief. *Studia Logica*, 86(2), 185-215.
- Evans, C. S. (2018). Moral Arguments for the Existence of God. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
- Evers, D. (2015). Street on evolution and the normativity of epistemic reasons. *Synthese*, 192(11), 3663-3676.
- Ferrero, L. (2009). Constitutivism and the Inescapability of Agency. *Oxford Studies in Metaethics*, 4, 303-333.
- Fine, K. (1994). Essence and Modality: The Second Philosophical Perspectives Lecture. *Philosophical Perspectives*, 8, 1-16.
- Finlay, S. (in press). Beyond Normativity: Can Metaethics Escape Samsara's Wheel? In S. Kirchin (Ed.), *The Future of Normativity*. Oxford University Press.
- Fleisher, W. (2018). Rational endorsement. *Philosophical Studies*, 175(10), 2649-2675.
- Foley, R. (1987). *The Theory of Epistemic Rationality* (Vol. 99): Harvard University Press.
- Foley, R. (1993). *Working Without a Net: A Study of Egocentric Epistemology* (Vol. 104): Oxford University Press.
- Foot, P. (1972). Morality as a System of Hypothetical Imperatives. *The Philosophical Review*, 81(3), 305-316.
- Fuller, L. L. (1958). Positivism and Fidelity to Law: A Reply to Professor Hart. *Harvard Law Review*, 71(4), 630-672.
- Gettier, E. (1963). Is Justified True Belief Knowledge? *Analysis*, 23(6), 121-123.
- Giere, R. N. (1989). Scientific rationality as instrumental rationality. *Studies in History and Philosophy of Science Part A*, 20(3), 377-384.
- Gödel, K. (1949). An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation. *Reviews of Modern Physics*, 21(3), 447-450.
- Goldman, A. I. (1967). A causal theory of knowing. *Journal of Philosophy*, 64(12), 357-372.

- Gould, S. J., & Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205(1161), 581-598.
- Greco, D. (2015). Epistemological Open Questions. *Australasian Journal of Philosophy*, 93(3), 509-523.
- Greenberg, A. (2020). Should I believe all the truths? *Synthese*, 197(8).
- Hájek, A. (2007). My philosophical position says <p> and I don't believe <p>. In M. S. Green & J. N. Williams (Eds.), *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford University Press.
- Hájek, A. (2012). Is Strict Coherence Coherent? *Dialectica*, 66(3), 411-424.
- Hare, R. M. (1964). The Promising Game. *Revue Internationale de Philosophie*, 18(70 (4)), 398-412.
- Hart, H. L. A. (1958). Positivism and the Separation of Law and Morals. *Harvard Law Review*, 71(4), 593-629.
- Hooker, B. (2016). Rule Consequentialism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition).
- Horgan, T., & Timmons, M. (1991). New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research*, 16, 447-465.
- Horgan, T., & Timmons, M. (1992a). Troubles for new wave moral semantics: The 'open question argument' revived. *Philosophical Papers*, 21(3), 153-175.
- Horgan, T., & Timmons, M. (1992b). Troubles on moral twin earth: Moral queerness revived. *Synthese*, 92(2), 221-260.
- Huemer, M. (2005). *Ethical Intuitionism* (Vol. 78): Palgrave Macmillan.
- Hull, D. L. (1986). On Human Nature. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1986, 3-13.
- Husi, S. (2013). Why Reasons Skepticism is Not Self-Defeating. *European Journal of Philosophy*, 21(3), 424-449.
- Isserow, J. (2019). Evolutionary Hypotheses and Moral Skepticism. *Erkenntnis*, 84(5), 1025-1045.
- Jackson, F. (1991). Decision-Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics*, 101(3), 461-482.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford, United Kingdom: Oxford University Press.
- Joyce, R. (2001). *The Myth of Morality*. Cambridge University Press.
- Joyce, R. (2006). *The Evolution of Morality*. Cambridge, MA.: MIT Press.
