

2021-05-28

Making Up Reasons: constructivism, functionalism, and reasons

Stein, Joshua David

Stein, J. D. (2021). Making Up Reasons: constructivism, functionalism, and reasons (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from <https://prism.ucalgary.ca>.

<http://hdl.handle.net/1880/113487>

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Making Up Reasons: constructivism, functionalism, and reasons

by

Joshua David Stein

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PHILOSOPHY

CALGARY, ALBERTA

MAY, 2021

© Joshua David Stein 2021

Abstract This dissertation expounds and defends a theory of what it is for some fact to be a reason. The goal of the dissertation is to combine the role of reasons in justifying behavior with the role of reasons in motivating behavior, to give a more robust, practically viable account of reasons that can be applied across a range of domains and disciplines. In advancing this view, the dissertation advances the unorthodox philosophical view that in order for some fact to be a reason, that fact must *both* justify *and* motivate the agent acting in the relevant way.

The analysis of being a reason advanced in this dissertation is formulated as follows:

Some fact R is a *pro tanto* reason for an agent S to do an act of type ϕ in a circumstance C *iff* (1) R is evidence that S ought to ϕ *and* (2) R entails S ϕ -ing in C promotes or protects something of value.

Condition (1) and (2) correspond to the justifying and motivating roles of reasons, respectively.

In defending this analysis, the dissertation takes a functionalist approach to considering the roles. On a functionalist approach, some fact is a reason if and only that fact functions as a motivation *and* justification for action.

The methodology of this dissertation includes assessing whether the analysis I have proposed above and various competing analyses are correct in their assessment of whether some fact is a reason in a range of cases. This approach to assessing the extension of the concept of being a reason helps to refine and clarify the analyses; it establishes where the differences in case judgments lie between analyses of reasons.

The dissertation then argues that the analysis expounded and defended in this dissertation is especially well-positioned to account for the way that reasons are used in personal deliberation and social interaction. As reasons are used to predict, anticipate, explain, and influence the behavior of other agents, an analysis of reasons should improve our understanding of these practical uses.

Preface

Preface

Reasons are a central feature of our internal lives and our social world.

When deciding how I am going to manage my summer, what my goals are going to be for publishing, where I am going to live, and so on, I consider a range of practical reasons. My mental health, financial security, and professional development goals are each composed of facts; those facts require consideration in making complex decisions. When I consider these facts, they move me (or not) to certain plans of action. These facts are reasons, justifying and motivating me to perform certain actions. These facts require some deliberation, including sorting through which facts warrant priority and which can be subordinated.

I set out writing this dissertation because I was convinced that reasons matter. Moreover, I was convinced that theories of reasons matter, that it made a difference to how we prioritize and evaluate reasons, and thereby held sway over what agents ought to do in many cases. My hope is that points of the dissertation illustrate how we should understand reasons, and how this impacts our judgments about what we (and other agents) ought to do. This is one way to see that the theory of reasons matters, that the differences between theories of reasons that treat reasons as immutable and independent of us (on the one hand) or as products of our desires and values (on the other) matters to what we ought to do and are bound to do, and therefore getting the theory right makes a difference.

As I wait to hear back about job offers, I consider the reasons of those involved in selecting job candidates. Do they care about the Ph.D. granting institution? The publications? The writing sample? The research projects?

Which of these considerations matters most and what sorts of less obvious facts might influence the process? How might these impact my prospects of getting hired, or change the way I prepare my applications? I consider the reasons of other actors, though I rarely know enough about those actors to make well-informed predictions about which facts will determine the outcome. Even given my limited knowledge, this is the best approach I have to navigating the application process.

Making sense of the reasons which move others to act and which justify their acts play a significant role in how I relate to others. It is a major component of the foundation for how we understand and relate to each other. A compellingly crafted literary character will have reasons that a reader can understand, reasons that inform actions which drive the plot; this is how we relate to the characters and develop a literary discourse around the author's reasons, as well as the character's, for acting in such-and-such a way. This is part of how we recognize that other people are deliberating, that they respond to reasons.

Ishiguro's meditation on stoicism resolves with a direct consideration of reasons not just for an action, but for adopting a broad set of actions and deliberate efforts to focus one's thought on the present, rather than the past:

... for a great many people, the evening is the most enjoyable part of the day. Perhaps, then, there is something to his advice that I should cease looking back so much, that I should adopt a more positive outlook and try to make the best of what remains of my day... What is the point in worrying oneself too much about what one could or could not have done to control the course

one's life took? Surely it is enough that the likes of you and I at least try to make our small contribution count for something true and worthy. And if some of us are prepared to sacrifice much in life in order to pursue such aspirations, surely that is in itself, whatever the outcome, cause for pride and contentment.¹

In attempting to placate my anxiety and constant second-guessing, I found myself moved by Ishiguro's invocation. While I am not always able to adhere to those reasons, recognizing the reasons helps in knowing what I ought to do and in motivating me to do it.

Ishiguro's protagonist's attempt to justify and motivate a decidedly stoic outlook and stunt neurotic reimagination of what might have been is driven by the consideration of reasons, of how (by his own lights) the later days of the protagonist's life may fair better. Even (perhaps especially) those of us who cannot overcome our neurotic dispositions to move ourselves to act in this way recognize the familiar pattern of thought and can relate to this character's efforts. Ishiguro's writing is moving because it provides both justification for the patterns of intentional thought and the emotional motivation to pursue those patterns.

The posture of this dissertation is towards an especially restrictive account of what it is to be a reason, because the dissertation requires a the combination of justification and motivation to capture the familiar character of reasons, in our personal experience and in relating to others. Some analyses of reasons hold that what it is for a fact to be a reason is *either just to justify an act or just to motivate an act*; these views are my primary

¹Kazuo Ishiguro, *The Remains of the Day* (1989), p. 244

interlocutors throughout this dissertation, arguing that just one or the other will not do. To capture the character of reasons, we need an analysis that includes both elements.

Some philosophers hold that the principal, central role of reasons is to justify actions. My deepest encounter with this view was in a seminar with the late Derek Parfit, who was a staunch advocate of the view that the central role of reasons was in justifying action, and that all else followed. Parfit held a quasi-Benthamite view, where there was a single unified element of good which grounded justifications for action and, therefore, grounded reasons. On this view, what made one fare better, maximizing pleasure over the long term (including health and other things he saw as instrumentally necessary to that pleasure) grounded reasons for action. He was a staunch advocate of oatmeal and a vicious opponent of smoking.

To maximize pleasure, or the instrumental elements of pleasure including (for example) health, was a core element of justification and was the mouth of the river out of which all of ethics flowed. Our moral obligations were parasitic on making it possible for others (especially those disadvantaged by circumstance) to live with pleasure and a range of preconditions that made pleasure more easily attainable, like health, education, shelter, and other such necessities. The justification for helping others fare well, that we set aside our biases towards the personal and the proximate. This analysis provided a useful bridge from the practical to the moral was deeply persuasive to me. The commentary on the practical reasons Parfit developed in his work left me with a deep dissatisfaction.

I come from a family of smokers; in particular, his work led me to consider the decisions of my late grandmother Hilda. Hilda passed away while

I was taking the course with Parfit. During the last few years of her life, her smoking had limited her mobility. That she would require supplemental oxygen was an entirely predictable consequence (my grandfather, also a smoker, had required oxygen as well during the last years of his life a decade before), but what struck me about my grandmother's decision to continue smoking was that it struck me as well supported by her reasons for action.

Perhaps she was, as a matter of fact, unduly influenced by her addiction to nicotine which developed out of decades of smoking, and so she may not have been reason responsive on the issue. But even when I set the possible addictive considerations aside, the decision to smoke struck me as entirely reasonable. In spite of Parfit's insistence that the impacts of smoking on an individual's welfare provided decisive justification to the contrary, her decision to smoke was well supported by her circumstances, preferences, and attitudes to risk. The circumstances of my grandmother's life, and of her motivations for action at that point in her life, rendered the decision reasonable. She survived the Holocaust (losing most of her family, including her first husband), moved to the United States, and raised two children and participated actively in raising each of her four grandchildren. She had survived long odds, and decided that the risk of losses to her mobility and life expectancy were worth it for the pleasure of enjoying those cigarettes. Given that she retained considerable independence into her 80s and lived well into her 90s, she won the gamble on her life expectancy, but even if she hadn't, then it still would have been an entirely reasonable gamble to make.

The alternative view on offer in that seminar with Parfit, advocated primarily by co-instructor Sharon Street, was that some fact being a reason

depends on the attitudes of the agent to whom the reason applies. My grandmother's reasons for smoking, or quitting, depended on her attitudes.

This is because, following the long tradition of arguments developed by eminent figures like [Davidson \(1963\)](#) and [Williams \(1979\)](#), to be a reason is to participate in a process that is at least partly psychological. The process by which individuals arrive at deliberate decisions to act in a particular way is not merely a matter of the procedures of justifying; rather, the process is tied up with the psychology of the agent and the mess of individuals' desires, aims, and beliefs regarding what and how they'd like their lives to be.

In the domain of the practical, I found this deeply persuasive. *Contra* Parfit, there was not a singular core value that applied to each and every individual with overwhelming strength. Rather, the value of pleasure was subject to interrogation under the light of each individual, who may they care a great deal about pleasure, how they conceived of it, how they weigh the short and long term, and may choose instead to subordinate pleasure to some other pursuit, or to conceive of short and long term pleasures (as my grandmother did) in a less risk-averse way.

The problem with this story was that some lights, in some domains, seemed to be broken in a way that I could not brush aside. In the domain of the practical, allowing individuals to pursue what they believed was best for themselves seemed fine. When this attitude-dependence transgressed into the domain of the moral, then it became an especially difficult problem.

Could the attitudes of a particular agent similarly ground reasons for moral decisions? And what about the circumstances where those attitudes ground morally degenerate reasons? I spent a good deal of my time trying

to develop a response to this worry. I will present the best answers I have been able to develop in the 7 or so years since in chapter 6.

The very broad strokes, though, were that I found both justification-oriented and motivation-oriented theories on offer to be inadequate in pivotal respects. They seemed to me to call some facts reasons when those facts just were not, and to make these mistakes in a deeply consequential way.

My estimation was about the over-inclusion of the most plausible constructivist theories. Most of the criticisms of the attitude-driven analysis, centered on motivation, already held that such a theory would hold that a lot of facts were reasons when those facts should not be considered reasons. Put another way, those theories of reasons would over-generate the extension of being a reason.

By contrast, my own critique of the justification-focused view put forward by Parfit was that it also took some facts to be reasons when those facts lacked any motivational force for the agent to whom they were supposed to apply. Parfit's analysis took some facts to be reasons in cases where they just did not do the requisite work.

The result of this realization was to try and create a theory of reasons that excluded the cases on each of the two competing approaches that struck me as inadequate. As such, what I develop throughout this dissertation is a decidedly more restrictive theory of reasons than either of the two principal interlocutors. My analysis requires *both* conditions be satisfied, and so takes as reasons only the intersection of the two analyses of reasons. As I acknowledge through this dissertation, I am not the first person to develop such a hybrid theory.

The point of difference in my own approach is the emphasis on the range of roles that reasons play: they justify and motivate; they explain and influence. This range of functional roles is a part of the reason (pardon me) that reasons merit such significant philosophical attention. The focus on developing the social roles, especially in the later sections of this dissertation, is a reflection of that motivation and the project's practical orientation.

Davidson put forward a theory of reasons that was principally oriented towards using reasons to explain behaviors. As such, he focused on the psychological elements and the use of motivations in his account. What he produced was a good analysis of how reasons play a psychological role in individual action, understanding the role that facts play in informing the psychological processes and eventually the deliberate actions of those individuals. But the approach that I develop in this dissertation has a significantly broader application, scaffolding the account developed by Davidson (and other motivation oriented theorists working on reasons) into an account not just of individual behavior, but also of the norms of justification, the social facts which act as reasons for individuals and communities, and the ways in which individuals relate to each other and understand each others' behavior.

Above I note that reasons are not just a part of our personal experience, but also a part of how we relate to one another. Recognizing the reasons under discussion in Ishiguro's advocacy of stoicism helps me to understand him (and his character) in a richer, deeper way than merely acknowledging the acts that these reasons justify and motivate. Further, considering that passage moves me, especially in the late days of working on my dissertation, when the inclination to obsess over what I might have done better, or

differently, overwhelms me. Ishiguro's passage helps me to understand, it provides a vehicle for explaining a behavioral disposition, but it also influences me to act.

Combining justification and motivation allows us to understand how we influence and are influenced by each other. We understand how it is that an appeal to our desires move us to act; we also understand that sometimes mere appeal to motivations without justification may be effective at causing us to act, but is normatively deficient. We can recognize that some instances of influence are effective, but unjustified, therefore producing behavior which is not well-reasoned. The layering on of the justificatory dimension gives us an understanding of the relation between what is well motivated and what is justified; unreasonable or unjustified acts are often very well motivated.

My first year of doctoral work saw the election of Donald Trump to the office of President of the United States, and so the entirety of my work exists in the shadow of the age of viscerally motivated and profoundly unreasonable behavior of people in my own country. Understanding the deterioration of the reasonableness of my country required consideration of the social psychological analysis that grows into, around, and out of theories of reasons (including Davidson's), along with the consideration of how that work fits into a broader political and moral world that is supposed to be informed by reasons. Parfit's prescient lesson, that reasons are a shared source-point for both the practical and moral domains, held especially firm in trying to make sense of both the motivational strength and justificatory deficiency of the facts many in my country considered to be reasons.

As with creating a hybrid theory of reasons, I am hardly alone in trying to orient a theory of reasons towards political issues. Raz (1999) and O'Neill (2015) are major influences in this regard, but the difference between my approach and their distinct approaches is the significant emphasis on integrating the elements of the personal psychology alongside social constructivism to develop a rich theory. Like Raz and O'Neill, the goal is to achieve a framework for talking about reasons in the political and moral sphere that is both psychological and sociologically explanatory *and* provides assessment of the justification. My theory of reasons is served directly by its ability to provide both parts of this framework, and the most significant consideration in favor of utilizing my theory of reasons is presented in chapter 5, as I expound on how my theory of reasons contributes to this framework.

Acknowledgements

Writing this dissertation has been an intensely personal process. The most important lesson that I have drawn from this project was finding my own voice in my thinking and writing again, instead of regressing back into the standard sounds of academic discourse that I knew were acceptable and familiar to others. Doing that was very difficult, and I have been fortunate to have enormous personal and academic support throughout the process.

The most substantial sources of support in my life during this process has been my family. My late grandmother Hilda Stein (and my grandfather Arnost, who passed before I started this road), my grandfather John Hersey, my stepmother C.J. Hirschfield, my brother Steven Stein, my sister Micah McKechnie, my brother-in-law Forrest McClellen, my aunt and uncle Joan and Dick Means, my cousins Emily Means, Paul Means, and Mary Willis White Means (and Ridgley).

The most significant source of support has been my father, David Stein. His support has made this work possible, and the core moral lesson of orienting ourselves towards the service of others (especially those more vulnerable) is a major influence on this project and my philosophical work more broadly. While the dissertation is largely an abstract project, the practical orientation that develops in the late sections comes out of this sense of moral commitment.

John Baker and David Liebesman provided intellectual and human support throughout this process, and have been the most significant driving forces in this process. The methodological and pedagogical differences between them have provided me with the opportunity to explore my own

ideas, consider alternative approaches, and consider the goals of this project and my career extending out beyond. I mention above the lesson of trying to find my own voice in this project, and the constant reminders from both of them that this is *my* project were invaluable.

While at New York University, I took a seminar that laid the foundations of this project. That seminar was led by Sharon Street and the late Derek Parfit. That seminar motivated my interest in metaethics and theories of reasons, and helped me to see the significant role that metaethics could play in a range of other disciplines (within and beyond philosophy). As is likely to be clear throughout the dissertation itself, their ideas are a huge influence on this work.

I owe broader debts to the members of the faculty who helped me grow professionally, especially the two graduate program advisors during my time in the program, Richard Zach and Ann Levey. Among the philosophy faculty at the University of Calgary, I am grateful for conversation and collegiality with Reid Buchanan, Megan Delehanty, David Dick, Marc Ereshefsky, Jeremy Fantl, Allen Habib, Ish Haji, Yoshiki Kobasigawa, Noa Latham, Mark Migotti, Ron Wilburn, and Nicole Wyatt. David Dick and Ish Haji also served as members of my committee, for which I am grateful. I am also grateful for Erin Gibbs Von Brunschot and Daniel Star for their work on the committee as external reviewers.

I also am grateful for the terrific community of fellow graduate students and friends in and outside of the department and throughout my graduate work. Celso Alves Neto, Alican Basdemir, Eric Bohner, David Boutland, Berly Brumble, Ananya Chatteraj, Evangelian Collings, Emily Duchesne, Husna Farooqui, Archie Fields III, Dan Goss, Jamie Lombardi, Alison Mc-

Conwell, Shelley Smith, Chloe Stephenson, Rachel Taylor-Fergusson, and Elena Tumaszk-Jordan. Philosophy is social, and the conversations with you all sharpened me as a philosopher and improved me as a person.

During my time at New York University, during the early development of the ideas which are partly constitutive of this project, I was significantly informed by the ideas and challenges of a range of professional philosophers: David Chalmers, Kit Fine, Harvey Lederman, S. Matthew Liao, Thomas Nagel, Jesse Prinz, Regina Rini, Steven Schiffer, Miriam Schoenfield, and Peter Unger.

I am enormously grateful for the knowledge and guidance of the support staff in the department, without whom we would be lost in the bureaucracy. My deep gratitude to Rebecca Lesser, Denise Retzlaff, and Jen Smith.

A short story to note the strange trip that I took towards this project, which I hope might have some value to readers coming to this as research. I came to the University of Calgary expecting to continue work in neuroethics and mental health ethics; as a result of some peculiar circumstances, I was unable to pursue that project. Instead, I happened in to developing a backburner project into a full dissertation. There is a lot of randomness in the course of academic life, and we make the most of the random occurrences. This project, which has marked about seven years of work in philosophy, is the product of such randomness, and I am grateful for that randomness.

I am disinclined to think that other students at early stages of the dissertation or thesis writing process will come across this dissertation. While it is always an aspiration to be widely read, that is rarely true of dissertations (and rarer still of acknowledgements), but in the unlikely event someone does come across this passage, I strongly recommend leaning in to the

randomness and chaos of this process and taking advantage of the opportunities to interact and explore through that process under the guidance of the more personable, thoughtful members of your faculty.

Table of Contents

Preface	iv
Preface	v
Acknowledgements	xv
Table of Contents	xix
1 Introduction	1
1.1 The analysis of being a reason	2
1.2 Reasons as facts	15
1.3 Constructivism, through and through	18
1.4 Functionalism about being a reason	26
1.5 Argumentative methodology throughout the dissertation ..	35
1.6 Outline of the dissertation	37
2 The reasons as evidence condition	40
2.1 Reasons as evidence, “counting in favor”	41
2.2 Reasons as evidence, unpacking the view	47
2.3 Separating the two halves of the biconditional	56
2.4 $R \Rightarrow E$, the deliberation argument	58
2.5 Mere pretext arguments against $E \Rightarrow R$	62
2.6 Ersatz Reasons, a stranger critique of attitude-independence	74
2.7 Reasons are evidence	78
3 The promoting and protecting what is of value condition	81
3.1 Promoting and protecting what is of value	82
3.2 Being of value	89

3.3	Promoting or protecting	104
3.4	Valuing as grounding being of value, an argument	117
4	Conjoining the conjuncts	130
4.1	Conjoining the conjuncts	131
4.2	Presenting the analysis and giving reasons	136
4.3	Interdependent evidence and promoting and protecting cases	142
4.4	Strictly better acts and the reasons as evidence thesis	146
4.5	Shared evidence and its explanatory value	156
5	Added value and the social analysis of reasons	161
5.1	Social roles of reasons and explanatory value	162
5.2	Extensional improvement on existing theories of reasons	166
5.3	Basic considerations in social ontology and explanatory value	170
5.4	The robust explanatory value in accounting for the social roles	177
5.5	Explanatory value, contra alternative theories	189
6	Counterexamples to the proposed analysis of reasons	192
6.1	Counterexamples to my proposed theory of reasons	193
6.2	Entailment of normative facts	194
6.3	Ideally coherent eccentrics	202
6.4	Abhorrent societies and further worries	221
7	Making reasons matter more	237
7.1	Revisiting the course of the dissertation	238
7.2	Making reasons matter more	242
7.3	Glimpses Beyond	246
7.4	A note on value	254
	Bibliography	256

Appendix	270
Survey of potential variations and responses to ideally coherent eccentrics	271

Chapter 1

Introduction

1.1 The analysis of being a reason

This dissertation provides, expounds, and defends an analysis of what it is to be a reason. This chapter provides some of the background for developing a theory of reasons; the early sections specify some methodological commitments and situate this dissertation in the existing literature on reasons, as a theory within the constructivist literature. The end of the chapter, section 1.6,¹ does the typical introductory work of providing an outline of the dissertation.

The analysis of what it is to be a reason, which is the core thesis of this dissertation, holds that some fact being a reason is constituted by two independently necessary and jointly sufficient conditions.

Some fact R is a *pro tanto* reason for an agent S to do an act of type ϕ in a circumstance C iff (1) R is evidence that S ought to ϕ and (2) R entails S ϕ -ing in C promotes or protects something of value.

The analysis contains two conjuncts. Throughout the dissertation, I refer to (1) as “the reasons as evidence condition” as a general shorthand; I refer to (2) as “the promoting and protecting condition.”

There are a handful of technical terms in the analysis that require definition, in particular: *pro tanto*, evidence, promote, protect, and something of value. At various points in the dissertation, I give accounts of each of these terms. I also give an account of my broad treatment of agents, act types,

¹For clarity, the format of these crossreferences begins with the chapter number, followed by the section and subsection. So a crossreference formatted “m.n.o” will refer to chapter m, section n, and subsection o.

and circumstances. Table 1.1 provides a quick reference for where one can find the analyses of these terms in the dissertation.

Table 1.1: Exposition of core terms, with references

<i>pro tanto</i>	section 2.2.1
evidence	section 2.2
promote/protect	section 3.3
something of value	section 3.2
agent/act/circumstance	section 1.1.1

Each of the two conjuncts might plausibly be given as its own independent analysis of being a reason if the conjunct were given as a biconditional (rather than merely a necessary condition). For example, “R is a reason for S to ϕ iff R is evidence that S ought to ϕ ” is an entire analysis, and is defended in the literature as such.²

In section 2.4,³ I argue that the claim “R is a reason... iff R is evidence S ought to ϕ ” fails as an analysis of reasons.

The argument advanced in the section holds that the analysis overgenerates. That is, the biconditional analysis that only includes condition (1) will result in sorting some facts as reasons, when they are just not. Being evidence without promoting and protecting something of value picks a set of facts I argue are not reasons.⁴

By contrast, I argue that my own analysis is more restrictive (because it requires *both* conditions (1) and (2) be satisfied and that this restriction avoids this overgeneration.

²See [Kearns and Star \(2008, 2009\)](#), which I interact with extensively in chapter 2.

³This cross-reference formatting is structured so that a chapter M, section N, and subsection O will be presented as “section M.N.O”.

⁴See sections 2.5 and 2.6.

However, for reasons developed through the chapter, the thesis that being a reason entails being evidence one ought to act is useful as a necessary condition. The structure of this argument against taking condition (2) as its own biconditional analysis in section 4.4 is also focused on the prospects of picking out facts which are not reasons. In that section, I consider the possibility that the reasons as evidence condition could be subsumed under condition (2). I then argue that we should not take such an approach, that we should instead take the two conditions as independently necessary and jointly sufficient.

A core methodological approach of this dissertation is to foreground considerations of what reasons do, how they function, what their *role* can be in practical and moral reasoning and deliberation. It is commonly claimed in the literature that reason can and do play three roles or kinds of roles: justification, motivation, and explanation of actions.⁵

At a very high level of analysis, this dissertation argues that the functional roles of justification and motivation are each independently necessary for some fact being a reason. Conjunct (1), the reasons as evidence condition, captures how reasons justify actions. Conjunct (2), the promoting and protecting condition, captures how reasons can motivate actions.

The later chapters of the dissertation, most explicitly chapter 5, argue that a major explanatory upshot and argument in favor of my analysis of reasons is that by accounting for both justification and motivation, one can capture a wide range of the explanatory functions as well. My analysis builds both the justificatory and motivational roles in (rather than just one

⁵See Alvarez (2017), whose survey I use at several points in this dissertation.

or the other), and gets the range of social roles and a robust explanation of social phenomena for free.

1.1.1 Preliminary note on regimentation

Throughout the dissertation, I use a regimentation for reasons and related entities. Being a reason is a property of a fact (F) in relation to an agent (S), an act (ϕ), and a circumstance (C). On this regimentation, we get the form: F is a reason for S to ϕ in C.⁶

I set each of these elements up as types. A fact of type F; an agent of type S; an act of type ϕ ; a circumstance of type C. For any of these types, one might just replace the particular element (fact or agent or act or circumstance) with a set. An agent of the type S can just be replaced with the set of all of the agents who have the property of being type S. We can just replace the type with its extension (the members of the set it picks out).

My hope is that this regimentation will avoid any metaphysical prolixity, because the regimentation itself does not have any metaphysical commitments beyond that there are agents, acts, and circumstances; hopefully, those commitments are shared by all philosophers working on reasons. The use of types should be as promiscuous as possible, so that anyone who has commitments to a particular metaphysical position (e.g. that there are only sets, and not types) could still make use of this analysis without issue.

⁶I owe this regimentation to John Baker. Baker's inclination is to include a separate ordered item for occasion in order to note the reason R as bound to a particular ordered four-tuple: agent, act-type, circumstance, and occasion. While I share Baker's commitment to particularism, there is useful value in not limiting each item in the four-tuple to a single occupant.

For background on moral particularism, see [Dancy \(2017\)](#).

1.1.2 Motivating clause (1); the justificatory role and reasons as evidence

Scanlon and Parfit, two of the most important contributors to metaethics in the early 21st-century, characterize reasons in terms of “counting in favor.”⁷ While I argue in section 2.1 that “counting in favor” is inadequate, this characterization of reasons gets at something essential about what it is to be a reason. Reasons stand in relation to agents and acts; if some fact F is a reason for S to ϕ in C , then F stands in a relation to S ϕ -ing in C . Scanlon and Parfit would say that F “counts in favor” of S ϕ -ing in C . One might more generally say that the fact F *justifies* S ϕ -ing in C , at least *pro tanto*.

The opening section of chapter 2 argues that analysis provided by [Kearns and Star \(2008\)](#) makes more explicit what Parfit, Scanlon, *inter alia* have in mind when they talk about “counting in favor.” In particular, section 2.1 argues that “counting in favor” is an imprecise way of talking about justifying, and that evidence is a more fine-grained and clear way of talking about justifying.

The reasons as evidence condition, conjunct (1) in the analysis, is the claim that being a reason entails being evidence for the proposition that S ought to ϕ in C . This has become popular as a whole analysis: Some fact is a reason S ought to ϕ iff that fact is evidence S ought to ϕ .⁸

The thesis that reasons are evidence (either as a necessary condition or as a biconditional) is tied up with the function of reasons as justifying

⁷See [Scanlon \(2014\)](#); [Parfit \(2011\)](#). I provide direct quotations and explication in section 2.1.

⁸Following convention, I sometimes generalize this as the thesis that “reasons are evidence” or “reasons as evidence” as in [Kearns and Star \(2008, 2009\)](#); [Brunero \(2009\)](#), though as [Kearns and Star \(2008\)](#) note, they defend only the biconditional and *not* the identity claim.

action. In determining whether one ought to ϕ , a reason is a fact that acts as evidence for ϕ -ing (or $\neg\phi$ -ing). Replacing favoring with being evidence is a significant improvement, by specifying what it means for a fact to “count in favor.”

Both “counting in favor” and being evidence are gestures at the role that reasons play in justifying action.⁹

Determining whether one should ϕ , ought to ϕ , would be right to ϕ (or whatever phrasing is appropriate in the context) involves consideration of a certain proper subset of facts. The contents of that subset and the further facts in virtue of which a particular fact is a member of that subset are of direct interest to an analysis of reasons.

If I am deliberating about what movie to watch, there are certain facts that are relevant and other facts that are not. The current temperature in Beijing is not relevant to my deliberation. Whether it is late at night and I anticipate some possibility of falling asleep is relevant. The exact articulation of a theory of relevance is a technical project outside of the scope of this dissertation. However, one simple way of articulating relevance is whether there are possible circumstances that fall within type C such that F makes a difference to whether S ought to ϕ those circumstances.

Hare (1978) notes that relevance is multiply ambiguous and often sensitive to modification. Moral relevance is different from practical relevance is different from explanatory relevance, and so on. The term may also vary significantly across contexts. His paper focuses on *moral* relevance. My discussion here focuses on practical relevance: practical relevance is simpler

⁹It can be the case that F justifies S ϕ -ing in C without either motivating S to ϕ in C or explaining why S ϕ s in C. This is the core of my objections to an analysis that only considers reasons as evidence, that argument is advanced in section 2.5.

because it does not require a theory about what it is for a reason to be moral (rather than merely practical).

A fact is practically relevant if its presence (or absence) might make a difference to whether the agent ought to act in those circumstances. F is relevant to S ϕ -ing in C if it makes a difference to whether agents of type S ought to do acts of type ϕ in circumstances of type C. This difference making view is a precondition of Hare's analysis of moral relevance (Hare, 1978, p. 73); for my purposes, this precondition gives a useful, though admittedly reductive, characterization of practical relevance.

One advantage of my theory of reasons (and other theories that connect reasons to being evidence) is that it relates reasons to a robust philosophical literature on justification in epistemology. Reasons are evidence for a certain kind of proposition, and therefore the philosophical literature on how evidence works (and substantive philosophical disagreements about the nature of evidence) can be brought to bear on the discussion. I discuss one potentially consequential disagreement in the literature in section 2.2.4. The broad idea, though, is that seeing reasons as evidence helps to draw a connection to other areas of philosophy, especially to the extensive philosophical literature on evidence.

1.1.3 Motivating clause (2); the motivational role and being of value

The second conjunct is my own concoction; it is grounded in discussion of constructivism about reasons and values. Some fact F is a reason *only if* that fact is related to something of value V. The relation is that F entails

S ϕ -ing in C will promote or protect V. The fact that F stands in such a relation to V is often an attitude-independent fact, and one that imposes a significant constraint on the existence of reasons. I discuss the nature of promoting and protecting at length in section 3.3, and how the factual constraints operate in section 3.3.4. While important, these constraints are not the major reason for including the promoting and protecting condition, clause (2) in the analysis. The purpose of the promoting and protecting condition is to give an account of how reasons motivate agents to act.

On its face, that F is related to something of value does not seem to require that F would motivate S to ϕ . The colloquial reading of “something of value” does not require that the thing motivate S to ϕ under any conditions. Suppose a shekel is of value and it is a fact (known to S) that S ϕ -ing in C would result in S being paid a shekel. It does not follow that this fact will motivate S to ϕ in C. However, I argue that being of value, as it should be understood by philosophers, should include conditions that would make it the case that if shekels were of value (properly understood), then S would be so motivated. This follows arguments advanced by Williams (1979) on the importance of individuals’ motivation sets with respect to their reasons.

In discussing what it is for something to be of value, I argue that being of value depends on the attitudes of agents. Further, some thing being of value depends on a narrow, more fine grained kind of attitude (in comparison to other constructivist theories); that kind of attitude includes the agent’s disposition to feel, in addition to dispositions to believe and act (Scheffler, 2010, Ch. 1). The idea that being of value is attitude dependent has already been articulated and defended at length in the literature.¹⁰ My account

¹⁰See Korsgaard (2003); Street (2017), which I discuss at length in section 3.4.2.

tries to provide a clear, fine-grained articulation of the way in which the particular attitudes account for and explain the motivational character of reasons.¹¹ This move serves a few purposes in the dissertation.

First, it means that my analysis will not be subject to the counterexamples that I raise for those analyses of reasons which hold that being a reason is biconditional with being evidence. The analyses that hold reasons are biconditional with being evidence do not provide for motivational force. Laying out this account of being of value gives my analysis the ability to account for the motivational role, and thereby avoids the class of counterexamples that I raise against the analyses of reasons as evidence.

The analysis of reasons as evidence neglects some of the properties present in subjective experience of reasons (including motivation); the account of valuing helps to capture many of these subjective features. When an agent is presented with a reason to act, this reason is not just taken as a piece of trivia and put in the belief box. Rather, it has certain properties that dispose the agent to act consistent with that reason. Accounting for the attitude dependence of being of value (and, therefore, promoting and protecting something of value) helps to flesh out the dispositional properties and the agent's response to those properties.

This ties together with the major argument in favor of my theory of reasons which develops in chapter 5. Reasons play instrumental roles in social interaction, and as such play a role in how we explain various kinds of social interactions. In order to play those roles, it has to be the case that reasons (as a category, not merely as particular facts) influence the actions

¹¹See [Williams \(1979\)](#); [Alvarez \(2017\)](#).

of agents. The inclusion of the motivational role allows for this significant explanatory upshot.

Second, this attitude-dependent view of some thing being of value provides space to unpack a range of claims about the nature of reasons, claims grouped under the constructivist school of thought. I say more about constructivism in section 1.3.

Constructivist theories that view reasons as dependent primarily on the attitudes of the agent are frequently viewed as being subject to counterexamples, these including worries about eccentric agents (see section 6.3) and abhorrent societies (see sections 6.4 and 7.4). My assessment is that, while the core of such constructivist theories (the view that reasons and values are attitude dependent) is basically correct, such theories underdevelop or fail to clarify certain attitude-independent constraints on reasons. By developing these constraints¹² in both the theory of reasons proper and in specifying conditions for being of value, the counterexamples can be rendered much less forceful or even be avoided altogether.

1.1.4 Attitude dependent, dispositional properties

Dispositional properties are central to the account of valuing and some things being of value. Rather than the broad range of dispositional proper-

¹²I do significant work to develop these constraints in sections 3.3 and 3.3.2, but this is only a touch on the surface in terms of the wider possibilities for such theories, including the connections to extensive rational choice literature. One might develop these constraints out of the traditional literature on rational choice theories, as in Bell (2021); Briggs (2019). One might alternately turn to the more recent literature in dynamic and social choice theories, as in Andreou (2020); List (2013). These are all survey articles, but hopefully also illustrate why I have made the prudential decision not to address these potential approaches to constraints on rationality at such length.

ties, like fragility, I have in mind the disposition of an object to evoke certain kinds of responses in agents.¹³

Some things are disposed to produce certain responses in an agent. When I see photos of my friends' babies, I have a disposition to experience a range of emotions. What I value is tied up in my dispositions to respond in certain ways. These are attitude-dependent dispositional properties; properties of the object and agent that are sensitive to the attitudes of the agent.

If F is a reason for S to ϕ in C, then F justifies ϕ -ing in C. But F being a reason is not just a matter of F justifying S ϕ -ing; it includes the role F plays in making S feel a certain way about ϕ -ing (or prospective outcomes of ϕ -ing) in C. These feelings include (but not limited to) those which motivate action. Accounting for the range of these dispositions is ambitious, but necessary.

Reasons provide for complex judgments about our own behaviors and the behaviors of others; flattening out an account of reasons to abstract away these judgments and experiences is a mistake and deprives us of a satisfying theory of reasons. The fact that providing basic social services improves the welfare of the worst off in our communities is a reason to support such social services; appealing to the facts about welfare, though, is not just a simple two-dimensional assessment of goodness.

When we talk about the values associated with the welfare of others, there are conceptual elements (beliefs about fairness, essential elements of a good life, etc.) and there are emotional elements. Seeing, or even thinking about, individuals suffering in a way that we recognize as unjust is not just a

¹³For more comprehensive discussion, see [Smith et al. \(1989\)](#); [Brower \(1993\)](#).

doxastic matter (though the beliefs that the suffering is unjust are doxastic), but also an affective one and is tangled up with our dispositions to act, to regard as reasons, and so on.

Simplifying how we understand reasons produces theories that are too thin, that are missing some of the explanatory elements that are useful. The richness of reasons provides a major argument for my theory of reasons.

In advancing both the combined theory in chapter 4 and the social upshots of the theory in chapter 5, I illustrate the value of this more complex approach that includes dispositional properties and out more complex emotional responses. One major argument for this dissertation, which is especially explicit in chapter 5, is that this theory has much more explanatory power and depth than the available alternatives.

1.1.5 Argumentative structure of the dissertation

The justificatory and motivational roles of reasons are the center of this dissertation, which is why section 2.1 and section 3.2 respectively are dedicated to advancing them.

The structure of the arguments that develop through this dissertation are, at times, circuitous. Perhaps I spent too much time watching David Lynch films and caved to a subconscious desire to develop the ideas out of their most obvious order. This somewhat scenic approach to developing the ideas is especially difficult in chapters 2 to 4. In each of those chapters, I provide a short survey section towards the front to hopefully make a more readable discussion. These are sections 2.1.1, 3.1.2 and 4.1.1.

While each dedicated chapter advances the functional roles on their own terms, at various points I present arguments that each of these functional roles only establishes a necessary condition (rather than a biconditional) for being a reason.

This dissertation focuses on constraints on being a reason that prevent over-generation, that stop a theory of reason from extending to facts that are not reasons. I criticize accounts focused on justification for having a wide extension that includes pretexts and ersatz reasons. I criticize accounts of reasons that focus primarily on motivation for having too wide an extension, both in the empty track argument in section 4.4.1 and in the eventual considerations of historical criticisms of constructivism discussed in section 6.3.

A consequence of setting up justificatory and motivational roles of reasons in terms of their necessity and arguing that reasons are at the intersection of those roles is that there is a clear set of constraints on reasons that are brought to bear throughout the dissertation. The formal constraints are discussed at some length in section 3.3.2; the substantive and bodily constraints are discussed throughout the dissertation, especially in considering counterexamples to constructivism.

Chapter 5 considers a major explanatory upside of the theory of reasons that I advance through this dissertation. Including both the justificatory and motivational roles of reasons explains the various social roles played by reasons. These roles are extensive, and include: the social psychological phenomena whereby we predict the behavior of other members of our communities and coordinate our behavior accordingly (section 5.4.1); the social influence we exert by presenting and evaluating the reasons of other

members of our communities (section 5.2.1); the social construction of a range of norms and institutions (section 5.4.3).

It is a major argument in favor of my theory of reasons that leveraging the justificatory and motivational roles of reasons produces extensive explanations of reasons as major elements of social life.

1.2 Reasons as facts

For prudential reasons, this dissertation largely avoids discussion of whether reasons are facts or propositions.¹⁴

I treat reasons as a subset of facts, following a set of conventions in metaethics. It may be that reasons are propositions. The distinctions between facts and propositions in this context are subtle, technical, and their lengthy discussion contributes (in my estimation) little value to this project.

Some internalist approaches take it that reasons are a proper subset of attitudes, where attitudes cover a broad range of psychological states including both non-propositional and propositional attitudes; Smith wants her child to have a nap in the afternoon and this attitude is a reason to take her son to the park. When one says, following the internalist locution, “S has a reason to ϕ in C” then one might hold that this is an ascription of an attitude, like S believes that p, or desires that p. Alternatively, when one expresses reasons in terms of content, one might similarly regard reasons as beliefs; that it is raining is reason to put on a jacket, and my belief that it is raining informs my belief that I should put on a jacket.

¹⁴For some background, see (Finlay and Schroeder, 2017, Sec. 1).

The view that reasons are attitudes is mistaken and results in an impoverished theory of reasons. Holding that reasons are attitudes gets the motivational features and perhaps (with certain appropriate constraints) justification, but such a theory struggles with the idea that reasons can be points of common access, that reasons can be things we share with other agents and use to coordinate and influence behavior. My argument for pushing this theory, advanced in chapter 5, requires this common access to reasons.¹⁵ By focusing on reasons as facts, we can get the observation that individuals can appeal to the same reason through the appealing to the same fact for each agent. Consider the following cases.

In the first game of the second match in the tournament, Steven draws black again. His new opponent moves pawn to d4. R still obtains, despite the distinction between the two occasions. R applies to the act of the same time, for the same agent, across two distinct occasions. However, just as we can vary the occasions, we can vary the agents and act types.

Suppose that Steven's opponent in the third match is Tonya. Tonya knows that Steven is experienced with the King's Indian Defense; she knows the fact that R. Tonya draws white, Steven again drawing black. Because Tonya knows that Steven is experienced with the King's Indian Defense, she has reason to make a move *other than* pawn to d4. I submit in this case that R is a *pro tanto* reason for Tonya to move this way, just as it was a *pro tanto* reason for Steven to open the King's Indian Defense in the previous matches.

¹⁵A certain behaviorist approach to attitudes could accommodate this, but would do so by rendering attitudes as equivalent sets of facts about agent behaviors and dispositions. Such a theory collapses back into a theory of reasons as facts, albeit a different set of facts.

I do not expect people to accept the claim that reasons are facts on the basis of intuition. There are arguments for this presented elsewhere in the literature, as discussed by [Finlay and Schroeder \(2017\)](#); my argument in favor of the thesis that reasons are facts is tied to the core of the dissertation itself. Namely, the explanatory value of a complex theory of reasons consists partly in that theory's ability to explain a range of social phenomena; illustrating how my theory produces these explanations is the major argument of chapter 5, but we can see the bones of that explanation in this discussion of reasons as facts.

Steven and Tonya considering R in making their decisions is an instance of coordinating behavior. In this case, it is Tonya coordinating her behavior in anticipation of Steven's likely behavior under certain conditions. Joint consideration of reasons, understanding and anticipating the reasons of others, etc. is a set of significant roles that reasons play in our lives. I discuss this at length in section 5.1. Social activity involves the assessment, evaluation, and coordination of others' behaviors in part on the basis of their reasons. Conceiving of reasons as propositions or attitudes can be compatible with this analysis, but risks creating confusion about the potential object. By starting with the fact, we ensure that the contents of both Steven's and Tonya's attitudes are joint, rather than merely correlated.

On the view that reasons are facts, reasons cannot be true or false. Rather, they either exist or do not exist. Agents may have beliefs about reasons just as they do about other facts. On this view, we say "there is a reason for S to ϕ in C" (following the standard externalist articulation) and alternatively we can talk about S believing that F is a reason for S to ϕ in C. Colloquially, we say that S has a reason to ϕ in C, following the internalist locution; this

is a loose talk equivalent of saying that S believes that S is justified and motivated to ϕ in C. I largely avoid adopting the internalist phrasing, to avoid certain other wrinkles.

1.3 Constructivism, through and through

This dissertation takes a constructivist approach to reasons, values, and many of the norms governing reasons and values.