- Joyce, R. (2016). Reply: Confessions of a Modest Debunker. In U. D. Leibowitz & N. Sinclair (Eds.), *Explanation in Ethics and Mathematics: Debunking and Dispensability*. UK: Oxford University Press
- Joyce, R. (2017). Fictionalism in Metaethics. In T. McPherson & D. Plunkett (Eds.), *The Routledge Handbook of Metaethics* (pp. 72-86): Routledge.
- Kelly, T. (2003). Epistemic Rationality as Instrumental Rationality: A Critique. *Philosophy and Phenomenological Research*, 66(3), 612-640.
- Kelly, T. (2005). Moorean Facts and Belief Revision, or Can the Skeptic Win? *Philosophical Perspectives*, 19(1), 179-209.
- Kelly, T. (2007). Evidence and Normativity: Reply to Leite. *Philosophy and Phenomenological Research*, 75(2), 465-474.
- Kilov, D., & Hendy, C. (2022). Pundits and Possibilities: Philosophers Are Not Modal Experts. *Australasian Journal of Philosophy*, 1-20.
- Kitcher, P. (1992). The Naturalists Return. *The Philosophical Review*, 101(1), 53-114.
- Knobe, J., & Nichols, S. (2017). Experimental Philosophy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition).
- Kornblith, H. (1993). Epistemic Normativity. *Synthese*, 94(3), 357-376.
- Kornblith, H. (2007). Naturalism and Intuitions. *Grazer Philosophische Studien*, 74(1), 27-49.

- Kornblith, H. (2019). Our Rational Nature. In H. Kornblith (Ed.), *Second Thoughts and the Epistemological Enterprise* (pp. 243-261). Cambridge: Cambridge University Press.
- Kyriacou, C. (2018). From Moral Fixed Points to Epistemic Fixed Points. In C. Kyriacou & R. McKenna (Eds.), *Metaepistemology: Realism and Anti-Realism* (pp. 71-95). Springer International Publishing.
- Kyriacou, C. (2019). Moral Fixed Points, Rationality and the 'Why Be Moral?' Question. *Erkenntnis*, 1-18.
- Lange, M. (2002). Baseball, Pessimistic Inductions and the Turnover Fallacy. *Analysis*, 62(4), 281-285.
- Laudan, L. (1981). A Confutation of Convergent Realism. *Philosophy of Science*, 48(1), 19-49.
- Laudan, L. (1986). *Science and Values: An Essay on the Aims of Science and Their Role in Scientific Debate*. Berkeley, CA, USA: University of California Press.
- Laudan, L. (1990a). Aim-less epistemology? *Studies in History and Philosophy of Science Part A*, 21(2), 315-322.
- Laudan, L. (1990b). Normative Naturalism. *Philosophy of Science*, 57(1), 44-59.
- Leite, A. (2007). Epistemic Instrumentalism and Reasons for Belief: A Reply to Tom Kelly's "Epistemic Rationality as Instrumental Rationality: A Critique". *Philosophy and Phenomenological Research*, 75(2), 456-464.
- Lewis, P. J. (2001). Why the Pessimistic Induction Is a Fallacy. *Synthese*, 129(3), 371-380.
- Lockard, M. (2013). Epistemic instrumentalism. *Synthese*, 190(9), 1701-1718.
- Lutz, M. (2018). What Makes Evolution a Defeater? *Erkenntnis*, 83(6), 1105-1126.
- Lutz, M., & Lenman, J. (2021). Moral Naturalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
- Lyons, T. D. (2017). Epistemic selectivity, historical threats, and the non-epistemic tenets of scientific realism. *Synthese*, 194(9), 3203-3219.
- Machery, E. (2017). *Philosophy Within its Proper Bounds*: Oxford University Press.
- Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*: Penguin.
- Maguire, B., & Woods, J. (2020). The Game of Belief. *The Philosophical Review*, 129(2), 211-249.
- Moore, G. E. (1903). *Principia Ethica* (reprinted ed.): Cambridge University Press.