“Constructivism” is a term that varies across domains; even within metaethics, the term can refer a handful of positions.¹⁶ Frequently, these positions are differentiated by the historical traditions with which they are associated (Bagnoli, 2017, Secs. 2-4): Humean, Kantian, Aristotelian. This approach to individuating traditions often requires some significant and contentious historical exegesis. I am skeptical that this approach is satisfactory in illuminating constructivism (rather than illuminating the historical roots of constructivism). Instead, this section (and the dissertation broadly) develops the constructivist stories that I endorse and make use of in this dissertation, with little if any reference to the historical traditions. Instead, I focus on the claims of the constructivist theories that I endorse (or reject) directly.

The versions of constructivism advanced below are consistent with a few features characteristic of constructivism generally. These features can be broadly characterized by a few claims:

- The explanandum of a constructivist theory (whether reason, value, norm, etc.) is dependent on the attitudes of agents.

¹⁶These positions are surveyed in: Bagnoli (2017); Street (2008).

- Reasons are brought about through active psychological or social acts performed by the agent or community of agents.
- The explanandum (reasons; being a reason) is subject to certain attitude-independent constraints, including both formal constraints and factual constraints on the success of construction.

These are each theses for which I actively argue throughout the dissertation.

1.3.1 Constructivism about reasons

Suppose that there is an agent S , an act type ϕ , and a circumstance C . There is a set of facts which apply to S ϕ -ing in C ; some of those facts are reasons and others are not.

Some fact F is a reason for S to ϕ in C partly in virtue of the fact that the set of goals, values, desires, etc. held by S are related to F and applicable to a deliberation about whether to ϕ in C . F is a reason for S to ϕ in C partly in virtue of some evaluative attitudes of the agent, and attitude-independent constraints on the relation between those facts and the act. This can even be read off as a summary of the way my analysis of reasons works outright; the fact that F is a reason for S to ϕ in C depends on the attitudes S holds and those attitudes satisfying some formal and factual constraints, like the facts about whether the act promotes or protects things of value and whether the fact is evidence that S ought to ϕ in C .

There are potential defeaters, like when the attitude depends on a falsehood (see section 3.3.4). The nature and scope of these defeaters is subject to some disagreement amongst constructivists, and so I advance my own view on the subject and argue for some constraints throughout the disser-

tation.¹⁷ Often, the approach to understanding these defeaters in terms of historical constructivist traditions lays out a continuum where “Humean” constructivism has fewer defeaters and “Kantian” constructivism has more defeaters. I try just to consider the potential defeaters proposed in the literature on their own merits, without worrying so directly about where on that continuum this might put the theory. Let us consider one potential defeater, as a toy case.

Suppose that Bernie is thirsty, wants a drink, and there is a full glass nearby. There is a reason for Bernie to drink from the glass *only if* the contents in the glass would actually satiate his thirst; if (for example) the glass is full of petrol or something similarly noxious, then there is no such reason to drink from the glass. The facts about the content of the glass constrain the existence of a reason, independently of Bernie’s attitudes.

Some constructivist theories hold that there is a very general categorization of favoring attitudes which can bring about the existence of a reason. In section 3.2, I argue for some specific constraints on the agent and evaluative attitude. On my view, it must be the case that the evaluative attitude includes dispositions of the agent to take the fact into consideration as a reason and to feel certain emotional responses to certain facts. These dispositional properties play a role in establishing the justification and motivation of the agent during consideration of the reason, because the dispositions to feel or think certain ways are constitutive of motivations to act. Basically, the idea that all favoring attitudes are sufficient means that some very weak favoring attitudes (which do not include the dispositions to feel or act) are

¹⁷These arguments can mostly be found in chapters 3 and 6.

going to result in calling some facts “reasons” when they do not satisfy the motivational role. I argue that this is a problem for other theories of reasons.

Constructivist theories of reasons also hold that there are attitude-independent constraints on the construction of a reason; being a reason depends on attitudes, but also on not running afoul of certain considerations. In order for F to be a reason, F must avoid certain defeaters. I unpack these defeaters at length through the second and third chapters; the exact content of defeaters can prove controversial among constructivist theories. Generally, the “Humean” theories are regarded as having fewer or less restrictive defeaters while “Kantian” theories are regarded as having more or more restrictive defeaters.

The boundaries between the different camps within constructivism are porous and often unclear. Because this dissertation is not committed to a particular historical tradition, it is more useful just to enumerate the constraints and walk through the arguments for some of those constraints. The constraints I consider and where I explain and explore them are enumerated in table 1.2.

Table 1.2: Exposition of constraints, with references

Evidentiary constraints	section 6.3.4
Formal/logical constraints	section 3.3.2
“only if” factual constraints	section 3.3.4
Constraints on attitudes	section 3.3.5
Possible substantive constraints	sections 4.5 and 6.4

1.3.2 Constructivism about values

Like constructivism about reasons, constructivism about values turns on a claim about attitude-dependence. Something is of value only if some agent holds the right sorts of evaluative attitudes towards that thing, subject to certain constraints and possible defeaters. My own view is especially narrow with respect to “the right sorts of evaluative attitudes,” as I discuss in section 3.2.

What an individual values may vary depending on the circumstances, but values often float independently of those circumstances. Smith values honesty across a range of circumstances; even if there are some extreme circumstances where he does not value honesty, one should not suppose that his valuing honesty is limited to a type of circumstances.¹⁸

On constructivism about values, some thing *x* is of value partly in virtue of some agent *S* holding a certain kind of attitude towards *x*. In chapter 3, I argue at length that values are dependent not just on a general category of favoring attitude, like desiring, but rather on a more specific attitude of valuing. By fleshing out the concept of valuing in section 3.2, I illustrate both substantial features of the attitudes in virtue of which some fact is of value *and* how it is that being of value (when correctly understood) accounts for the motivational role of reasons and how such reasons have significant force when offered in social interactions.

Taking favoring attitudes generally will not suffice to generate the motivational force, and so some constructivist theories shift to motivating under

¹⁸Perhaps one might be able to formulate it in terms of circumstances through massive disjunction, but I try and avoid doing that sort of thing in this dissertation for reasons of parsimony.

“ideal” conditions or put in a very strong *ceteris paribus* clause. My approach is to focus on a narrower understanding of the attitude, to say: S holds a certain kind of attitude towards x; there are constraints on this attitude, like the fact that the attitudes S has towards x cannot depend on a false belief¹⁹ and that the constraints include dispositions to feel certain ways. But if the constraints are met and the defeaters are absent, then x is of value.

Some constructivist views disagree about the kinds of attitudes that give rise to something being of value; some constructivist views leave this open, and just use some synecdoche for the attitude. I argue that the appropriate attitude is valuing, in particular valuing as described by [Scheffler \(2010\)](#). I unpack Scheffler’s account, and how I use it, throughout chapter 3.

At a high level, the purpose of giving such a specific account of valuing is to illustrate a significant point of difficulty in philosophical discussions of motivation. Many eminent philosophers have discussed the desires or motivations of agents in broad terms, as though the constituents of an agent’s motivation set were all more-or-less homogeneous; sometimes this is useful for their own goals, but it often obscures significant differences, especially when it comes to how individuals make decisions and how individuals are motivated to act. The purpose of my analysis in terms of valuing is to provide some grist for philosophical and psychological discussion about variation in motivation. Sometimes, when considering F as a reason for S to ϕ in C, S will recognize a doxastic commitment entailed by F such that S is motivated to ϕ in C; sometimes there are physical and involuntary elements of S responding to F such that S is motivated to ϕ in C. My approach to

¹⁹I defend this particular version of the false belief constraint in section 3.3.4. As far as I know, the existing constructivist literature has not provided a precise version of this constraint.

understanding valuing helps to parse the variations in cases, recognizing that some instances of valuing motivate automatically while others require active deliberation.²⁰

The nature of an agent's motivation in seeing someone in pain and wanting to help that person is different from the agent's motivation in considering whether to complete their homework today, rather than putting it off for later in the week. The richness of psychological and phenomenological differences is a feature of theories of motivation, including how reasons can provide motivation. A theory of reasons should include some consideration of these psychological and phenomenological differences; my analysis of things being of value creates this space.

1.3.3 Constructivism about norms

Constructivism about norms is distinct from constructivism about reasons and values. Constructivism about reasons and values can be psychological; constructed reasons and values can also be social (I will discuss how the social construction of reasons and values works in chapter 5, especially section 5.4.3). By contrast, constructivism about norms must be a social constructivism.

Constructivism about norms is not contentious and so I do not dedicate energy to defending it. Rather, I give the standard articulation of the established philosophical views in section 5.3, and then implement elements

²⁰This seems like a broad gesture, but it helps in making sense of difference between philosophers who focus on "ideally considered" attitudes (like Williams and Rawls) and those who focus on the actual, immediate attitudes and dispositions (like Street and Prinz). While I do not get too far into this particular issue, chapter 3 provides a framework for sorting through these disagreements fruitfully.

throughout chapter 5. *Pace* some moral realists, some norms *just are* the products of human minds and behaviors. That norms are constructed is a standard observation. That norms play a role in the construction of reasons is interesting and important to this dissertation, especially in section 5.3 as we consider the incorporation of social ontology.

That there is such-and-such a norm is a fact; this fact may itself be a reason for an agent to act. Norms include social regulation of what acts are appropriate (or not) in a given circumstance; these are the regulative norms. Norms include the conventions governing what constitutes the performance of acts of certain types, like swearing oaths and making chess moves; these are the constitutive norms.

Constructivism about reasons and values includes the thesis that norms may play a role in whether some fact is (or is not) a reason, and in determining the character of that reason. To defend this thesis, one must clarify what norms are and how they can play such a role.

While I do not argue directly for constructivism about norms, it is important to keep this element of constructivism in mind because constructivism about norms is necessary for much of the work in chapters 5 and 6. The arguments that run through those chapters make use of ideas in social constructivism about norms, especially the claims that (1) such construction grounds facts and (2) the social facts so constructed can themselves be reasons.

1.4 Functionalism about being a reason

This dissertation adopts a particular methodological approach to the analysis of reasons. The dissertation takes being a reason as a functional property; a fact is a reason in virtue of that fact's ability to play a role in justifying and motivating behavior.²¹

Following [Fodor \(1989\)](#) and the early account of functionalism as developed in philosophy of science and mind, what I mean here by a functional role is the disposition of the object with the property to behave in a particular way under a set of conditions. In Fodor's canonical example, being an airfoil is specified by the disposition to create drag under conditions of airflow. "[T]he physics of airflow is specified with the necessary degree of causal detail, but the composition of the central object, is only functionally specified." ([Strevens, 2008](#), p. 165-166) Analogously, the causal structure of a reason is a part of the theory of reasons, just as the causal elements of the physics of airflow are part of the theory of aerodynamics, but the property of being a reason is functionally specified.

[Alvarez \(2017\)](#) distinguishes the range of properties. She treats these properties as independent: being a justification, being a motivation, and being an explanation. The article notes that this partition is common in the literature, especially a distinction between justification and motivation.²²

²¹This is not the whole of the functional role; the social functions are also a part of this, but are much harder to specify in a pithy and engaging manner. The social functions include social scientific and psychological explanation (section 5.4.1), influence of behavior (section 5.2.1), and the existence of institutional structures proper (section 5.4.3).

²²There is some semantic slippage between a few terms for justification. Some philosophers use the term "normative" while others use the term "justificatory." I focus on justification throughout this dissertation, both for uniformity and because the conceptions of normative reasons at play in the philosophical literature get slippery. For example, in some cases, a philosopher might use "normative" reason to refer to reasons enshrined by society to motivate certain behaviors.

She does, however, treat being a justificatory reason and being a motivating reason as functional properties. A justificatory reason for S to ϕ in C is a fact which can justify (or play a particular role in justifying) S ϕ -ing in C. A motivational reason for S to ϕ in C is a fact which can motivate S to ϕ in C.²³ These are functionally specified properties.

Treating being a reason (or more narrowly being a justificatory reason, a motivational reason, etc.) as a functional property is not unusual in the literature, though to the best of my knowledge this has never been made explicit.

I adopt the explicit approach to treating being a reason as a functional property for two reasons. First, making functionalism explicit helps in clarifying the extension of the analysis of being a reason, as one can assess if and insofar as a particular fact has the functional property. If we are trying to determine whether some fact counts as justifying, then the appropriate approach is to determine whether it justifies the act under the specified conditions. *Mutatis mutandis* for the other functional roles. I hope this comes through as a virtue in the relative clarity of analyzing cases.

Second, analyses of reasons that do not adopt this approach wind up with a range of commitments that are counter-intuitive or just strange. I want to avoid adopting those commitments, for example the possibility that there are reasons that can never be held by any agent (as I will discuss in section 2.6). Some significant cases of such commitments are discussed

²³R “can motivate” S to ϕ in C depends on the attitudes S has. This is an uncontroversial kind of attitude dependence. Being motivated by a reason is dependent on the individual’s attitudes; this is trivially true because the dispositions of the agent to be motivated are themselves attitudes.

at length in the dissertation, though usually in the context of presenting arguments against various competing theories of reasons.

1.4.1 The justificatory role

Reasons justify action. That R is a *pro tanto* reason for S to ϕ in C entails that R is a *pro tanto* justification for S ϕ -ing in C. It may be subject to defeaters on deliberation, and therefore not an all-things-considered or *pro toto* justification. Sometimes *pro tanto* reasons fail.

The reasons as evidence condition is at the core of the analysis and is an attempt to capture what I argue is the best approach to the justificatory role of reasons.

There are two ways of talking about justification. First, justification may refer to a factual relation between a fact and an attitude or act. The fact F justifies S believing some proposition p. The fact F justifies S doing some act of type ϕ . Given the focus on reasons for action (rather than reasons for belief), the dissertation is more concerned with justifying acts than beliefs.²⁴

A fact justifies S ϕ -ing in C based on its relevance relation and a series of constraints. These constraints are discussed at length in section 3.3.2. Often, these relevance relations are intuitive. Facts about the weather in Beijing are not relevant to my decisions about how to go about my day in Calgary. In some cases, there may be some challenges in establishing the relevance relation, or some disagreement about whether the fact is relevant. Much

²⁴If one has a view that believing is voluntary, then the analogy will be much closer between believing and acting. I will not tackle epistemic voluntarism in this dissertation and have no strong opinion on the issue. If it turns out that epistemic voluntarism is true, along with some other theses, then there may be an appropriate application of reasons to believing.

of the existing first-order discussion in practical reasoning has to do with assessing the relevance of certain facts and the application of constraints on practical reasoning to facts.

Second, justification may refer to the social act of answering a “why” question about one’s behavior. S justifies S doing ϕ in C to some other agent T. If T asks S, “why did you ϕ ?” and S responds by articulating a set of reasons, then S has justified their behavior to T.

Through the discussion of justification, I focus on the first of these; in chapter 4 and section 5.4.1 I give some discussion of the second as a social practice. The two are closely related, the fact F justifying S doing ϕ in C is also the sort of fact that S invokes to justify doing ϕ in C to T. The reason for focusing on the first is that the second requires further conceptual work around the social elements, standards of evaluation, and the like. The second (justifying *to* T) comes into play in discussing the social roles of reasons.

1.4.2 The motivational role

Williams (1979), *inter alia*, argued that being a reason requires the existence of some relevant motivational force. R is a reason for S to ϕ in C *only if* there are certain conditions under which R would motivate S to ϕ in C.²⁵ Motivation is a tricky concept.

The core of my analysis does not require adopting a specific account of motivation. My analysis uses the constructivism and attitude-dependence claims as a way to illustrate how reasons motivate. This allows for further

²⁵Williams (1979) specifies that these conditions include transparency that does not always practically obtain. So S may not in fact be motivated.

development and refinement in future discussions. For the purpose of this dissertation, the goal is to leave some of the unresolved questions open. I do not give and defend a specific, technical account of motivation. Instead, I give an account of valuing (as an attitude of agents). Valuing gives rise to a broad, rich, and intuitive sense of motivation.

This account provides enough of an account of motivation to address worries raised about justification-only theories of reasons. In section 2.5, I argue that theories of reasons that only provide treatment of the justificatory role result in treating mere pretexts as reasons, when pretexts are not reasons. I argue that pretexts fail as reasons both intuitively and on the basis of practical judgment.

Rather than specifying that motivation involves the disposition to act or the disposition to desire acting or something else, the account of valuing that I develop in this dissertation²⁶ serves to illustrate how reasons may produce each of these in different circumstances; as a result, valuing does double duty both in establishing the view that reasons are attitude dependent and in providing an account of motivation. There are several ways that this can be developed into a theory of motivation, including through the use of dispositions to act and dispositions to experience certain emotions. The result is a richer theory of reasons, more resonant with our experience of having and being motivated by reasons.

²⁶This account is drawn heavily from [Scheffler \(2010\)](#), as I make explicit throughout chapter 3.

1.4.3 Explanation, other social roles, and the master argument for the dissertation

Reasons for an agent to act matter to that agent; reasons for and against a chess move matter to the player who has to make the move. But reasons also matter to other agents.

The dissertation notes that, following [Alvarez \(2017\)](#), there are roughly three sets of roles that reasons play: justificatory, motivational, and social. The preceding sections focus the analysis on two of those roles (justificatory and motivational). A major argument for keeping the two roles as conjuncts is that the conjunction of these roles allows for a full and rich explanation of the social roles of reasons. This is the argument that develops through the later chapters, most explicitly chapter [5](#), but also through chapters [4](#) and [6](#).

The above discussion of Tonya considering reasons for Steven to move is illustrative, but that discussion is also simplistic. Game theoretic approaches present a range of simplified toy cases, as well as an understanding of how those cases increase in complexity to produce a range of more sophisticated social situations. Prisoners' dilemmas involve direct consideration of such a case, where each individual prisoner has competing *pro tanto* reasons predicated on competing strategies, minimizing risk or maximizing benefit subject to facts we know about the circumstances. Each prisoner has to consider the similar reasons of the other prisoner (in the standard case, they are shared reasons) in order to develop a plan for action.

Throughout chapters [5](#) and [6](#), I make use of the social dimensions of reasons to address significant cases of behavioral coordination through informal social norms and through the creation of formal norms and insti-

tutional structures. This explanatory upshot is a major argument in favor of considering the theory of reasons I advance in this dissertation; it provides a theory that is accessible to existing literature on the way in which the social and legal order functions.

A major influence for this dissertation is the often under-appreciated relationship between metaethics (especially theories of reasons) and the broad range of issues in political and legal philosophy. Rawls, Raz, and other major figures have touched on these relations in their work,²⁷ but often the implications for social sciences and political philosophy remain understated. The purpose of the lengthy expository articulation in chapter 5 is to put a more explicit point on the value of theories of reasons (and my theory of reasons, in particular) to political and legal philosophy, as well as the social scientific domains that interface with such philosophical domains.

Metaethics is clearly and frequently associated with ethics. Good metaethical theories can expand that value out into domains often treated as estranged second-cousins of ethics.

1.4.4 Unifying the roles

Some contemporary approaches to reasons have separated out these dimensions of reasons.²⁸ Those approaches create a taxonomy of justificatory reasons, motivational reasons, and explanatory reasons. While many such

²⁷I owe this point in part to conversations with John Baker, who noted that much of Rawls' later work on foreign policy (and the developments on that issue in post-Rawlsian, liberal political philosophy) are fundamentally oriented towards problems of reasons that emerge between and across different communities. This is explicit in [Rawls \(1989, 1993, 1997\)](#). Some other influencers who are explicit on this point: [Enoch \(2015\)](#); [O'neill \(1986\)](#); [Raz \(1990\)](#).

²⁸Surveyed in [Alvarez \(2017\)](#).

instances overlap, the goal of these approaches is to suggest that being justificatory does not require motivational, and vice versa.

Reasons, properly understood, are the intersection of (at least) the justificatory and motivational roles. The dissertation provides standard analytic philosophy arguments based on extension, namely that the focus on mere justification or mere motivation results in an unsatisfactory extension, calling some facts reasons when we should not do so. But the theory also provides some illustration of the ways in which the justificatory and motivational roles of reasons taken together help to explain a range of explanatory practices.²⁹

Jonathan Dancy writes,³⁰

If we do speak in this way, of motivating and normative reasons, this should not be taken to suggest that there are two sorts of reason, the sort that motivate and the sort that are good. There are not. There are just two questions that we use the single notion of a reason to answer. When I call a reason “motivating”, all that I am doing is issuing a reminder that the focus of our attention is on matters of motivation, for the moment. When I call it “normative”, again all that I am doing is stressing that we are currently thinking about whether it is a good reason, one that favours acting in the way proposed. (Dancy, 2000, p. 2-3)

²⁹One might regard these as an argument for the truth of theory, at least in the sense articulated by James (1895), but one might also regard this as a general argument for accepting the theory independent of its truth value. I take the latter approach in presenting the arguments throughout the dissertation.

³⁰Notable, Alvarez (2017) quotes this passage in her discussion.

In discussing this passage, Alvarez notes that this puts Dancy in clear disagreement with those philosophers who do clearly individuate motivating reasons and justificatory (or normative) reasons as a matter of theory. Dancy's endorsement of the view that there is no distinction between motivating and justificatory reasons is a commonly accepted view among many motivational internalists and those who center the attitude-dependent view of reasons (including all constructivists). However, many philosophers explicitly view motivating reasons and justificatory reasons as distinct. On Parfit's view, for example, some fact F may be a justificatory reason for S to ϕ in C completely independently of any facts about whether F would or could motivate S to ϕ in C .

One of the major advancements in this dissertation is that it argues explicitly for the unification of the justificatory and motivational roles of reasons, both as elements of the analysis and as broader programmatic considerations in understanding the extension of being a reason. I agree with Dancy that presenting reasons as either justificatory or motivating is just a matter of emphasis and attention; however, for the purpose of presenting an analysis of being a reason, I think it useful to keep both roles in the foreground. Further, given the existence of philosophers who view the two functional roles as separable, it is important to explicitly argue against programmatic separation of justification and motivation.

In section 2.5, I argue explicitly that justifications without motivations are not reasons; if someone splits off the justificatory role then the extension of justifications will over-generate to include pretexts (as I present them in that section). In section 4.3, I argue that the motivational role without an explicit justificatory role, even with constraints that may allow for the

conditions of justification to be folded in, will result in a practical deficiency in our theory of reasons. This argument pushes towards the claim made by Dancy, and the central contention of the dissertation; reasons are both justificatory and motivational.

In chapter 6, I also make use of parts of the justificatory and motivational roles to help address objections to constructivism generally. This illustrates that taking the two together also provides a better account of the hard cases for constructivist theories of reasons in the existing literature.

1.5 Argumentative methodology throughout the dissertation

This dissertation is methodologically pluralist in its approach to argumentation; that is, the dissertation approaches arguing for the analysis of reasons and various subordinate these, through a variety of different approaches.

The first approach to argument, and the most conventionally accepted approach in my present departmental and professional context, is to argue based on the extension of my analysis and its competitors. The argument goes, “analysis A holds that some fact F is a reason, but F is not a reason; therefore A fails as an analysis of being a reason.” This concerns the extensional adequacy of the theory, which is a minimum condition for the success of an analysis. For an analysis to be extensionally adequate, it must pick out all and only the things it purports to analyze. An analysis of reasons should pick out all and only reasons when applied, and some of the available

analyses do not do so. As a result, those analyses fail (or are at least inferior to analyses that get the extension right).³¹

The second approach to argument is driven by the pragmatic considerations. Analyses, like scientific theories, can be evaluated in terms of what they can explain, the detail of the explanations that they offer, and the potential applications. A major point of argument, perhaps the master argument of the dissertation (as it is built in to the structure of the dissertation), is that my analysis is much better than available alternatives in terms of its practical value. Chapter 5 is dedicated entirely to exploring the explanatory upshots of this theory of reasons for issues in the social domain; later sections of chapter 3, including sections 3.3.5 and 3.4.3, provide an illustration of how my analysis' more fine-grained approach to attitude-dependence (focusing on valuing, rather than attitudes generally) provides richer explanation of the motivational role.

There are periods where the argumentative course of this dissertation may look like it advances only one or the other of these. For example, during section 4.4.1, it may seem that my own analysis is concerned solely with the rational reconstruction of the extension of being a reason, and that the principal argument is that my analysis gets the extension right. Neither of these arguments is meant to be treated independently; both are substantial advantages to my analysis, and should be treated as interdependent.

³¹My own approach involves focusing on easy case judgments. Work on conceptual analysis often involves the recognition that some cases may produce reasonable disagreement or be otherwise unclear. As such, the easiest place to start is with cases that are subject of broad agreement.

1.6 Outline of the dissertation

The first chapter, the chapter that you are currently reading, provides some notes on basic philosophical commitments of the dissertation. In section 1.1, I have provided the analysis itself and throughout the rest of chapter 1 provided relevant background.

The analysis has two conjuncts, two parts. The first part is the claim that being a reason entails being evidence that one ought to ϕ , what I call “the reasons as evidence condition” throughout the dissertation. I expound and defend this claim in chapter 2. The chapter focuses on arguments in favor of the reasons as evidence thesis, both as a necessary condition (per my analysis) and as a biconditional (in analyses developed by [Kearns and Star \(2008\)](#)). I argue that being evidence is necessary for being a reason in section 2.2, on the basis that being evidence one ought to ϕ best captures the justificatory role of reasons.

I argue that the biconditional analysis fails in section 2.5, because a fact can be evidence that S ought to ϕ without being a reason for S to ϕ . This argument, like many of the arguments in this dissertation, is driven by the presentation of cases that cut against an existing analysis.

The second conjunct of the analysis is that being a reason entails promoting or protecting something of value, which I refer to as “the promoting and protecting condition” throughout the dissertation. I expound and defend this part of the analysis in chapter 3. I also argue that being of value requires a particular relation to the evaluative attitudes of agents in section 3.4, and further that if values are attitude dependent then my analysis entails that reasons will also be attitude dependent.

The chapter argues that attitude-dependent analyses of values are correct (in section 3.4.2), but that such analyses also add to a theory of reasons by explaining the motivational role of reasons. In contrast to some alternative attitude-dependent theories of reasons, my analysis spends time developing the particulars of the sort of attitude which gives rise to a thing being of value, and holds that these particulars of our attitudes are important to understanding the complexity of our motivations. Section 3.4.3 argues that the motivational roles of reasons are best explained by these attitude-dependent analyses of values and by focus on a particular, rich account of valuing.

These two chapters make up the defense of each of the conjuncts in the analysis of reasons.

In chapter 4, I consider the two conjuncts of the analysis taken together. I argue that the analysis gets the extension of being a reason right, by virtue of avoiding the prospective over-generation of competing theories of reasons. I consider the prospect of separating the two conditions, and focus on the prospect that the promoting and protecting condition might entail the reasons as evidence condition, such that the separate reasons as evidence condition is unnecessary in section 4.4, however I argue that there is good independent reason for adopting the conditions as independently necessary and jointly sufficient. This argument considers a few potential corner cases the analysis may get right (which other theories neglect), in section 4.3, but the argument primarily focuses on the explanatory upshot of the analysis of being a reason which uses the conjunction. The explanatory upshot argument is initially developed in section 4.5 and it is the subject of chapter 5.

Chapter 5 advances a further practical advantage of the analysis and a major argument in favor of the analysis. The chapter shows how coupling the two forms of analysis with standard claims in social ontology can produce a robust account of the social roles played by reasons. This includes explanation of the social dimensions that are often discussed in the social sciences (sections 5.2.1 and 5.4.1) and the normative dimensions of philosophy of law and ethics (section 5.4.2). This practical advantage is a major argument in favor of my analysis of reasons.

Chapter 6 considers some prospective counterexamples to and worries about my proposed theory. Because the theory I am presenting here is novel, these prospective counterexamples are adapted from broad objections to constructivism about reasons. I explained my commitment to constructivism about reasons in section 1.3 and its influence is pervasive, especially in the discussion of values. There are many objections to constructivism in the literature, and I try to take my time to develop and respond to the strongest and most representative counterexamples.

Chapter 6 provides an indirect argument for constructivists to accept my analysis of reasons; my analysis provides tools that more directly and effectively address the most significant and enduring examples to constructivism. At the very least, that chapter provides a range of tools for constructivists with competing theories to address the counterexamples. The advantages of that defense of the theory subsist in the strength of the analysis, that the analysis is especially fine-grained in its handling of the roles of reasons, and gives some gritty detail about the process by which facts come to occupy the roles of reasons.

Chapter 2

The reasons as evidence condition

2.1 Reasons as evidence, “counting in favor”

You know my method. It is
founded upon the observation of
trifles.

Sir Arthur Conan Doyle, 1892,
*The Adventures of Sherlock
Holmes*, “The Boscombe Valley
Mystery” p. 214

A platitude in characterizing reasons holds “facts give us reasons when they count in favour of our having some belief or desire, or acting in some way.” (Parfit, 2011, pp. 1) In some cases, this is used as an analysis of reasons, that to be a reason *just is* to count in favor (Scanlon, 2014, pp. 30); more precisely, to be a reason to ϕ just is to count in favor of ϕ -ing.

“Counting in favor” is imprecise; it uses accounting as a metaphor. Agents do the counting and accounting; facts do not count. When one says a reason R “counts in favor” of an act ϕ , this means R is (or should be under certain specified conditions) counted in favor by some agent or group of agents, under some specified circumstances. Talk about facts “counting in favor” is misleading; the facts are the sorts of thing we are disposed to count in favor, or perhaps believe we ought to count in favor, of ϕ -ing. When deliberating about whether to ϕ , we are inclined to count some facts in favor of ϕ -ing. But the facts themselves are things counted; they don’t themselves count.

Section 1.4.1 introduced the justificatory role of reasons. A fact plays the justificatory role when that fact justifies *pro tanto* S ϕ -ing in C. When I

consider my second cup of coffee, considering the amount of work I need to get done, my current resting heart rate, whether I have eaten breakfast already, and other such facts play a role in justifying drinking (or refraining from drinking) that second cup of coffee.¹ On the account of reasons as “facts that count in favor,” the justificatory role illustrates that these facts are reasons to drink (or to refrain from drinking) that second cup of coffee.

Some fact is a reason in part because that fact can play a role in justification for acting in some way or deliberation about whether to act in some way. This is an attempt to be more precise, but preserve the general point of “counting in favor” articulated above. The facts themselves do not count in favor.

Some philosophers (including Parfit, Scanlon, and Kearns and Star) hold that to be a reason *just is* to play this role in justification. Kearns and Star, for example, provide the biconditional claim that some fact F is a reason for S to ϕ *if and only if* F is evidence that S ought to ϕ . I argue that Kearns and Star’s biconditional analysis fails, but that the justificatory role is a necessary condition. So, one part of the biconditional (the necessary condition) is true, but not the other (the sufficient condition).

I argue that this biconditional analysis fails as an analysis of reasons. However it succeeds in providing a tractable, fine-grained, and accurate account of what Scanlon and Parfit are after with “counting in favor” by explaining how the fact justifies acting.

Table 2.1 provides a survey of the cases discussed in this chapter (the left column) and whether each of the theories of reasons considered in this

¹This only applies to the second cup. The first cup is frequently involuntary. Some behaviors are not reason responsive, but rather responsive to some involuntary bodily inclination.

Table 2.1: Extensions of cases, Chapter 2

Cases	My Theory	Kearns and Star	Right-making
Ordering In (Toy Case)	$\exists R$	$\exists R$	$\exists R$
Ananya's Fire Alarm	$\exists R$	$\exists R$	$\neg\exists R$
Timmy's False Alarm	$\exists R$	$\exists R$	$\neg\exists R$
<i>Paris, Texas</i> (Weightlessness)	$\exists R$	$\exists R$	$\exists R$
Pretext Cases	$\neg\exists R$	$\exists R$	$\exists R$
Ersatz Reasons Cases	$\neg\exists R$	$\exists R$	$\exists R$

chapter (my own, the biconditional advanced by Kearns and Star, and right-making theories) holds that there is a reason (marked by $\exists R$) or not ($\neg\exists R$) in each case. The first few cases (the case of whether there is a reason to order takeout and Ananya's fire alarm) are meant to provide illustrations of the reasons as evidence thesis and how the analysis works. The *Paris, Texas* case clarifies a technical issue in the discussion of Kearns and Star's theory, some confusion in the literature over *pro tanto* reasons; my treatment of the case in section 2.2.1 illustrates the resilience of Kearns and Stars analysis against these putative counterexample.

The mere pretexts² and ersatz reasons³ cases make up most of the end of this chapter, as they are my substantive argument that the extension of Kearns and Stars biconditional analysis of reasons as evidence is mistaken. The argument holds that taking the reasons as evidence thesis as a biconditional results in claiming certain facts are reasons when they just are not. That is, by having an extension that includes the pretext reasons cases, the biconditional approach overgenerates the extension of being a reason by including pretexts.

²See section 2.5.

³See section 2.6.

By contrast, my analysis is more restrictive and the other necessary condition (to be fleshed out in section 3.1) does not include the offending extensions.⁴ We get to keep the useful account of the justificatory role laid out by Kearns and Star, while avoiding these counterexamples. In sections 4.5 and 5.1, I argue that the reasons as evidence thesis *as a necessary condition* provides significant value to the explanation of reasons. While holding that being evidence is necessary for being a reason fails to provide a comprehensive analysis of reasons, that condition provides useful insights and tools for theories of reasons.

As I revise this chapter, it is worth pointing out the significance of mere pretexts in American political contexts. Daunte Wright was killed after a pretextual stop “justified” by an air freshener, as there is a legal prohibition on hanging anything from the rear view mirror that obscures a driver’s view. Systematic racial bias in policing often hinges on the recognition of certain pretexts which are disproportionately applied to Black people. When I say in the later sections that treating pretexts as reasons is not merely an intuitive and conceptual mistake, but also a serious practical mistake, this is an especially consequential, moral, and political point in our present context. To treat pretexts as reasons, based on their putative role as justificatory without consideration of whether the pretexts in fact play a motivational role is not merely a matter of conceptual disagreement, but a matter with some significant stakes. The fact that the conceptual issue is somewhat

⁴It is unclear whether, on the view developed by Kearns and Star, all facts that are right-makers are also evidence. If it is true that all right-makers are evidence (but not the converse; the converse is certainly false), then Kearns and Star have a more extensionally permissive view than theories that take reasons as right-makers. As we can see from the fire alarm example, this is correct.

baked into American jurisprudence following *Whren v. United States* is a serious moral and political deficiency.

2.1.1 Fitting this chapter into the broader project

As discussed in section 1.1.5, there are parts of the argument where it may not be obvious how the moves in the dissertation fit together, and so it is useful to provide some primer on the arguments in this chapter and how they serve the dissertation as a whole.

The most explicit goal of this chapter is to expound and defend the first condition in the analysis, the reasons as evidence condition: If some fact F is a reason for S to ϕ in C , then F is evidence that S ought to ϕ in C . This is presented in the deliberation argument, adapted from Kearns and Star, and advanced in section 2.4.

- If F is a practical reason for S to ϕ in C , then F can play a role in reliable deliberation about whether S ought to ϕ in C .
- If F is not evidence that S ought to ϕ in C , then F cannot play a reliable role in deliberation about whether S ought to ϕ in C .
- \therefore If F is a practical reason for S to ϕ in C , then F is evidence S ought to ϕ in C .

There is quite a bit going on in this argument, but my view (shared by Kearns and Star) is that this establishes that some fact being evidence is necessary for that fact being a reason. This is the first condition of my theory of reasons.

Contra Kearns and Star, I do not believe being evidence is *sufficient*. This is also a requirement of my theory of being a reason, as the reasons

as evidence condition must be only necessary and not sufficient (or else my theory is false). I defend this position in section 2.5 by counterexample. There are two sets of counterexamples.

- **Mere pretexts** are cases where the fact F justifies S ϕ -ing in C , but would not (given the contents of S 's motivation set) motivate S to ϕ in C . These cases would satisfy the sufficient condition of being evidence S ought to ϕ in C , but mere pretexts are not reasons.
- **Ersatz reasons** are cases where the fact F makes it true that S ought to ϕ in C , but cannot be known by S (or, on some stipulations, by *any* agent). On some views of evidence, ersatz reasons will satisfy the sufficient condition of being evidence that S ought to ϕ , but ersatz reasons are not reasons.

The explicit purpose of this is to establish:

- Being evidence is necessary for being a reason ($E \Rightarrow R$).
- *It is false* that being evidence is sufficient for being a reason ($\neg (R \Rightarrow E)$).

These claims are fairly explicit through the development of the chapter. The relationship of these arguments to the justificatory role of reasons, as understood in functionalist terms in section 1.4.1. The purpose of the deliberative argument borrowed from Kearns and Star, and in circumscribing mere pretexts and ersatz reasons out of the extension of being a reason is to show how it is that being a reason entails the deliberative role. The master argument is structured approximately as follows.

- In order to account for the justificatory role of being a reason, a theory of reasons must show how a reason F for S to ϕ in C would matter to deliberation about whether S ought to ϕ in C .
- *Contra* some alternative theories (e.g. the reasons as right-makers approach discussed in section 2.2.3), the reasons as evidence condition ($E \Rightarrow R$) shows how a reason F for S to ϕ in C would matter to deliberation. See, especially, section 2.4.
- $\therefore E \Rightarrow R$ satisfies a necessary condition for a successful theory of being a reason.

That is the major purpose of this chapter; it works to illustrate how the reasons as evidence condition explains the justificatory role of reasons. Various subsections of the chapter, like the discussion of followability and ersatz reasons in section 2.6, serve to illustrate points of explanatory strength in the reasons as evidence condition, and thereby strengthen the case for using reasons as evidence to explain the justificatory role.

2.2 Reasons as evidence, unpacking the view

The canonical formulation of the reasons as evidence thesis, developed by Kearns and Star, holds that being a reason is biconditional with evidence for a certain kind of proposition. That kind of proposition includes the subject, deontic modal operator, and the act type for which the fact is a reason.

Necessarily, a fact F is a reason for an agent A to ϕ *iff* F is evidence that A ought to ϕ (where ϕ is either a belief or an action). (Kearns and Star, 2009, pp. 216)

The fact that it is snowing and -40°C outside is reason for me to order food in rather than going out. The fact that it is snowing and -40°C outside is evidence I ought to order delivery rather than going out. On the biconditional theory advocated by Kearns and Star, the weather conditions are a reason if and only if the weather conditions are evidence I ought to order delivery. For a shorthand, from this point on I refer to this biconditional analysis with “ $R \iff E$ ”, primarily to distinguish it from the individual conditionals.

The weight and strength of the ought may vary subject to context. The force of “I ought to eat breakfast” is different than the force of “I ought to call my grandmother,” which is different from “I ought to save the drowning child.” Even if we hold that “ought” itself is singular and it is only the domains of reasons (practical, prudential, moral, etc.) that differ, then this still leaves us with space for varying the character of the ought subject to considerations of context.⁵ The practical reason for drinking coffee results in a weaker ought than the moral reason to aid a person in danger. All of these work as evidence, as justifying action.

2.2.1 Contributory and *pro tanto* reasons

Kearns and Star refer to reasons without adding any modification. $R \iff E$, as they present it, does not specify whether it is discussing *pro tanto*, contributory, *pro toto*, all-things-considered, or other varieties of reasons.

⁵Gauthier (1963) argues that there is a single character of ought which is influenced by reasons. I have a moral reason to call my grandmother and a prudential reason to call my grandmother, and they both are evidence for the same proposition: that I ought to call my grandmother. They are not evidence for separate propositions where one ought is moral and another ought is prudential. This is a separate axis from the point here, which is that the oughts in “ought to paint” and “ought to save a life” are different in certain respects. Further, moral considerations are stronger, for example, than practical ones in forcing that ought.

Unfortunately, this leaves the analysis open to some worries about whether there ought to be some modification. Let us consider some evidence. (Par-don me.)

Kearns and Star (2008) includes a summary passage, “reasons *contribute* to all-things-considered oughts, and normative reasons are essential *inputs* to successful processes of practical deliberation.” (Kearns and Star, 2008, Conclusion, emphasis in original) This indicates that reasons are not the justification for an all-things-considered judgment, but rather contribute to the formulation of such an all-things-considered judgment.

Throughout this dissertation, where the term “reasons” appears without modification, it refers to *pro tanto* reasons. As far as I can tell, this is the same idea that Kearns and Star have in mind in discussing that reasons *contribute* to practical deliberation. Let us consider one objection raised to Kearns and Star’s reasons as evidence thesis, and how specifying that it is a theory of *pro tanto* reasons defuses the objection.

2.2.2 Weightlessness “counterexamples” and *pro tanto* reasons

One alleged counterexample to $R \iff E$ is that some fact is only evidence for a proposition when the matter is undecided but a reason for S to ϕ obtains regardless of whether there is a decisive judgment that S ought to ϕ . Therefore, there are cases where F is not evidence that S ought to ϕ (because

the issue is decided), but F is still a reason for S to ϕ (because being decided does not dissolve the reason).⁶

In this class of counterexamples: a new fact is introduced after an issue is decided and that fact has no influence in justifying. It is “weightless” because the issue is already settled. Consider its application in a case.

That *Paris, Texas* is showing is a reason for Miyuki to go to the cinema. However, the evidence that Miyuki ought not to go to the cinema is conclusive. In that case, the evidence that she ought to go to the cinema is just misleading; Miyuki need accord it no weight in any further deliberation. From $R = EO$ [Kearns and Star’s $R \iff E$] it seems to follow that she need accord the fact that *Paris, Texas* is showing no weight as a reason for going to the cinema in any further deliberation. But this is wrong. In Gregory’s words, that *Paris, Texas* is showing is not a mere ‘red herring’; it really does count in favour of going to the cinema. (Whiting, 2018, pp. 2201)

There is a difference between “counting in favour” (on one hand) and having weight in deliberation or justifying (on the other hand). This objection to the reasons as evidence thesis rests on a straightforward mistake: misunderstanding that $R \iff E$ is an analysis of *pro tanto* reasons, or reasons which contribute, rather than *pro toto* or all-things-considered reasons.

Kearns and Star propose various technical solutions to this objection; I think the core response is simple. The *pro tanto* reason applies only insofar

⁶This objection is put forward as a challenge by Gregory (2016). A more in-depth discussion of a range of similar counter-examples is surveyed and addressed by Kearns and Star (2013).

as it justifies the action relative to other considerations. In a case where the justification is settled, the extent to which the *pro tanto* reason applies (or, to borrow Kearns and Star's phrase, "contributes") is limited by the other considerations. This does not mean the *pro tanto* reason is not evidence, just that it is only evidence on a *pro tanto* basis.

A *pro tanto* reason is evidence that S ought to ϕ absent countervailing considerations. If there are countervailing considerations that defeat the evidence, then this so be it. One can have evidence that p while there is definitive evidence $\neg p$.