- Murphy, A. (2021). Chris Cowie, Morality and Epistemic Judgement, (OUP), 2019. *Ethical Theory and Moral Practice*, 24(2),
- Murphy, M. (2019). Theological Voluntarism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition).
- Nadler, J. (2008). Hart, Fuller and the Connection between Law and Justice. *Law and Philosophy*, 27(1), 1-34.
- Nagel, T. (1997). *The Last Word*: Oxford University Press.
- Niiniluoto, I. (2017). Optimistic realism about scientific progress. *Synthese*, 194(9), 3291-3309.
- Nolfi, K. (2018). Why Only Evidential Considerations Can Justify Belief. In C. McHugh, J. Way, & D. Whiting (Eds.), *Normativity: Epistemic and Practical* (pp. 179-199). Oxford: Oxford University Press.
- Nozick, R. (1993). *The Nature of Rationality* (Vol. 45): Princeton University Press.
- Paakkunainen, H. (2018). Doing Away With The 'Shmagency' Objection To Constitutivism. *Manuscrito*, 41(4), 431-480.
- Parfit, D. (2011). *On What Matters: Two-Volume Set*: Oxford University Press.
- Plunkett, D., Buchak, L., & Lombrozo, T. (2019). When and why people think beliefs are "debunked" by scientific explanations of their origins. *Mind & Language*.
- Prawitz, D. (2012). Truth as an Epistemic Notion. *Topoi*, 31(1), 9-16.
- Pritchard, D., Turri, j., & Carter, J. A. (2022). The Value of Knowledge. In Edward N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition).
- Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth* (Vol. 109): Routledge.

- Putnam, H. (1975). What is Mathematical Truth? In *Mathematics, Matter and Method* (Vol. 1, pp. 60--78): Cambridge University Press.
- Quine, W. V. (1969). Epistemology Naturalized. In *Ontological Relativity & Other Essays* (pp. 69--90). New York: Columbia University Press.
- Quine, W. V., & Ullian, J. S. (1970). *The Web of Belief*: Random House.
- Ramsey, F. (1926). Truth and probability. In A. Eagle (Ed.), *Philosophy of Probability: Contemporary Readings* (pp. 52-94): Routledge.
- Rinard, S. (2017). No Exception for Belief. *Philosophy and Phenomenological Research*, 94(1), 121-143.
- Rinard, S. (in press). Eliminating Epistemic Rationality Philosophy and Phenomenological Research.
- Rosati, C. S. (1995). Naturalism, Normativity, and the Open Question Argument. *Noûs*, 29(1), 46-70.
- Russell, B. (1948 [2009]). *Human Knowledge: Its Scope and Limits*: Taylor & Francis.
- Saatsi, J. (2005). On the Pessimistic Induction and Two Fallacies. *Philosophy of Science*, 72(5), 1088-1098.
- Scanlon, T. (1998). What We Owe to Each Other (Vol. 66). Belknap Press of Harvard University Press.
- Scanlon, T. M. (2014). *Being Realistic About Reasons*: Oxford University Press.
- Schloss, J., & Murray, M. J. (2009). *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion*: Oxford University Press.
- Schroeder, M. A. (2007). *Slaves of the Passions*: Oxford University Press.
- Schwitzgebel, E. (2021). Belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
- Shafer-Landau, R. (2003). *Moral Realism: A Defence*. New York: Oxford University Press.
- Shah, N. (2003). How Truth Governs Belief. *The Philosophical Review*, 112(4), 447-482.
- Shah, N., & Velleman, J. D. (2005). Doxastic Deliberation. *The Philosophical Review*, 114(4), 497-534.
- Sharadin, N. (2018). Epistemic instrumentalism and the reason to believe in accord with the evidence. *Synthese*, 195(9), 3791-3809.
- Siegel, H. (1989). Philosophy of science naturalized? Some problems with Giere's naturalism. *Studies in History and Philosophy of Science Part A*, 20(3), 365-375.
- Siegel, H. (2019). Epistemic Rationality. *Metaphilosophy*, 50(5), 608-630.