There is still a *pro tanto* reason R for S to ϕ in C even if that reason comes about when we are already aware of decisive *pro tanto* reason or *pro toto* reason for S to $\neg\phi$. The fact that, given the already available and decisive evidence, this new evidence does not warrant consideration is irrelevant to whether the evidence exists. Such a case may be considered straightforwardly where evidence is reason for a juridical action, as in a criminal proceeding.

Donald is on trial for offering a *quid pro quo* with Ukraine in a phone call. His schedule holds Donald was golfing at the time. There is a summary of the call provided and verified by Donald and other parties of the call at the time. The call summary is decisive evidence that Donald did crime. The schedule is non-decisive evidence that Donald did not do crime.

There can be evidence for acquittal, even when there is decisive evidence for conviction.⁷ There is evidence that $\neg p$; this does not imply that $\neg p$. Only if there is *decisive* evidence that $\neg p$ might we say this implies $\neg p$.⁸

So, the response here is that the reasons as evidence thesis refers to facts that are part of (i.e. *pro tanto*) and/or contribute to the justification or deliberation.

Further, such a reason having weight in some instance is not a matter of whether it can, in the actual world, play a role in deliberation.⁹ Rather, having weight is a feature of the reason in relation to other sets of facts. The fact that Donald's schedule says he was golfing has weight; that weight is not significant in comparison to the other available evidence. However, there are counterfactual situations, circumstances where the available evidence does not include the decisive evidence, in which Donald's schedule claiming he was golfing *contributes to* or is *part of* the eventual judgment, where it does have weight.

The purpose of addressing this objection was not that the objection itself is interesting. Rather, the objection and response illustrate a shared feature of $R \iff E$ and my own analysis. The object of the analyses are *pro tanto* reasons, rather than the whole of the justification.

⁷Colloquially, "evidence" is sometimes used to refer to objects, rather than facts. A physical copy of the schedule may be "evidence" in the colloquial usage. My usage here (as is the convention in philosophy) is to take evidence as referring to facts.

⁸Even this may be controversial subject to rules of evidence (Kelly, 2016, Secs. 3, 4).

⁹One might hold that being a *pro tanto* reason can be even weaker than this, that there can be a *pro tanto* reason even for necessary falsehoods. Perhaps the fact that there is a prime number between 90 and 100 is evidence that 95 is a prime number, but it is necessarily false that 95 is a prime number. I am skeptical of this for reasons that are unpacked section 4.3, though not directly applied to mathematical cases; it is also possible that this case turns on the metaphysics and epistemology of mathematics which is far afield of this dissertation. See Horsten (2019), especially Secs 3 and 4.

It also establishes a feature of reasons that becomes salient later in this analysis: the ways and extent to which those facts justify action. This includes, but is not limited to, the weight of one fact relative to others.

2.2.3 Reasons as evidence, rather than right-makers

$R \iff E$ holds that reasons are evidence for the truth of a proposition propositions, rather than “right-makers.”¹⁰

A right-maker is just a fact in virtue of which S ought to ϕ in C or would be right to ϕ in C.¹¹ In theory, every right-maker should be evidence that S ought to ϕ , unless there is some reason that right-maker is inaccessible to an agent. (I discuss these potential cases in section 2.6.) Evidence that S ought to ϕ in C is not always a right maker. Some fact may indicate that S ought to ϕ in C without making it the case that S ought to ϕ in C.

$R \iff E$ and Kearns and Star’s writing on the subject clarifies what it means to say that a fact “counts in favor,” when other philosophers (including Parfit and Scanlon) present the platitude. $R \iff E$ also clarifies the distinction between reasons as evidence and reasons as right-makers. Evidence need not make the proposition for which it is evidence true; evidence can just be an indicator that the proposition is true.

¹⁰See: [Kearns and Star \(2013\)](#); [Whiting \(2018\)](#).

¹¹The use of right-making (like truth-making) may be taken in specific views where for some fact F to be a right-maker, it must ground the rightness of S ϕ -ing in C *pro toto*. However, my use here is open-ended to theories that are not so strict, for example theories where some right-maker may be a partial ground such that F makes it right for S to ϕ in C only *pro tanto*. This functions like the views of truth-makers that allow for partial grounding of the proposition, as outlined in ([MacBride, 2020](#), Sec. 1.6). There are a range of theories of right-making, but the argument provided here by way of the fire alarm case applies to all such arguments.

There is a considerable extensional difference between evidence and right-makers. The fact that the fire alarm is going off is reason for Ananya to leave the building. However, the fact that the fire alarm is going off does not make it right for Ananya to leave the building.

The fact that there is a fire or a coordinated fire drill would make it right to leave the building. The fire alarm is evidence because it indicates the facts that make it right to leave the building (Kelly, 2016, Sec. 3).¹² Right making facts may also be evidence; that there is a coordinated fire drill makes it right for Ananya to leave the building, and that fact is evidence and a *pro tanto* reason to leave the building. A fact may be evidence for a proposition without being a right-maker or truth-maker for that proposition.

The check-engine light is an indicator that there may be a problem with the engine. The check-engine light is not a right-maker for the proposition “there is a problem with the engine.” Are these indicators reasons? $R \iff E$ is better than the reasons as right-makers thesis here. An alarm can be a reason, just as smoke can be a reason to believe there is fire.

2.2.4 Nature of evidence

Following Kelly: “Evidence, whatever else it is, is the kind of thing which can make a difference to what one is justified in believing or (what is often, but not always, taken to be the same thing) what it is reasonable for one to believe.” (Kelly, 2016, Sec. 1)

There are a range of subtle differences between the claim “evidence that *p* makes it obligatory (absent defeaters) to believe that *p*” and “evidence

¹²I do consider the fallible nature of this indicator at greater length in section 4.2, along with potential variation in readings according to my theory of reasons.

that p makes it reasonable to believe that p” and “evidence that p makes it permissible to believe that p.”¹³ The particular question of what evidence *does* admits of a general characterization, that evidence stands in relation to a proposition such that it has normative implications for an agent believing that proposition, that it justifies S believing that p.

Consider this in the context of $R \iff E$.

A fact F is reason for S to ϕ in C. Does F make it *permissible* for an agent to believe that S ought to ϕ in C? Does F make it *obligatory* (absent defeaters) for an agent to believe that S ought to ϕ in C? Those are radically different things.

This issue cannot be resolved without significant work in epistemology and is an ongoing debate in the literature. I cannot get into the nitty-gritty of it here, and the disagreement would take us into a significantly different philosophical domain. I am of the view that this disagreement should inform my theory of reasons, as well as Kearns and Star’s $R \iff E$, as it does meaningfully change what the existence of a reason entails.

This disagreement about the nature of evidence, and how justification of belief works, does not change the deontic modal operator in the reasons as evidence thesis. Rather, the distinction between justification as (on the one hand) permission or (on the other hand) obligation adds an additional deontic modal operator to the left. If evidence for a belief creates a permission, then the fact F makes S permitted to believe that S ought to ϕ in C. If evidence for a belief creates an obligation, then the fact F makes it such that S ought to believe that S ought to ϕ . The existing debate about the nature of evidence changes the first of the two modal operators, but not the second.

¹³See: [Kelly \(2014\)](#); [White \(2014\)](#)

The second modal operator is what matters to Kearns and Star's analysis and to my analysis.

That the fire alarm is going off justifies Ananya's belief that she ought to leave the building. The difference is in how "justifying Ananya's belief" is cashed out in modal terms.

Justifying makes permissible If justifying makes believing *permissible*, then the fire alarm makes it the case that Ananya is *permitted to believe* that Ananya ought to leave the building *ceteris paribus*.

Justifying makes obligatory If justifying makes believing *obligatory*, then the fire alarm makes it the case that Ananya *ought to believe* she ought to leave the building *ceteris paribus*.

The first deontic modal operator changes; the second does not. As this is not a dissertation on epistemology, I want to bracket that question as a substantial one for epistemology with the acknowledgement that there is a difference between "Ananya *ought to believe p*" and "Ananya *is permitted to believe p*."

2.3 Separating the two halves of the biconditional

$R \iff E$ is a biconditional relation. R is a reason for S to ϕ in C *if and only if* R is evidence that S ought to ϕ in C .

Kearns and Star note that because $R \iff E$ is a biconditional, one might treat each of the two conditionals independently (Kearns and Star, 2009, p.

224). Because I think one of the two conditionals is false, this is useful for my analysis.

Consider the following conditionals.

A reason is evidence ($R \Rightarrow E$) If the fact F is a reason for S to ϕ in C, then F is evidence that S ought to ϕ in C.

Evidence is a reason ($E \Rightarrow R$) If the fact F is evidence that S ought to ϕ in C, then F is a reason for S to ϕ in C.

Following my previous use of the biconditional, I will also use the $R \Rightarrow E$ and $E \Rightarrow R$ as shorthand.

On my analysis, every reason is evidence but not every piece of evidence is a reason. I make use of Kearns and Star and $R \iff E$ partly because Kearns and Star work to establish each of these two claims independently. I think their argument for $R \Rightarrow E$ is successful and I make use of that argument below. In my own analysis, I formulate $R \Rightarrow E$ as justificatory condition: some fact F is a reason for S to ϕ in C *only if* F is evidence that S ought to ϕ in C.

I argue that $E \Rightarrow R$ fails and, therefore, so does the biconditional. I advance my argument to that effect in section 2.5. For now, let us bracket the disagreement and focus on the agreement and value of $R \Rightarrow E$.

Section 2.4 considers Kearns and Star's deliberation argument in favor of $R \Rightarrow E$. The thesis and the deliberation argument play a significant role in my analysis of reasons, especially in fleshing out the justificatory role of reasons.

I am motivated to make use of Kearns and Star because of the success of their analysis in fleshing out "counting in favor" and the wide ranging use

of “counting in favor” in the reasons in literature. They provide a successful, more fine-grained analysis of what it means for a fact to count in favor of acting and keeping that analysis is useful for ensuring continuity with existing analyses of reasons as justification.

2.4 $R \Rightarrow E$, the deliberation argument

Kearns and Star argue independently for each of the two claims in the biconditional: $R \Rightarrow E$ and $E \Rightarrow R$. In arguing for $R \Rightarrow E$, they advance what they call “the deliberation argument” (Kearns and Star, 2009, p. 225). Their argument goes (and I quote directly) as follows.

1. Practical reasons to ϕ can play an important role in reliable practical reasoning concerning whether or not one ought to ϕ .
2. If practical reasons to ϕ are not evidence that one ought to ϕ , then they cannot play this important role in reliable practical reasoning.
3. Therefore, reasons to ϕ are evidence that one ought to ϕ .

This argument establishes that $R \Rightarrow E$ is true, and this deliberation argument is consistent with my own motivation for holding that $R \Rightarrow E$, with a few qualifications and clarifications.

The first clarification has to do with their use of the term “deliberation” in naming the argument. “Deliberation” is replaced by “reliable practical reasoning” in the argument, and Kearns and Star are right to make this change. Generally speaking, deliberation is regarded as a forward looking exercise about whether (at some point in the future) one ought to ϕ , but

practical reasoning is not always forward looking. Rather than talking about “deliberation,” my preference is to talk about “justification” and allow that this includes both the forward looking consideration of what one ought to do and the backward looking consideration of whether one acted rightly. This is perfectly compatible with Kearns and Star’s argument, as they are focused on “reliable practical reasoning” regardless of whether the action lies in the future or the past.

There are a handful of subtly different verbs one might employ which all broadly converge around the process of S deciding whether to ϕ . Deliberate, evaluate, reason, and justify all might be used here. While Kearns and Star prefer to talk about “reliable practical reasoning,” I prefer to talk about justification for reasons provided in the preceding section and to draw attention to the justificatory role of reasons.

A second clarification is that there is some content missing in Kearns and Star’s account. Most significantly, it is missing a quantifier in the first premise. An uncontroversial interpretation of that premise would hold “[some] practical reasons to ϕ can play an important role...” While it is no doubt true that some practical reasons can play such a role (as a matter of fact, many do actually play such a role), this would not produce the conclusion. Rather, what Kearns and Star mean to say is “[all] practical reasons to ϕ can play an important role...”

Set aside that first premise for a moment. The second premise is constructed to be uncontroversial, given the role of evidence as justificatory and the idea that such deliberation is about justifying a belief about what one ought to do. If some fact being evidence is characterized by its role in justifying belief, then that second premise should be straightforward. If all

practical reasons can play a role in deliberation, then they are justifying the proposition about what the agent ought to do. The potential disagreement regarding weight of the deliberation argument rests on the first premise.

The modality of “can” gets a bit tricky here; the modality of “can” is a point of subtle divergence between my view and Kearns and Star’s. For me, “can” is not a matter of logical possibility, but of proximate practical and psychological possibility. A fact that raises the likelihood of p being true is not a reason if there are no agents in the world where that fact occurs who could grok, understand, or consider that fact.

S could ϕ is not just a property of ϕ as a possible action, nor is it just a property of S . Rather, it is a property of S ’s situation in the world, including circumstances and dispositional properties. This is one of the lessons in recognizing functional properties. That an object can create drag under conditions of air flow is not a property of the object or its geometry; being an airfoil also requires some background facts about the world, the way that airflow works in that world. Circumstances matter to whether S can ϕ .

Kearns and Star seem to understand “can play an important role” as a property of the fact if there are some possible worlds in which the fact is present and agents in that world could use the fact in deliberation. I am narrow-minded. If there is a fact F in a world w where there are no agents who can use F in deliberation, then it is wrong to say that F can be used in deliberation. This is partly a development of arguments about the attitude-dependence of reasons, and whether reasons that can be taken up by no agents like us (in the broad sense of having our sensory apparatus, or even being the products of evolution) can play an important role in deliberation.

My view, following the debunking arguments I take up in section 3.4.2, is that these are neither evidence nor reasons.

Suppose that there are some cases at the intersection of *conceptually possible* to play a role in practical reasoning and *practically impossible* to play a role in practical reasoning. For example, suppose that F could be evidence that S ought to ϕ in C, but it is practically impossible for any agent to know that F. On such a case, one might say that it is conceptually possible that F is evidence S ought to ϕ in C, but it is not practically possible for F to play a role in reliable practical reasoning, since no agent could know that F.

I am skeptical that there are cases that can satisfy this schema; the reason I do not provide a clear and concrete example here is just because I cannot think of any that seem compelling. Such a case would have to establish that there is a logically possible world in which S can ϕ , but there is a practical limitation on the S-type agents such that no S in that world could ϕ . I do not take my own intuition on this subject to be decisive; if there are such cases, then that would be an interesting wrinkle.

A reason that satisfies the above schema would have to split across modalities: it would have to be logically possible for S to believe that F such that F would be evidence for S believing that p *and* it would have to be practically impossible for S to believe that F such that F would be evidence for S believing that p. The fact has to exist in the penumbra between being logically and practically accessible. I am not convinced that this can be sensibly discussed, because it would require radically varying the psychology of the agent S once we leave the domain of the practical. If we vary the S-type agent too much, in order to accommodate that logical-but-not-practical possibility, then we lose the relevant features of the S-type agent and we

wind up talking about radically different types of agents, who would have radically different cognition and epistemic access. That's a sort of logical consideration that, by its nature, makes it difficult to apply to agents like us.

It is entirely possible to read the modality of "can" in either of these ways, with little difference in the cases except those weird possibilities discussed above. There is a potential point of agreement with Kearns and Star that I will develop in section 2.6, which I take a corollary of their deliberation argument. We seem to agree that being evidence is a function of possible use in deliberation.

2.5 Mere pretext arguments against $E \Rightarrow R$

This argument proceeds from a worry about mere pretexts for action and to the conclusion that the second part of the biconditional, $E \Rightarrow R$, is false. Mere pretexts are facts provided as justification for an agent to act, even though those facts would not motivate that agent to act. The argument maintains that the reasons as evidence thesis does not distinguish between mere pretexts and facts that justify acting and motivate agents to act. Failing to distinguish between mere pretexts and facts that can play the motivational role results in serious practical mistakes.

Suppose Officer Smith is running a speeding trap. He sees Tim, who he believes to be under the influence of drugs, coming through the trap.¹⁴

¹⁴In introducing this chapter, I briefly mention *Whren* as an illustration of the legal acceptance of this principle. As a matter of fact, as long as the pretextual stop is in advancement of "some additional law enforcement objective" (Scalia's phrase in the decision) then it is regarded as legal. This fails to distinguish the conceptual problems of pretexts adequately, and has caused serious problems in addressing violations of protections against search and seizure. Notable, in *Whren*, the justification for the stop was regarded as legitimate, as was the secondary law enforcement purpose (drug enforcement), but the questions

2. The reasons as evidence condition

Tim is going slightly over the speed limit and is otherwise driving normally, non-erratically. With any other driver, Officer Smith would not be inclined to make the stop. Because he regards Tim as a drug user, Officer Smith is motivated to stop Tim. If a judge in traffic court questions Officer Smith about why he made the stop, Officer Smith would appeal to the slight speeding. The slight speeding is a pretext, which justifies his behavior but does not motivate it.

Colloquially, referring to some fact as a mere pretext (or even just a pretext) is to cast aspersions on its legitimacy as a justification, to imply that it is spurious. But a pretext need not be a spurious justification. In this case, the speeding is regarded by the legal order as a legitimate justification for making a stop.

There is often a distinction made between the reasons *there are* for an agent to act and reasons an agent *has* for acting. Clearly, one challenge is that the mere pretexts are not reasons the agent *has*, but there remains an issue of whether the pretext is a reason for action. The purpose of the moves below is to establish that because (*ex hypothesi*) the agent would never have these reasons, it is a mistake to talk about them as reasons *there are* for that agent to act.

Pretextual police stops are regarded as illegitimate in jurisprudence because they obscure the actual motivations of the officer and (in many circumstances) the pretexts obscure motivations which are prejudicial. If Officer Smith is motivated to stop Tim because Tim is Black and uses the

of whether either such justification was appropriately motivating remains. Notably, the decision in *Whren* requires only consideration of the *justification* for the stop, and not whether the stop was in fact prejudicially motivated. This rests on an interpretation of the Fourth Amendment where “reasonable” means merely justified, rather than justified and motivated.

slight speeding as a pretext, then this is a serious problem. Officers making stops that might be justified based on pretext, but are in fact *motivated* by (for example) race are regarded as illegitimate because officers are acting in an official capacity. The officer is responding to reasons that they should not. To give a pretext is to give a reason that is acceptable to the social standards, when in fact the reason to which the agent was actually responding was illegitimate.¹⁵

The ordinary language notion of pretexts does not have a regimentation. The regimentation offered here is focused on advancing the argument and distinction between mere pretexts and reasons, rather than capturing the ordinary language concept. A fact F is a pretext for an agent S to do an act ϕ in a circumstance C iff F is evidence that there S *pro tanto* ought to ϕ in C and F would not motivate S to ϕ in C and there is some non-identical fact G that would motivate S to ϕ in C .

On this regimentation, mere pretexts are always reasons on the reasons as evidence thesis and never reasons on my analysis.

On $R \iff E$, being evidence S ought to ϕ in C is sufficient for being a reason. The mere pretext cases provide an objection to $E \Rightarrow R$ by presenting a counterexample. This is sufficient to undermine $R \iff E$ as well.

The fact that Tim is speeding is evidence that Officer Smith ought to stop Tim. We know that, under normal conditions, the slight speeding would not be sufficient to render a *pro toto* reason for Officer Smith to stop Tim, but it is a *pro tanto* justification, a fact that contributes to determining (under some circumstances) that Officer Smith ought to stop Tim.

¹⁵The notion of an agent *having* (or being motivated by) certain reasons having an impact on the evaluation of the permissibility of an act is a familiar one. I illustrate this at length in [Stein \(2019\)](#).

But we also know that the speeding would not have motivated Smith to make the stop. Not only did Smith not have F among his reasons, but Smith would never have F among his reasons because it would never be motivating. It fails to satisfy the motivational role (which is independently satisfied by Officer Smith's other beliefs). We can flesh this example out either in terms of appeal to Officer Smith's psychology (that he would not be motivated to make a stop, absent other facts) or counterfactually (that he would not make the stop absent other conditions). This ties in to the role of reasons as psychological explanations addressed at greater length in chapters 3 to 5.

On my theory of reasons, mere pretexts will never be reasons because they fail to satisfy the motivational role. Being evidence that S ought to ϕ in C is a *necessary* condition for being a reason, but it is not sufficient

It is outside of the scope of this section to get more precise nature about the relation between the fact and the motivation set, as argued by Williams (1979); I discuss this directly in section 3.2. Further, there is some question about distinguishing actual motivation and those which may be adopted through "a sound deliberative process" (Williams, 1995, pp. 35). To bracket this point, the present discussion uses "could not motivate" as a gesture to the general motivational role.

To say F *would not* motivate S to ϕ in C is to appeal to both a contingent psychological fact and to apply the practical sense of modality of can. What can and cannot motivate an agent is a matter of the psychology of that agent; as such, what can motivate S is contingent on certain facts about the psychology of S. To say that F *would not* motivate S is to bracket the

issue of modality: Is it conceptually impossible for F to motivate S? Merely practically impossible?

Talking about Smith's motivations in terms of logical possibility is a bit strange. After all, what would motivate S to ϕ depends on the psychology of S and the psychology of S is variable. One might take the "could" in the broadest possible terms, and get basically anything to count on this analysis as long as some logically possible psychological configuration would satisfy the conditions.

The more useful approach is to consider S in terms of the relevant psychological features, and make those part of the type (as discussed in the regimentation). The explanatory value arguments advanced in section 5.4.1 provide illustration of the practical upshot for using the practical modality; the practical value of considering empirical psychology is the strongest (and, in my view, decisive) argument for framing this way.

2.5.1 The Iraq Case; articulating pretexts

The goal of the pretext cases is to separate the two parts of the biconditional set out by Kearns and Star. The first half of the biconditional is the claim that if F is a reason for S to ϕ in C, then F is evidence that S ought to ϕ in C, or $R \Rightarrow E$. I endorse that direction.

The second half of the biconditional is the claim that if F is evidence that S ought to ϕ in C, then F is a reason for S to ϕ in C; or $E \Rightarrow R$. I think $E \Rightarrow R$ is false, and offer the pretext cases as a set of counter-examples. The core idea is that some facts are evidence S ought to ϕ in C, but could not motivate S to ϕ in C. They are pretexts and can be given by S as justifications; in

evaluating whether it was appropriate for S to ϕ in C, they are evidence that S ought to ϕ . Intuitively and on reflective judgment, pretexts (even ones that adequately justify S ϕ -ing) are not reasons.

The argument that pretexts are not reasons is straightforward enough. Treating pretexts as reasons is a practical, explanatory, and sometimes even moral mistake. Consider a case.

Suppose George really wants to bomb Iraq. George is motivated to bomb Iraq because it is in his financial interest to do so. Following a terrorist attack, George is informed that some evidence is provided that the leader of Iraq was involved in said terrorist attack. The evidence of a connection between Iraq and the terrorist attack is evidence that George ought to bomb Iraq. George is not and could not be motivated by the connection to the terrorist attack.

Is the terrorist attack a reason for George to bomb Iraq? On the reasons as evidence thesis, yes; on my analysis, no.

If one followed the approach in [Alvarez \(2017\)](#), then one might distinguish the attack as a justificatory reason but not a motivational reason. On this view, the defense might run that the reasons as evidence thesis is an account of justificatory reason, as distinct from motivational reasons. For some fact E, E may be a justificatory reason and not a motivational reason. I do not want to write off this approach altogether, though given that evidence has to do with justifying (as illustrated in the preceding discussion), this would not be very interesting.

There is some value to focusing entirely on justification. In some areas of epistemology, focusing on the conditions of justification and completely ignoring motivations may be the best way to proceed, especially if the project

is oriented towards finding some limits of idealization. This is not consistent with the goals of my project, because it deprives such a theory of reasons of much of its practical value, which grounds my argument in chapters 4 and 5.

The purpose of rejecting pretexts is to get at an important difference in the extension of being a reason; pretexts are not reasons in the conventional sense and to treat them as such would miss major programmatic goals in metaethics. Focusing solely on the justification in such a case has serious adverse consequences.

2.5.2 Revisiting the external/internal defense

The major potential response a defender might make to block these pretext counterexamples is argue that the pretext cases turn on a slippage between what reasons *there are* for S to ϕ in C and what reasons S *has* to ϕ in C. The response can argue that pretexts are reasons for S to ϕ in C, but they are not reason S *has* to ϕ in C; this explains the duplicity of responding to an inquirer with a mere pretext. “Why did you ϕ in C?” is a response asking for the reasons that you had for ϕ -ing rather than reasons there were which might have justified ϕ -ing. Therefore, the objection to offering pretexts in court or to the American public is that one is lying, by trading on this slippage.

This objection to the pretext argument holds there is a distinction between the external sense of a reason (the reasons there are) and the internal sense of reasons (the reason an agent has) that covers the pretext cases. It

is only the internal sense that requires motivating; the external sense only requires justifying and therefore can include mere pretexts.

This attempts to formalize the separation in justificatory and motivational reasons, as discussed in section 1.4.4.

Being a justification means that the fact *is* a reason for S to ϕ in C. Failing as motivation means that S *does not have* the fact as a reason to ϕ in C. There are a few problems with this line.

The first is that it ignores the historical philosophical arguments against the distinction between internal and external conceptions of reasons (at least, a distinction that would separate external conceptions from the motivational role), like those advanced by Williams (1979).

Finlay and Schroeder (2017) give a survey of reasons to be suspicious of the prospect of purely external reasons, but there is one principally derived from Williams in (Shafer-Landau, 2003, p. 181-182).¹⁶ In order to blame individuals for acting wrongly, it must be possible for the individual to do otherwise. However, individuals are limited by their psychology, including both their knowledge and their desires (broadly understood). If there are genuinely external reasons, then individuals must be blameworthy under some circumstances for not acting in accordance with those reasons (after all, they are not doing what the evidence holds they ought to do). However, such blame is unfair to individuals who are limited by their psychology in ways that result in violations of these external reasons. Therefore, we arrive at a *reductio*.

This is similar to arguments for attitude-dependence of reasons considered in section 3.4.2, but drawn with a normative (rather than descriptive)

¹⁶This is also summarized in (Finlay and Schroeder, 2017, Sec. 2.3).

structure that indicates blaming practices would be wrong by virtue of their unfairness.

The purpose of this argument is to help reestablish the condition that Kearns and Star and I advance in the deliberation argument above: in order for some fact to be a reason, that fact must be the sort of thing that could play a role in deliberation. Put another way, it must be the sort of fact that *can be* a reason for that agent under some circumstances.

The second concern is that it neglects what it means to say F is a reason *for* S to ϕ in C. The idea that F would not motivate S to ϕ in C, based on constraints of S's motivation set, indicates that while F may be a reason for some agent to ϕ in C (subject to motivation set), it is not a reason *for* S.

One might try to get around this response by supposing that what matters is not the motivations instantiated in the state, but some broader set of possible motivations that S might have. Williams himself intimates something like this by talking about motivations an individual might hold given adequate reflection. I think Williams is right to allow some flexibility in motivation set. As I will discuss at length in section 3.2.2, the motivation set is dynamic. But the motivation set of an agent is constrained by some contingent psychological facts about the agent. It is not the case that all metaphysically possible motivations are available to Officer Smith or President George. That would seriously over-generate in the extension of being a reason.

Suppose failing to turn recommendation letters on time has serious adverse consequences for the graduate students on behalf of whom the letters are submitted. Suppose that, while Professor Walker knows this and understands that it is evidence he ought to turn the letters in on time, that he is

not motivated to do so. One might then say: if Professor Walker cared about the well-being of his graduate students, then the facts about the impact of such letters *would be* a reason for him to turn in their recommendation letters on time. If Walker had a different set of psychological states, then Walker would have different motivations. This is trivially true.

These counterfactual conditions are subject to discussion. Since Professor Walker does not value the well-being of his students, that reason may not obtain for him, unless it does so by virtue of some other element in Professor Walker's motivation set. There are conditions where, if Walker (or the world around him) were relevantly different, then the reasons would be relevantly different, but we are dealing with the conditions for some fact being a reason in the world where Professor Walker is who he is and has such-and-such a psychology.

2.5.3 Naïveness objection to pretexts as reasons

The pretext examples provided above carry intuitive force. It seems wrong to say that mere pretexts are reasons. Sometimes seemings are deceptive; it is worthwhile to provide some further argument for why this intuitive force works. It is not just that pretexts do not seem like reasons, but that to treat them as reasons involves making a practical mistake in evaluating the behavior of others. Suppose some fact E is a pretext for S to ϕ in C; if we treat E as a reason for S to ϕ in C, then we will give undue clout to E and be inclined to empathize with S ϕ -ing in C on the [pretextual] justification E when we ought not.

In covering the first impeachment trial of Donald Trump (an event ripe with pretexts), CNN coverage included the claim “Elizabeth Warren’s question may have been part of the reason Lisa Murkowski came to a “no” vote.”

The article refers to a question submitted by Senator Warren (D-MA) regarding the deleterious effect of the failure to call witnesses on the legitimacy of the legal proceedings and the United States Constitution. This was provided as a *pro tanto* reason for Senator Lisa Murkowski (R-AK)’s subsequent vote against calling witnesses in the impeachment trial. Let us strain the imagination with a few stipulations about the incident.

Sen Warren’s question would not and did not motivate Sen Murkowski to vote “no.” Sen Warren’s question is evidence that Sen Murkowski ought to vote “no” on calling witnesses, because of the respect in which a lengthy trial would harm the credibility of involved institutions.¹⁷ What we have here is the failure to accord with anything in Murkowski’s motivation set; the fact was not in the set of reasons Murkowski had for voting no.

One might object to the CNN statement as potentially misleading readers to believe that Murkowski *had* this fact about the reasons for her vote, when she did not. Given that we have noted the relevant implicature (as in the judicial case discussed above) of giving reasons in terms of reasons one *has*, thereby excluding pretexts *inter alia*, this is a reasonable response. Distinguishing the reasons Murkowski had helps to track with the appropriate linguistic convention. Rather, she cited it as justification for a vote that was already determined by independent motivators. The failure of the

¹⁷Let us set aside that there are some circumstances where a fact F can be evidence that S ought to ϕ and that S ought to $\neg\phi$, especially when working in different respects, like the moral and practical dimensions of some deliberation. F may be a practical reason for S to ϕ in C *and* a moral reason for S to $\neg\phi$ in C.

reporting is not limited to its use of a conversational implicature presenting “part of the reason S did ϕ ” in [falsely] implying psychological explanation of action. The worry is deeper than poor communication.

On the reasons as evidence thesis, “F is a reason for Sen Murkowski to vote “no” on hearing witnesses in the impeachment trial” is true and does not provide any insight into Sen Murkowski’s psychology. Rather, it just provides an account of some *pro tanto* justification for Sen Murkowski voting “no.” Running these together is a mistake.

In some contexts, it is useful to consider *only* the justifications for the action. But focusing only on the justification shifts the scope of the explanation and we will inevitably lose the broad explanatory scope that I develop in chapters 4 and 5.

When Alvarez (2017) asks, “Why are you always lying?” to respond with a pretextual answer is to lie. That is, saying “I was lying to spare your feelings” when one is (in fact) motivated to lie out of expedience or some other preference satisfaction, is to be duplicitous. In asking the question, Alvarez is asking for an answer that *both* justifies and motivates, with emphasis added to the motivational (rather than justificatory) role. In any context, providing an answer to even the abstractly framed “Why might S ϕ in C?” that is simply pretextual is a practical mistake. Such a fact is a justification S might offer to avoid or reduce social sanction; the fact is not a reason for S to ϕ in C.

2.6 Ersatz Reasons, a stranger critique of attitude-independence

One question unresolved in the nature of evidence is whether some fact F being evidence that p entails that an individual could believe or know that F . If some fact F is evidence that Donald committed a crime, then does that entail that some individual could know that F ? It is unclear how one might even talk about a fact that cannot (by its nature) be believed.¹⁸ Semantically recognizable entities are usually easy things to believe. There are some potential examples, in cases of fiction and in literature on ineffable (typically religious) experience, but I find those cases create further confusion.

There may be certain facts about the world that we cannot believe, because we are just not psychologically able to form the belief. Can those facts still be reasons for an agent to act? I argue below that being able to believe that F is a minimal condition on F being a reason. As such, this builds in a minimal way in which what can be a reason is dependent on the minds of agents, because it depends on the cognitive limitations of actual agents.

Consider the claim that in order for some fact F to be a reason for S to ϕ in C , it must be possible for S to believe that F . This is a minimal condition and may be controversial for certain realist theories of reasons,¹⁹ though it is

¹⁸There is some complexity around a notion in the pragmatic theory of truth which C.S. Pierce called “buried secrets” (Capps, 2019, Sec 5.2). A buried secret is a fact which, from our current position, cannot be known. Which of Cesar’s assassins struck his heart is something that we cannot know from our present position. These sorts of facts may be candidates for ersatz reasons, but are tied up in some other debates about the metaphysics of their truth values.

¹⁹In particular, consider realist theories that just bite the bullet on queerness and debunking arguments (which I consider in section 3.4.2). Such theories may grant the queerness and/or debunking arguments but respond, “well, it is impossible for S to believe that F , but it is still possible for F to be a reason for S to ϕ in C .” My response is that this would

regarded as obvious by constructivist theories of reasons. On this version of mind dependence, there is an element of contingency on the minds of the agents, because what S can and cannot believe is a contingent fact about the psychology of S.

As far as I can tell, Kearns and Star are not committed on this point. One might suppose that their analysis entails that a fact be believable, but I do not think this is obvious.

One might hold that $R \iff E$ entails believability because one might suppose that the first premise of the deliberation argument,²⁰ that “[all] practical reasons to ϕ can play an important role in reliable practical reasoning concerning whether or not one ought to ϕ ” entails that all practical reasons can be objects of belief. In order for a fact to play a practical role in deliberation, it must be possible for an agent to believe that fact. We cannot use unbelievable things in deliberation, so the structure of that argument advanced by Kearns and Star would seem to indicate their commitment to such a view.

Alternatively, though, one might suppose that the reasons as evidence thesis does not require that any agent be able to believe that F. It might simply be that F is a fact which grounds a justification for a belief that p, independent of any possible or actual attitudes. Some views of evidence do not require any such possibility.²¹ One might suppose that being evidence is just the relation between a belief function and some elements of probability

undercut the deliberation argument, as well as the value in a theory of reasons developed in chapter 5.

²⁰See my discussion in section 2.4.

²¹The most prominent examples of such a view of evidence and epistemic justification are Williamson (2000) and probably Schafer (1976). Survey discussion is provided by Kelly (Kelly, 2016, Sec. 3).

theory (which are also attitude-independent). If this is so, then one can have a robustly attitude-dependent theory of the nature of evidence.

I think such a theory would be mistaken, but that is a result of my broader constructivist inclinations, and therefore should not be trusted without further independent motivation. However, even if one accepts that evidence need not be the possible object of a belief, one might still reasonably hold that *reasons* need to be objects of belief. Remember that, on my theory, being evidence is a necessary condition for being a reason, but a reason also needs to be able to satisfy the motivational role. One could use the motivational role as a way to establish that a reason must be believable. This is already an established position in the constructivist literature.

O'Neill's version of Kantian constructivism makes this very explicit.²² Let us briefly consider a claim made by O'Neill, regarding Kant. O'Neill writes, "[On Kant's view] Reasons need to be in principle followable by their presumed audiences. That is why they must be lawlike, or have the form of laws." (O'Neill, 2015, p. 118)

What O'Neill has in mind is distinguishing between the ability to *comply* with a reason and the ability to *follow* a reason. Start with the example of and analogy to laws. There is a law that requires one to alternate which side of the street one parks the car. Suppose that Zachary parks his car in accordance with that law. Does his parking in accordance with the law mean he is following the law?

Colloquially, we might just say that one who does not violate the law is "following the law," but this misses O'Neill's point. O'Neill is distinguishing

²²One might argue that this is really *Kant's* view, as O'Neill does. However, I have no idea what Kant is on about when it comes reasons, and so I do not want to presume a correct exegesis. For that reason, I stick with O'Neill.

2. The reasons as evidence condition

following the law with merely complying with the law.²³ If Zachary parks on the far side of the street because he's lazy and does not want to make a U-turn and his parking on the far side of the street happens to align with the law, then Zachary is complying with the law. However, he is not *following* the law.

Zachary is following the law only if he is aware of the laws regarding parking and is aware that he can only park on that side of the street tonight, because of street-sweeping laws. He is following the law only if he is aware of the law in making his decision.

Consider an alternative using reasons instead of laws. When I wake up and go through my morning routine, parts of that routine are involuntary. I do not consider that I am teaching when deciding whether to wash my hair; in point of fact, I rarely decide to wash my hair, as I am too bleary to run the decision procedure.

The fact that I am teaching today is a reason to wash my hair, brush my teeth, and so on. I happen to be performing acts that comply with that reason; I am acting in accordance with that reason. But I am not *following* that reason.

However, that I am teaching on a given Tuesday is *followable*. Were I fully awake and caffeinated enough to engage in the deliberative process, I could consider the reason. The precise details of what O'Neill means by "followable" aside, a minimal condition for a reason being followable is that an agent can believe it.

²³My linguistic intuition about the ordinary use of "following the law" suggests following the law includes mere compliance. The ordinary use is not what is at stake here.

On certain realist views about reasons, some fact F can be reason for S to ϕ in C even if it is impossible for F to believe (much less know) that F . S could never follow F ; F could never actually play a role in justification of S ϕ -ing or deliberating about whether to ϕ . On my commitment to functionalism, as discussed in section 1.4.1, I deny such things are reasons. Based on the deliberation argument provided by Kearns and Star, as discussed in section 2.4, the spirit of $R \iff E$ they set out would also be incongruous, because such facts could not play a role in practical reasoning about whether to ϕ .

2.7 Reasons are evidence

The core of this chapter is the separation of $R \iff E$ into its two constituent conditionals. This chapter endorses one condition and the rejects the other.

Let us close this section with three related, but non-identical claims that this chapter endorses. Each of these claims builds on $R \Rightarrow E$ in a way that explains $R \Rightarrow E$ as integral to my analysis of reasons.

1. If some fact F is a reason for S to ϕ in C , then F is evidence S ought to ϕ in C .
2. That some fact F is evidence S ought to ϕ in C is a necessary condition for F being a reason for S to ϕ in C .
3. That some fact F is a reason for S to ϕ in C obtains partly in virtue of the fact that F is evidence S ought to ϕ in C .

As a strictly extensional matter, item 1 and item 2 are equivalent. Some philosophers will note (rightly) that item 2 is stronger; item 2 suggests a

direction of fit to the relation between the property of being a reason and the property of being evidence, rather than simply noting that the presence of the former implies the presence of the latter.

Item 3 is stronger still. The arguments advanced above, in particular section 2.4, are presented as direct arguments for item 2. One might suppose that these arguments succeed in establishing item 2 but fail to establish the stronger item 3 on the basis that being a necessary condition is not the same as obtaining *in virtue of* being evidence.²⁴

What establishes item 3 is not merely the arguments advanced in this chapter, but rather their occurrence in tandem with the broader accounts of reasons in terms of justifying. Section 1.4.1 notes that some fact justifying S ϕ -ing in C is constitutive of a major functional role of reasons; in fact, as section 2.1 illustrates, for many accounts of reasons justifying is *the* major functional role of reasons. Some accounts of reasons take reasons to be the things that count in favor; some accounts of reasons more precisely hold that the central role of reasons is in justifying. While I think the analyses that center justifying (without treatment of motivating) fail, items 2 and 3 are part of my account as well as those justification oriented views.

As table 2.2 illustrates, my own analysis ultimately disagrees with Scanlon, Parfit, Kearns and Star, and other such analyses about the centrality of the justificatory role. Hopefully the table provides a useful summation of the core thesis and quick reference.

²⁴As a general philosophical point, this may be true. The absence of certain defeaters may be a necessary condition for some proposition being true, but it may still be wrong to say that proposition obtains [even partly] *in virtue of* the absence of such defeaters. But there may be reasonable disagreement on this issue.

2. The reasons as evidence condition

$R \Rightarrow E$	Accept	section 2.4
$E \Rightarrow R$	Reject	section 2.5
$R \iff E$	Reject	Entailed by rejecting $E \Rightarrow R$
item 1	Accept	section 2.4
item 2	Accept	sections 2.4 and 2.7
item 3	Accept	sections 2.4 and 2.7

Table 2.2: Core theses of the chapter and my position

The principal difference between my analysis and these others is that my analysis holds that motivation is also central. However, justification is still a central element for the reasons advanced through this extensive discussion in the reasons literature, and arguments developed by these various figures.

Chapter 3

**The promoting and protecting
what is of value condition**

3.1 Promoting and protecting what is of value

It's a good reason for postponing pleasures that you will then have more time in which you can enjoy looking forward to them. I remember exactly when, at the age of eight, I changed over from eating the best bits first to eating them last.

Derek Parfit, ([Parfit, 1984](#), p. 255)

The second conjunct in the analysis of reasons holds:

R is a reason *only if* R entails S ϕ -ing promotes or protects something of value.

This necessary condition can roughly be separated out into two parts. The first part is the factual constraint that S ϕ -ing must promote or protect something; the second part is that the thing must be of value. Most of this chapter focuses on the latter part; the former part is expounded in section 3.3 and various other elements of the constraints are discussed throughout the chapter, as in sections 3.3.2 and 3.3.5. The latter notion, though, is where the controversy (and therefore the meat) is in this dissertation, as the attitude-dependence of value is an essential element to this constructivist project.

The term “something” here is not meant to be technical. Like the ordinary use of the term, it is meant to be broad. In discussing what it is for something to be of value, I borrow largely from [Scheffler \(2010\)](#) on what

3. The promoting and protecting what is of value condition

it means for an agent to value something; Scheffler's analysis focuses on values, those abstract objects that we try to instantiate, like trustworthiness, kindness, enthusiasm, and so on. I allow that more things than the values themselves may be of value. "Something" as used in this necessary condition is a catch-all for any and all things that satisfy the analysis of being of value section 3.2.2.

Below, when talking about promoting and protecting, I use examples drawn from chess. In the chess case, one typically takes winning and not losing the game to be the relevant thing of value. I discuss other things of value that might be at play in evaluating chess decisions in section 3.3. The goal here is to establish some provisional, intuitive pluralism (within limits) about what is of value.

Unlike the reasons as evidence condition expounded in chapter 2, the reasons as evidence condition, the promoting and protecting condition is my own concoction. Like the reasons as evidence thesis, promoting or protecting something of value is independently necessary.

For the purposes of explaining the promoting and protecting element, and avoid creating confusion around what it means for something to be of value, at various points I just take some values for granted, like the value of winning and not losing in chess. Something being of value is sensitive to context, attitudes, and the like.