- Singer, D. J. (2018a). How to be an Epistemic Consequentialist. *Philosophical Quarterly*, 68(272), 580-602.
- Singer, D. J. (2018b). Permissible Epistemic Trade-Offs. *Australasian Journal of Philosophy*, 1-13.
- Sosa, E. (2003). The Place of Truth in Epistemology. In M. D. a. L. Zagzebski (Ed.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. New York: Oxford University Press.
- Sterelny, K., & Fraser, B. (2017). Evolution and Moral Realism. *British Journal for the Philosophy of Science*, 68(4).
- Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, 127(1), 109-166.
- Street, S. (2009). Evolution and the Normativity of Epistemic Reasons. *Canadian Journal of Philosophy*, 39, 213-248.
- Streumer, B. (2013). CAN WE BELIEVE THE ERROR THEORY? *The Journal of Philosophy*, 110(4), 194-212. Retrieved from
- Streumer, B. (2017). *Unbelievable Errors: An Error Theory About All Normative Judgments*: Oxford University Press.
- Swindlehurst, Z. M. (2020). The knowledge norm of belief. *Thought: A Journal of Philosophy*, 9(1), 43-50.
- Talbot, B. (2014). Truth promoting non-evidential reasons for belief. *Philosophical Studies*, 168(3), 599-618.

- Van Inwagen, P. (1998). Modal Epistemology. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 92(1/2), 67-84.
- van Roojen, M. (2018). Moral Cognitivism vs. Non-Cognitivism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition).
- Vavova, K. (2014). Debunking Evolutionary Debunking. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (Vol. 9, pp. 76-101).
- Vavova, K. (2015). Evolutionary Debunking of Moral Realism. *Philosophy Compass*, 10(2), 104-116.
- Vavova, K. (2021). The limits of rational belief revision: A dilemma for the Darwinian debunker. *Noûs*, 55.
- Veluwenkamp, H. (2017). Parfit's and Scanlon's Non-Metaphysical Moral Realism as Alethic Pluralism. *Ethical Theory and Moral Practice*, 20(4), 751-761.
- Vickers, P. (2013). A Confrontation of Convergent Realism. *Philosophy of Science*, 80(2), 189-211.
- Vineberg, S. (2022). Dutch Book Arguments. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition).
- Vogelstein, E. (2012). Subjective Reasons. *Ethical Theory and Moral Practice*, 15(2), 239-257.
- Way, J. (2012). Transmission and the Wrong Kind of Reason. *Ethics*, 122(3), 489-515.
- Wedgwood, R. (2002). The Aim of Belief. *Philosophical Perspectives*, 16, 267-297.
- Wedgwood, R. (2013). The Right Thing to Believe. In T. Chan (Ed.), *The Aim of Belief* (pp. 123-139): Oxford University Press.
- Weisberg, M. (2006). Water is Not H₂O. In D. Baird, E. Scerri, & L. McIntyre (Eds.), *Philosophy Of Chemistry: Synthesis of a New Discipline* (pp. 337-345). Dordrecht: Springer Netherlands.
- Whiting, D. (2010). Should I Believe the Truth? *Dialectica*, 64(2), 213-224.
- Wierzbicka, A. (2006). *English: meaning and culture*. New York: Oxford University Press.
- Williams, B. (2002). *Truth and Truthfulness: An Essay in Genealogy* (Vol. 78): Princeton University Press.
- Williamson, T. (2000). *Knowledge and its Limits*: Oxford University Press.
- Williamson, T. (2004). Philosophical 'Intuitions' and Scepticism about Judgement. *Dialectica*, 58(1), 109-153.
- Williamson, T. (2007). *The Philosophy of Philosophy*: Wiley-Blackwell.
- Wray, K. B. (2015). Pessimistic Inductions: Four Varieties. *International Studies in the Philosophy of Science*, 29(1), 61-73.
- Zollman, K. J. S. (2010). The Epistemic Benefit of Transient Diversity. *Erkenntnis* (1975-), 72(1), 17-35.