Suppose moving my knight to a particular position increases my chances of winning the game; the facts that imply moving my knight increases my chances of winning are reasons for making the move. The move improves movement for my bishop, establishes control of the center of the board, etc. All of these imply improvement in the prospects of the game. Further,

3. The promoting and protecting what is of value condition

the facts that are not so related, like whether my coffee has gone cold or whether I remembered to buy bleach on my grocery run, are not reasons for making the move.

Alternatively, suppose that moving my bishop (rather than my knight) would reduce my odds of losing the game. A move in chess can reduce your odds of losing without improving your odds of winning; while winning and losing are disjunctive, they are not exhaustive, because there are stalemates. Moving the bishop protects my odds of not losing, and the facts that entail the protection of my odds are reasons to move to the bishop.

Increasing and reducing odds are not the only way of promoting and protecting, as I will discuss in section 3.3, but they are a set of cases.¹ If a state is of value, and doing ϕ in C makes that state more likely, than that is a form of promoting that state. *Mutatis mutandis* for reducing probabilities and protecting.

Though I take it for granted at some points in this chapter, winning and not losing are things of value in a game of chess, but they need not be the only thing; in fact, they need not be of value at all in some contexts. What is of value depends on the aims of the agent, facts about circumstances, and other such factors. When playing chess with my niece, what is more valuable is her learning, enjoying, and appreciating the game. In such cases, there may be reasons to do otherwise than moving the knight or bishop, even if such moves make my odds of winning and of losing worse.

As I note, the goal of including “something” is to acknowledge that there is a range of things that can be of value. In superficially similar circum-

¹The idea of “promoting” utility is often used quite generally in the consequential ethical literature.

3. The promoting and protecting what is of value condition

stances, like chess games, the things of value in those circumstances may vary significantly. A tournament chess match and a chess match at a family gathering with my niece influence what things are of value.

3.1.1 Core theses of the second condition

The second necessary condition in my analysis is a new proposal (though influenced by some strands in the existing literature) and so it is important to start with some general theses to get started. These general theses will be refined and clarified as the chapter proceeds.

The goal here is to provide some footing for the theory and situate the theory in the extensive existing literature on reasons and values.

1. Whether S ϕ -ing in C promotes or protects something is sometimes an attitude-independent factual consideration.
2. Strictly on the formulation of the promoting and protecting condition, the condition is compatible with both attitude-independent and attitude-dependent theories of value.
3. We have independent reason to believe that something being of value is attitude-dependent.

Item 1 is simple. There is a fact about whether S ϕ -ing in C has certain consequences. Irrespective of our attitudes towards S ϕ -ing in C or the consequences, the two are related (or not) in a way that is necessary for a fact being a reason.

Suppose Sam is considering whether to jiggle the electrical input on his laptop. There is a fact of the matter about whether jiggling the input will

3. The promoting and protecting what is of value condition

make a difference (either positively or negatively) to that input's ability to supply power to his laptop. This fact is independent of what Sam believes about the power source; it does not matter what Sam wants the outcome to be.² The existence of the reason for his jiggling the input depends necessarily on the causal relation between the jiggling and the outcome. The facts are constraints on being a reason, and are unpacked in section 3.3.2.

The analysis of reasons that I propose is theoretically detachable from the attitude-dependent theories of values, as I note in item 2, but detaching it from an attitude-dependent, constructivist theory of value would be a mistake. Section 3.2 unpacks a theory of being of value, where being of value is attitude-dependent. The purpose of developing this attitude-dependent theory of being of value is to provide an account of how reasons motivate actions; to divorce the attitudes of the agent from something being of value will gut the ability to explain how reasons motivate, and therefore fail to satisfy a major goal of this dissertation, to account for the motivational role, as explained in section 1.1.3.

Following the historical tradition in metaethics, I unpack the direct arguments in favor of an attitude-dependent theory of values in section 3.4.2. The goal here is to tie together the motivational role, valuing, and reasons in a way that provides for the robust explanation of each that is often thinned out by mere description of general attitudes, to give some explanatory meat. I also argue that there are significant constraints on something being of value.

²If Sam is jiggling the input for his own pleasure, then this is a different matter entirely. Promoting his own pleasure is just fine, on accounts of what is of value. The example has to do with his reasoning being motivated and ostensibly justified by the effect on the electrical flow.

3. The promoting and protecting what is of value condition

The chapter illustrates how attitudes give rise to being of value in section 3.2; it then offers a survey of promoting and protecting in section 3.3 and how promoting and protecting provides constraints on being a reason for some agent to act in section 3.3.2 and constraints on something being of value (which in turn influences whether something can be a reason) in section 3.3.4.

In section 3.4.2, the chapter presents the independent arguments for attitude-dependent theories of value developed in the literature, and holds that these arguments succeed and therefore validate the theory of valuing presented throughout the chapter. Finally section 3.4.3 attempts to give an illustration of why my theory of reasons wants to maintain this attitude-dependent account, in particular how it offers an explanation of the motivational role.

3.1.2 Fitting this chapter into the broader project

This is likely the chapter where a reader is most easily led off the path of the dissertation and down a rabbit hole, never to be seen again. Like chapter 2, the purpose of this chapter is to show how one of the two conditions of my theory of reasons (in this case, the promoting and protecting condition) develops an explanation of one of the functional roles of reasons (in this case, the motivational role of reasons).

There are two major interlocking themes in this chapter, which try to give a finer grain to an existing tension in the literature. First: There are attitude-independent constraints on being a reason, including (but not limited to) the fact that S ϕ -ing in C will promote or protect some thing.

3. The promoting and protecting what is of value condition

Second: Reasons are attitude-dependent, because it is necessary that certain attitudes (specifically valuing) are present to give rise to something being of value.

The first is developed explicitly throughout this chapter, and so hopefully will be evident in the reading. The second is both less obvious (at least at some point) and far more important to see in the motivational role.

There are independent reasons for holding that some thing being of value is attitude dependent. The most commonly cited reason is a set of objections, starting with Mackie's famous queerness argument (Mackie, 1977, Ch. 1), entrenched in the constructivist literature. I present these arguments in section 3.4.2.

The second set of arguments is developed through a consideration of the functionalism about

- On attitude-independent theories of value, it need not be the case: if x is of value, that S ϕ -ing in C would promote or protect x , and S knows (even under ideal conditions) that x is of value and that S ϕ -ing in C would promote or protect x , then S would be motivated to ϕ in C , even without any contravening considerations.
- On both logical and phenomenological bases, most theories of reasons hold that S knowing that S doing ϕ in C would promote or protect something of value is sufficient to motivate S to ϕ in C .
- \therefore Attitude-independent theories of value cannot accommodate these logical and phenomenological bases for motivation

3. The promoting and protecting what is of value condition

Some theories of reasons are wholly disinterested in motivation, and so would not care about these features (or take them to be irrelevant to being a reason). Such theories might care only about the fact that a reason justifies $S \phi$ -ing in C and think the motivational role is vestigial. In section 3.4.1, I argue we should not set motivation aside, and instead that we should take this shortcoming of attitude-independent theories to explain motivation to be a serious explanatory weakness.

This serves the broad project of the dissertation by arguing that attitude-dependence generally (and the constructivist, constraint-including variety of attitude dependence in particular) is necessary to provide an adequate explanation of the motivational role of reasons.

The hope is that this chapter provides an illustration not just of the general, abstractly formulated claim that reasons motivate agents to act, but that this element of the theory provides a richer account of *how* reasons motivate agents to act. Specifically, that by spelling out which attitudes make things of value, and how those attitudes involve dispositions to feel and believe certain things, we can see that reasons motivate through both doxastic and non-doxastic elements of our psychology. This is the major appeal to the phenomenology of believing, knowing, and having reasons and values, and my major appeal in closing the chapter with section 3.4.3.

3.2 Being of value

Some things are of value. Some things matter. This locution is sometimes viewed as philosophically provocative, as when Street writes “Nothing really matters, but that’s not what matters” in [Street \(2017\)](#). It is useful in being

3. The promoting and protecting what is of value condition

both provocative (philosophically) and entirely banal and uncontroversial (colloquially). This is a banal truism in most of our lives, but provocative in some contexts in metaethics. My hope is that this discussion illustrates how the statement is both a truism (as something that our lives require for deliberating about action) and provocative (as something subject to substantive disagreement in metaethics).

There are lots of different kinds of things that can be of value. My computer is of value. My relationship with my father is of value. The welfare of people is of value. Honesty is of value. Kindness is of value.

Being of value is metaphysically dependent in part on the attitudes of some agents. If no agent has the right kind of attitudes towards something, then that thing cannot be of value. I *do not* hold that having the right attitude alone is *sufficient* for being of value. There are other conditions that have to be satisfied or must not be violated.³

On my analysis, something being of value is part of the necessary conditions for the existence of a reason. The fact that kindness is of value plays a role in the existence of reasons where performing an act will promote or protect kindness.⁴

3.2.1 Desiring, the general favoring attitude

It is useful here to provide quick and dirty distinction between evaluative attitudes and favoring attitudes. An evaluative attitude is any attitude that

³Schafer (2014) claims that some varieties of constructivism, “Humean” constructivism, holds that there are no non-trivial attitude-independent constraints. This is incorrect.

⁴Instantiation is a kind of promotion and protection; this is especially clear in cases where we talk about the abstract objects, the values that are Scheffler’s focus, like honesty or kindness.

3. The promoting and protecting what is of value condition

ascribes a normative property (e.g. being good, permissible, right, etc.) in its content. A favoring attitude is any attitude that disposes an agent to act in pursuit of one state or another.

When I believe that it is wrong to donate to the Salvation Army because of their views and policies on homosexuality, I am holding an evaluative attitude. That I am not disposed to make such a donation is a favoring attitude. These things may come apart, as we will see below in discussing cases where individuals favor ϕ -ing inconsistently with their evaluative attitudes or *vice versa*. These cases have an important implication for my view of being of value as, following Scheffler, I think some sense of evaluative *and* favoring attitudes are necessary for an agent valuing something.

Not every favoring attitude contributes to a thing being of value; nor does every evaluative attitude. So believing that it is right to vote is an evaluative attitude; wishing that someone would close the window is an evaluative attitude.⁵ But the mere believing or wishing is not sufficient to make something of value.

Being indifferent towards something is a favoring attitude;⁶ being indifferent towards something will not make that thing of value. It is conventional to talk about the favoring attitudes towards things as the basis for something being of value, on attitude-dependent theories, but the convention of including all favoring attitudes or all evaluative attitudes is a mistake.

General favoring attitudes may in some circumstances be too weak to give rise to a thing being of value. In the following sections, I argue the

⁵I leave out instances that do not have such clear content. Is an itch an evaluative attitude? Is a hunger pang? These are favoring attitudes that are correlated with a certain evaluative attitude, e.g. that I should eat. But they are not themselves evaluative.

⁶I recognize that this is a weird use of favoring, where “disfavoring” would also count as a favoring attitude, but I have not got a better term.

3. The promoting and protecting what is of value condition

mere presence of the favoring and/or evaluative attitude is not sufficient. In order to make some thing of value, the evaluative attitude should be supplemented with dispositions and a higher-order evaluative attitude. To distinguish between favoring and evaluative attitude, in contrast to a richer combination of attitudes and dispositions, I stipulate a distinction between desiring and valuing, respectively.

In this dissertation, I use *desiring* as a broad catch-all for the range of positive evaluative and favoring attitudes. This is inconsistent with many uses surveyed in [Schroeder \(2017\)](#), though it is consistent with the broader synecdochic uses embraced by some philosophers, including [Anscombe \(1957\)](#) and [Nagel \(1970\)](#).

Consider the notion of “wanting.” Baker wants a glass of water. Suppose that this want is an unmotivated inclination not based in some other reason (like the belief that he is dehydrated) or expectation of outcome (like that drinking the water will be pleasurable). He just wants some water. This is a favoring attitude towards drinking water.

On some views of reasons, the presence of this want for Baker is sufficient in itself to ground a reason for Baker to act. This is where my view is more radical than other attitude-dependent approaches. I think this is not enough, by itself. Rather, the fact that Baker has this want operates in conjunction with some further necessary conditions (which other theories [wrongly] take for granted) ground the existence of a reason for Baker to drink water.

My arguments against the view that mere favoring or evaluative attitudes are sufficient are implicit at many points throughout the chapter, but it is important to make them explicit here.

3. The promoting and protecting what is of value condition

The reason for Baker to drink water, based on this unmotivated desire, holds only if we also accept a few other assumptions. First, that the satisfaction of desires is good *ceteris paribus*, and that drinking of water is a case that falls within the scope of that *ceteris paribus* condition. Second, that Baker's unmotivated desire for water has some force (however weak) on his behavior. Third, that the case is not defeated by some other facts about bodily imperatives or inclinations.

In the first condition, one supposes that the satisfaction of unmotivated desires is good, at least *ceteris paribus*. Suppose that one holds the view that bending to unmotivated desires is bad, develops weakness of character and will, and other such things. This should not be a strange view, since some ascetic theories of value hold such a claim. On such a view, the fact that Baker has this unmotivated want is not a reason for Baker to drink water. Further, on that view the fact that Baker has an unmotivated want is an affirmative reason for Baker to abstain from drinking the water.

Many people hold, I think intuitively and correctly, that the satisfaction of desires is good *ceteris paribus*.⁷ My point here is just that we should not ignore this condition and its role in Baker wanting water. That the satisfaction of such desires is good is both subject to some debate and necessary for establishing the existence of a reason.

The second condition is a bit more complicated. Suppose that Baker has the want, but is in no way motivated to act on it under any possible circumstances, even if there were no opportunity cost, no labor, nor any such countervailing consideration. If S desires x, but has no inclination to

⁷I say "*ceteris paribus*" here because there may be circumstances where the satisfaction of desires is bad, perhaps because of the content of the desire, or because of some higher-order desire, or some other such fact.

3. The promoting and protecting what is of value condition

act in accordance with getting x, then it gets strange to talk about S desiring x as grounding a reason. I submit that the pretext examples considered in section 2.5 are evidence of this, as they fail to satisfy the motivational dimension. I supplement this point with an alternative argument from our phenomenology of recognizing reasons in section 3.3.5.

The third condition is stranger still, because it raises a concern about what it is to want water without feeling anything in the body related to wanting water. Suppose I want water but am not thirsty, or feel no pleasurable expectation at the thought of water, or any other such bodily sensation. I suppose it is possible that we can talk about favoring in this way, as purely intellectual and non-bodily, but I am skeptical that this provides an accurate representation of our ordinary experience of desires.

I will not suggest that the reader need to keep all of these points in mind; that would be a lot to carry on what is admittedly the most complicated and treacherous chapter in this dissertation. Rather, these claims work as cairns that keep us on the metaphorical trail and provide points of reference and elucidation in rereading.

3.2.2 Valuing, the specific evaluative attitude

I use *valuing* following the account laid out by Scheffler (Scheffler, 2010, ch. 1). Scheffler's account of valuing is not meant to be an analysis, but rather to provide a way of fleshing out what it means to talk about values (those abstract objects like friendship, trustworthiness, etc.) held by individuals and communities. He does not take this account to provide necessary and sufficient conditions, but rather to provide a heuristic.

3. The promoting and protecting what is of value condition

One point of notable difference: I used “things of value” more broadly than Scheffler uses “values.” When Scheffler provides his analysis, he is talking about a set of abstract objects which (while not clearly delineated) is understood to be narrow. Scheffler does not consider hydration to be a value; it might be a constitutive element of health, and health might be a value, but hydration proper would not be. By contrast, I think we should regard the potential set of things of value very broadly. As we will see, what individuals can value (on the terms of Scheffler’s analysis) and find as motivating is considerably broader than just the things that we might properly identify as values, in the narrow sense.

Scheffler’s account of valuing usefully includes both certain evaluative attitudes *and* certain favoring attitudes. The former are covered by item 1 and item 3; the latter are covered by item 2 and item 4.

Scheffler holds that valuing has four elements (Scheffler, 2010, pp. 29, directly quoted):

1. A belief that X is good or valuable or worthy.
2. A susceptibility to experience a range of context dependent emotions regarding X,
3. A disposition to experience these emotions as being merited or appropriate,
4. A disposition to treat certain kinds of X-related considerations as reasons for actions in relevant deliberative contexts.

Scheffler’s items 2 and 4 explain the motivational role; items 1 and 3 help to make sense of treating a reason as evidence for certain normative proposi-

3. The promoting and protecting what is of value condition

tions and as explanation of individual behavior. As such, Scheffler's account of valuing does a lot of work to move towards the thick account of reasons that we are pursuing. If valuing grounds being of value in the context of the second condition (as I argue it does throughout this chapter), then the second condition will provide us with a rich account of the motivational role of reasons.

Let us start by considering a case where an individual has a desire for something and the belief that the desire is not warranted. This is not valuing, since it violates item 3 and probably also item 1.

If Baker desires a drink of water in the sense specified above, then Baker may well be susceptible to experience a range of emotions when he drinks. But suppose Baker adopts the belief that the desires associated with drinking water are not warranted (undercutting item 1) and loses the disposition to experience the emotions as warranted. Should we say that Baker values that drink of water? No. He may desire that drink of water, but that desire is now outside of the proper subset of desires that is valuing.

There have been some efforts to make sense of potential restrictions on the notion of valuing as used by Korsgaard, as the psychological state, in distinction to (though underlying) the state of something being of value.⁸ Korsgaard writes, "My own view is that talk of the existence of values at this level is just misleading shorthand for something else, namely valuing, which is a thing that we *do*. To say that something is valuable in this sense is to say that people appropriately value it." (Korsgaard, 2003, p. 68) Korsgaard

⁸The most oft cited discussions are Korsgaard (1983); Korsgaard et al. (1996). People struggle with the exegesis of both of these; I have found the views elucidated in the more recent Korsgaard (2003) more useful and broadly consistent with my (hopefully correct) reading of her older work.

3. The promoting and protecting what is of value condition

then goes on to emphasize that “appropriateness” here does not require the stipulation of ontologically objective, attitude-independent value properties in the object, but simply puts some constraints on valuing and being of value.

Korsgaard is a constructivist about value.⁹ In particular, she has a constructivist position about values that usefully emphasizes two core tenets. First, being of value is necessarily attitude-dependent. Second, being of value is subject to attitude-independent constraints. I advance both such theses in the course of this chapter.

Scheffler’s account illustrates a way in which one might have the relevant desires (or more specifically the wants and favoring states) but still fail to value something, not in virtue of formal constraints (which we will return to below) but in virtue of the psychological states themselves.

A number of philosophers have raised this issue across a range of contexts; some theories incorporate second-order attitudes (e.g. Baker wants to go without his emotional responses to drinking)¹⁰ which may serve as ground for countervailing *pro tanto* reasons for action, reasons that push against the reason for Baker to have a drink of water in all (or nearly all) circumstances to block the possibility of a *pro toto* reason. The existence of such second-order attitudes is certainly a feature, and plays a role in making sense of the actual deliberation of individuals, their reasons, and their actions. It will play a role in the consideration of a range of problem cases throughout our discussion. But to shift immediately to those cases is to sell short the complexity of valuing.

⁹She is often regarded as the prototypical Kantian constructivist, because of her very active use of Kant in discussions of values and reasons [Korsgaard and Korsgaard \(1996\)](#).

¹⁰This account of second-order desires can be found in [\(Schroeder, 2017, Sec. 2.2\)](#).

3. The promoting and protecting what is of value condition

This account of valuing gets right a core, and frequently overlooked, set of circumstances.

When I was a child, my paternal grandparents listened to Wagner. Growing up Europe before World War II, Wagner was a significant part of their Czech culture; after the war, they had gone on their first date to an opera, a major form of those cultural elements. They would put on the cassettes and listen. I knew it was important to them and to the family, but I never had the positive emotional response to Wagner. I believed that Wagner was good, satisfying item 1, and I believed that such emotional responses were warranted, satisfying item 3. I just did not have those emotional responses. I could not make myself satisfy item 2.

I did not value Wagner; I believed Wagner's work to be good and important, and even an instance of other aesthetic values that I thought were worth promoting. I valued those experiences with my grandparents, as I did experience emotions as a result of my time with them, but I did not value Wagner. I might desire to listen to Wagner under certain conditions; there would be other things for which Wagner was instrumentally valuable, but the music itself was not a subject of my valuing.

There were things that I valued for which Wagner was instrumentally necessary, like the nostalgic feelings my grandparents had towards Wagner. I valued Wagner's music as a means to some other things that I value.

There is a distinction between instrumental values and things that have value in themselves. At some level, Scheffler's account of valuing will exclude a range of instrumental values, distinguishing what one "really" values (the things that satisfy Scheffler's conditions) from what is necessary to achieve those values. Though this distinction gets a little messy, as there are

3. The promoting and protecting what is of value condition

circumstances where an instrumental value may itself become an object of valuing (in Scheffler’s stronger sense) because of the association.¹¹

Perhaps, over time, seeing responses to Wagner produces an association and range of dispositions in my experience such that I satisfy the condition. Even though listening to Wagner itself is not properly the source of these dispositions to feel and act, it may satisfy the conditions of Scheffler’s analysis. Dispositions are not static; they change over time. So do our beliefs.

The attitude dependence of values and reasons entails that the reasons for agents to act are themselves dynamic over time. As the attitudes of agents change, so will their reasons. This is important throughout chapter 6.

The second-order values may do some work here. Perhaps I valued enjoying Wagner; perhaps I even valued valuing Wagner. But that did not change absence of the first-order attitude, based on the absence of the emotional disposition.

Table 3.1: Extensions of cases, Chapter 3

Cases	item 1	item 2	item 3	item 4
John wants water	N	Y	N	Y
Wagner	Y	N	N	Y
Nagel’s tragedies	Y	Y	Y	Y
Sam’s liturgy	Y (defeated)	Y	Y	Y
Bodily Imperatives	Y	Y (at issue)	Y	Y
Abusive husband	Y	N	N	N
Kramer’s Anger	Y	Y (with emphasis)	Y	Y

¹¹Korsgaard draws out this distinction (Korsgaard, 1983, p. 169-173) and provides some parsing for attitude dependent theories of value. She does use the term “intrinsic” throughout, rather than referring to things having value in themselves. She notes that this can cause semantic confusion over our valuing x being intrinsic and the property of x having value being intrinsic. This semantic slippage is the reason I use the alternative locution and focusing on *valuing* the thing in itself instead.

3. The promoting and protecting what is of value condition

The table provides a survey of the cases considered in illustrating Scheffler's analysis of valuing throughout the course of this dissertation, along with defeaters that prevent valuing from constituting being of value. The purpose of these is to illustrate the relative complexity of Scheffler's account of valuing and in particular to put emphasis on the dispositions in item 2 and item 4 that help to flesh out the relationship between values and motivation.

3.2.3 Nagel's "War and Massacre" and moral tragedy

puzzles

There are a set of cases that might be presented as counter-examples, but I take rather as illuminating an important feature of valuing, being of value, and the importance of evaluative attitudes in contexts.

There are circumstances where an individual has a range of potential choices that, on their face, do not seem to promote or protect any value. In those cases, the major consideration is between a range of things of disvalue, of varying severity, and the reasons for action implicated in justifying or motivating the act is a matter of triage, of trying to manage the various adverse possibilities as best we can. The worry about such cases is that my theory of reasons will fail to categorize the reasons for action in those cases as reasons; they clearly are. The fact that one must choose between two terrible acts does not mean that there are no reasons as play. I contend that, even in this case, the promoting and protecting condition is still at play, and there are still values to consider.

Moral tragedy cases are structured by drawing up two potential acts, ϕ and ψ , where neither act would promote or protect things of value (or, at

3. The promoting and protecting what is of value condition

least, of significant enough value to render the act permissible) if considered only on their own. Absent consideration of exigent circumstances, S ϕ -ing would be impermissible. *Mutatis mutandis* for S ψ -ing. However, there are some facts of the circumstance (for example, having to choose between either ϕ or ψ to avoid some alternative worse than either) such that one of those acts becomes the best possible choice. So-called “Buridan’s ass” cases provide a similar structure.

This functions as a putative counterexample to my analysis, because neither ϕ -ing nor ψ -ing seems to promote or protect anything of value. I argue that this counterexample fails because one must consider promoting or protecting things of value *in context*, which includes consideration of the circumstances.

In Thomas Nagel’s “War and Massacre,” he raises worries about the nature of certain moral positions being taken as severe and inviolable. He writes(Nagel, 1972, Sec. 3),

In 1958 Miss Anscombe published a pamphlet entitled *Mr. Truman’s Degree*, on the occasion of the award by Oxford University of an honorary doctorate to Harry Truman. The pamphlet explained why she had opposed the decision to award that degree, recounted the story of her unsuccessful opposition, and offered some reflections on the history of Truman’s decision to drop atom bombs on Hiroshima and Nagasaki, and on the difference between murder and allowable killing in warfare. She pointed out that the policy of deliberately killing large numbers of civilians either as a means or as an end in itself did not origi-

3. The promoting and protecting what is of value condition

nate with Truman, and was common practice among all parties during World War II for some time before Hiroshima. The Allied area bombings of German cities by conventional explosives included raids which killed more civilians than did the atomic attacks; the same is true of certain fire-bomb raids on Japan.

The policy of attacking the civilian population in order to induce an enemy to surrender, or to damage his morale, seems to have been widely accepted in the civilized world, and seems to be accepted still, at least if the stakes are high enough. It gives evidence of a moral conviction that the deliberate killing of noncombatants—women, children, old people—is permissible if enough can be gained by it. This follows from the more general position that any means can in principle be justified if it leads to a sufficiently worthy end.

Nagel frames all of this as an objection to Anscombe's moral absolutism. I find this objection, coupled with my consideration of my paternal family in Theresienstadt and Auschwitz-Birkenau, decisive in undercutting Anscombe's absolutism.

But the closing of Nagel's section is especially potent. "Despite its [absolutism's] appeal, it is a paradoxical position, for it can require that one refrain from choosing the lesser of two evils when that is the only choice one has. And it is additionally paradoxical because, unlike pacifism, it permits one to do horrible things to people in some circumstances but not in others."¹²

¹²One worry here is about connexity, the principle governing deliberating between reasons in cases where S has no inclination between two possible acts ϕ and ψ , but failing to choose one or the other will result in something far worse than either. (These are the

3. The promoting and protecting what is of value condition

“A moral tragedy occurs when, all things considered, every viable option one is confronted with involves a serious moral violation.” (Orend, 2001, pp. 28) The examples provided by Nagel are no doubt an instance of moral tragedy. Following Nagel and Walzer (Walzer, 2015, ch. 16), one can argue that there is a *pro toto* reason to engage in repugnant behavior based on the view that one *must* do a repugnant act (because all of the options are repugnant) and therefore consideration is among *pro tanto* reasons for repugnant acts in that context.

If the solution there is so elegant and so simple, and requires no appeal to the theory of reasons (mine or some other) then why mention it at all?

Setting aside the motivations for my frustrations with Anscombe and general inappropriate ambition, there is a philosophical purpose here.

The analysis of reasons at the core of this dissertation is, by design, more restrictive than many of the others on the market. The hope is to avoid over-generation worries directed at attitude-dependent (including constructivist) theories of reasons. Moral tragedies look like counter examples against more restrictive theories of reasons because they seem to provide clear cases of reasons where the facts under consideration are not related to things of value, but only of greater and lesser disvalue.

Further still, the restrictive nature of valuing unpacked above aggravates this problem. Can it be the case that there are things of value to be promoted and protected in considering the perpetration of grotesque acts of carnage, as the moral tragedy requires? Surely it is not the case that the widespread destruction of civilian life is, itself, valued by reasonable people trying to

so-called “Buridan’s Ass” cases.) Something in the canons of rationality requires that S does either ϕ or ψ , rather than simply allowing the worst outcome by doing neither.

3. The promoting and protecting what is of value condition

work out how to fight a war against the Third Reich. (We will directly address the possibility that some people do value such destruction later in the dissertation, with appropriate discomfort.)

But the restrictive theory only seems like a mistake. It is not in fact a mistake.

There may be things an agent values at play in performing a morally repugnant act. Saving lives, preventing genocide, etc. are at play even in the commission of horrible acts. Even if, all things considered, the act is immoral on the basis of some consideration, these facts about what is of value and the fact that a horrific may promote or protect what is of value still obtain.

3.3 Promoting or protecting

Promoting and protecting are not technical philosophical notions; rather, they are broad synecdochic terms for intuitive notions. Acts influence the world around us. Promoting and protecting describe forms of influence on different things in the world, including things that are of value.

In the initial examples pertaining to chess, I talk about promoting and protecting game states of winning and losing in terms of odds. This is straightforward, since winning the game or losing the game represent straightforward binary states. Strictly speaking, there are not degrees of winning or losing, though there may be more or less dramatic ways of winning or losing. As such, promoting and protecting are easy to flesh out just in terms of maximizing the odds of achieving one of those binary states.

3. The promoting and protecting what is of value condition

When we talk about more complex things, whether values as abstract objects like trustworthiness or things of value like financial security, then we quickly see that there are other ways of promoting and protecting that are not reducible to odds. Consider financial security; one might promote financial security by saving or building credit or investing. There are some respects in which these promote odds (e.g. long term investments protect the return on investment through exploiting good odds over time), but one might also talk about the volume of savings instead of odds and return on investment.

In addition to telling a probabilistic story (as in the chess cases), one might also tell a maximizing and minimizing story about promoting and protecting. If $S \phi$ -ing in C increases the volume of x , then $S \phi$ -ing in C promotes x . If $S \phi$ -ing in C prevents the volume of x from diminishing, then $S \phi$ -ing in C protects x . This is not exhaustive of the ways in which one might promote or protect things of value. It is merely meant to illustrate diversity in the ways in which we may promote or protect things of value.

Odds and volume are two ways of talking about varied dimensions of promoting and protecting. An alternative, perhaps more colloquially familiar, approach to promoting and protecting has to do with signaling. When I say “we should maintain transparency in how the philosophy department operates in order to promote trust,” it would be odd to cash this out in terms of the odds of trusting states or the volume of trust. Trusting states lend themselves to neither clear odds nor clear assessment of volume. What the commentary on promoting trust is getting at is the public signaling that trust is an appropriate response, and that there is a shared community goal of trusting one another.

3. The promoting and protecting what is of value condition

Promoting and protecting are meant to be understood in this pluralist way. That is, all of these uses are acceptable and some will be appropriate in different contexts.

3.3.1 Promoting, protecting, and attitudes to risk

There is some discussion of “promoting” as a general (and, as far as I can tell, non-technical) term in the literature on moral theory. In particular, this phrase is used by consequentialists as a way of talking about promoting some desirable consequence (for example, utility). In some of the classic literature that makes use of this language, especially [Bentham \(1789\)](#); [Mill \(1863\)](#), there is a sort of monism about both the thing of value promoted and a lack of specificity about potential pluralism in modes of promotion.

One of the developments in 20th century theories, particularly economic and political theories that develop with Rawlsian liberalism, is that there is a clear trajectory towards pluralism both in the things of value and in the ways that things of value may be promoted or protected. I advance and defend a particular version of pluralism about value in section [3.4](#), but it is useful to say a bit about the pluralism regarding promoting and protecting. I cannot dwell on this too long, as it leads far afield into discussions of decision theory, mathematics, and economic theory that are less immediately relevant to the present project.

Rawls’ core argument for a principle of distributive justice based on the original position and veil of ignorance rests on the claim that (in the original position and behind the veil of ignorance) one should adopt a risk-averse posture towards the design of a society([Rawls, 1971](#), p. 71-73). He advances

3. The promoting and protecting what is of value condition

this as a substantive, first-order argument. By appealing to certain attitudes and circumstantial principles, Rawls argues that there is strong *pro tanto* reason to take an averse attitude in the original position. It is wrong to conceive of Rawls' argument as a view about the necessity of one's attitudes, for reasons of Rawls-exegesis that would take us far afield,¹³ but what does emerge out of this argument is a criticism of Rawls on attitudes to risk.¹⁴

Let us set aside the case of the original position and the veil of ignorance and instead consider a simple case. Suppose Baker and I are both skiers of identical skill levels in identical circumstances; we are both considering whether to take a double-black-diamond (extremely high difficulty) run. Both of us share the same values of fun and the same concerns about potential injury and, for both of us, the risks associated with fun and injury are identical. The sole difference between Baker and me is our attitudes to risk: Baker is more risk averse such that he is disinclined to risk injury in pursuit of fun, and I am inclined towards risk such that I will more often risk injury in pursuit of fun.

There are a few ways one might handle this case, but they break down into two camps based on a judgment about the case. The first is that one might judge there just is a fact of the matter about the correct attitude to risk. It is exclusive that either Baker has the right attitude, I have the right attitude, or neither of us does; it cannot be the case that both Baker and I are right. We might call this monism or monistic realism about attitudes to risk.

¹³Rawls (1974) provides especially strong evidence for this exegetical position.

¹⁴See Traub et al. (2005) for a survey that includes some criticisms of Rawls on attitudes to risk. See Waldron (1986) for an especially astute discussion.

3. The promoting and protecting what is of value condition

Alternatively, one might be inclined to judge that both Baker and I can be correct in our attitudes to risk simultaneously, and that insofar as there is a fact of the matter about appropriate risk, that fact is dependent on our attitudes to risk such that both can be correct. This does not necessarily mean that both are correct, as we will see in section 3.3.2 discussing constraints. There can be constraints on attitudes to risk such that some attitudes are incoherent or wrong; there can be substantive arguments in favor of preferring one attitude to risk over another. However, there can also be variation in attitudes to risk that results in difference in what there is reason (either *pro tanto* or *pro toto*) for an agent to do.

The technical details do a few things within the potential structure of a theory of reasons, some of which are directly tied to my own assessment advanced here and some of which are loose ends that one might tie later in developing this theory.

- Attitudes to risk provide another potential axis for attitude-dependence. My analysis only focuses on the attitude-dependence of being of value, as discussed in the section 3.4. However, one might advance the thesis that promoting and protecting are themselves attitude-dependent (under constraints).
- Disagreement about attitudes to risk provide a means for assessing meaningful disagreement about what there is most reason to do, both in personal and social situations. This provides a separate means for addressing reasonable disagreements presented in sections 5.4.2 and 6.3.4, though I do not consider this approach directly.

3. The promoting and protecting what is of value condition

- Attitudes to risk have a range of formal constraints that they may introduce which can be brought to bear on the promoting and protecting condition. I only provide a broad gesture on the constraints in section 3.3.2, and this leaves open an avenue for significant, fruitful development of the relationship between these formal domains and the grittier domains of applied metaethics and political theories.

In any case, this sort of material often falls under considerations of capital-R-Rationality, which is closely related to the present project, but which would be too ambitious (in the pejorative sense) to pursue in this dissertation.

3.3.2 Constraints on reasons, the promoting and protecting condition

There is some literature on the rationality of desiring, especially valuing under rankings of states that one desires. Some constraints prevent internal incoherence in things being of value; some govern the order of an agent's preferences to prevent incoherence in reasons. These may inform constraints on reasons, though they most clearly apply as constraints on the relationships between competing *pro tanto* reasons.

Acyclicity is a constraint on the ordering of preferences. Following Hansson and Grüne-Yanoff, "There is no series A_1, \dots, A_n of alternatives such that $A_1 \succ A_n \succ A_1$." (Hansson and Grüne-Yanoff, 2018, Sec. 5)

Suppose Sally was voting in the 2020 Democratic primary for President, in the early days when there are a dozen or so candidates; the set of candidates Biden, Bloomberg, Buttigieg, Gabbard, Harris, Klobuchar, Sanders, Steyer, Warren is ordered. It should not be the case that for any candidates,

3. The promoting and protecting what is of value condition

Sally produces an ordering that creates a cyclical preference; she cannot (for example) have an ordering that produces Sanders \succ Biden \succ Sanders.

Connexity is a constraint that requires a mechanism to resolve cases of genuine parity, cases where there is no ordering such that an agent values x over y or *vice versa*. If S puts O and N at parity (that is, has no preference between the two) and prefers both to P , and failure to act consistent with either O or N gives rise to P , then there has to be some way of reconciling either O or N . I will not defend any particular version of the story of a logic handling this question. Some such thing is a constraint.

Suppose that Tom is indifferent between Sanders and Biden, but prefers both to not voting. In order to prevent the states where he does not vote (like the states where Buridan's ass starves to death) he has to make a choice. Perhaps the mechanism for that choice is arbitrary; perhaps the mechanism is random.¹⁵ But there has to be some choice. It is not acceptable for Tom's inaction to result in his not voting.

These constraints are not intended to be exhaustive,¹⁶ but rather serve as a gesture at some of the formal constraints we might reasonably expect governing ordering of our values, as well as the relations between reasons in deliberation.

¹⁵My maternal grandfather John provides a useful approach for addressing arbitrariness. One performs a coin flip to make sense of the case and then, if one experiences regret at the course of action selected by the arbitrary method (the coin flip) then revert to the other option. The idea is to use arbitrary selection techniques to determine if there was a weaker preference, and assessing that weaker preference by way of regret.

¹⁶There is an entire literature on rational choice dedicated to exploring potential formal constraints of this nature. For a fairly canonical survey, see [Coleman and Fararo \(1992\)](#). For attempts to probe the role and limits of rational choice theory in the domains discussed in this dissertation, see [Boudon \(1998, 2003\)](#).

3.3.3 Factual success-conditions for valuing, being of value

There are a number of ways in which an agent valuing some thing x may fail to make x of value. One case is when an agent's valuing is related to a false belief. If S satisfies Scheffler's belief condition, item 1, on the basis of something false, then that may undermine the thing being of value.

Here I will explore the factual constraint. Below, I find that an agent's valuing fails to give rise to being of value if and insofar as the falsity of belief results, or would result, in the attitude of valuing being undermined by the claim.

An ordinary version of the factual constraint (which I will reject for reasons argued below) is that an agent S valuing a thing x will fail to make x of value *if* S holds some false beliefs about x . This is too restrictive. One might hold lots of false beliefs about things we value that are extraneous to our valuation of that thing. If Sam values his morning coffee and also believes that coffee reduces cancer risk, then Sam's valuing coffee would fail to give rise to his coffee ritual being of value *even though*, were his false belief to be corrected, he would still value that morning cup of coffee. This counterexample indicates that we need a less restrictive version of the condition.

In addition to the ordinary factual constraint above, there are two more sophisticated versions.

An agent S valuing x will fail to make x of value if S came to value x in virtue of some false belief p . On this factual constraint, the [false] belief was a condition for S valuing x .

3. The promoting and protecting what is of value condition

An agent S valuing x will fail to make x of value if S values x only if S [falsely] believes that p. On this factual constraint, the [false] belief is necessary for S continuing to value x.

These two are extensionally different. Below, I argue that the last formulation (the only if formulation) is correct.

Elizabeth values saucers of mud only if she believes that doing so promotes the favor of the great Mud God; she is mistaken about a matter of fact, as there is no great Mud God. Suppose she discovers that there is no Mud God and, as such, no longer values saucers of mud. She may still have the relevant emotional responses, but she is no longer inclined to believe the saucers of mud are good and she is no longer inclined to experience her emotional responses towards saucers of mud as being merited (failing Scheffler's items 1 and 3). Under such circumstances, we should say that the saucers of mud were never of value, as this disposition to lose her valuing in the face of a false belief does undercut the saucers being of value.

3.3.4 Argument for the “only if” formulation of the factual constraint

A more serious example where the “only if” condition is not satisfied; an agent comes to value something as a result of a false belief but finds that the belief being undermined does not undermine the valuing.

Suppose Sam comes to value the Jewish liturgy because of his religious beliefs, as a child taught that the stories of the Exodus were true. But also suppose that Sam would still value the liturgy even if he knew those religious beliefs to be false; even if he believed there was no God, no covenant with

3. The promoting and protecting what is of value condition

Abraham, no Exodus, then he would still believe the Jewish liturgy was good (Scheffler's item 1) and that his emotional responses to the liturgy were merited (Scheffler's item 3).

In the case, the fact that Sam holds a false belief does not create a problem for the liturgy being of value. The false belief may be implicated in Sam's current attitudes towards valuing, but the counterfactual situation where he does not hold that belief results in no difference. Sometimes beliefs that are implicated in our valuing are vestigial; like the appendix, if removed they result in no change.

Nagel's distinction between motivated and unmotivated desires here works both for desires (broadly understood, per my stipulation) and the more narrow case of valuing (per Scheffler's account).

Nagel writes that some desires are *motivated* while others are unmotivated (Nagel, 1970, 2.V). Our motivated desires are the product of "decision" and "deliberation," while our unmotivated desires "simply assail us." (Nagel, 1970, pp. 29) Nagel's distinction, though, focuses on the conditions that bring about the desires, rather than the conditions necessary to sustain those desires. My goal here is to transition focus in the context of making things of value to the latter. The constraint on being of value is not whether the relevant attitude (what I call "valuing") is produced through either a deliberative process or without any rational deliberation whatsoever. As such, those unmotivated desires are often resistant (though not impervious) to changes in doxastic states.

Whether an attitude is sufficient to ground something being of value is not a matter of how the attitude was acquired, but of what keeps the attitude in place or would displace it. This is not a story about the journey to an

3. The promoting and protecting what is of value condition

experience of valuing, but of the ongoing experience of valuing and the conditions that sustain that experience for the agent.

This constraint is importantly different from the *promoting and protecting* condition, as promoting and protecting some thing of value is a relationship between something being of value and the existence and nature of a *pro tanto* reason. This constraint is, to borrow a metaphor, “upstream” of promoting and protecting, because a thing has to be of value before promoting or protecting that thing can play a role in some fact being a reason. The conditions enumerated here must be satisfied antecedent to the conditions on being a reason.

It is important to bring up the conditions on something being of value, because that is a major constituent (on my analysis) to being a reason. On constructivist theories of value, as I discuss in section 1.3, something is of value in virtue of the attitudes held by some agents.

This is because the attitude-dependent picture of values is the crux of serious over-generation worries about theories of reasons. The idea that a thing can be of value (and, thus, provide reasons) based on the attitudes of agents creates worries about the over-generation of values, because lots of agents value weird things. The account provided above should not directly assuage those worries about over-generation;¹⁷ in fact, I suspect it feels to many like I am unwittingly setting myself up for devastating counterexamples. At the very least, I am witting. We will return to these counterexamples in chapter 6 and section 7.4.

¹⁷Some constructivist theories have imposed very strict, substantive constraints on the ways in which attitudes can give rise to being of value. Sometimes this includes a substantive constraint on an agent (or their attitude) being reasonable. Sometimes this includes the hypothetical consideration of an “ideal” agent of similar mental character. I am deliberately not appealing to these here, and will not adopt them.

3.3.5 Bodily constraints and influences on values

One of the challenges facing Scheffler's account of valuing is that it includes certain conditions that are not clearly either doxastic or evaluative. Dispositions to feel certain emotional responses provide a useful and involuntary constraint on what we value. These dispositions also illustrate how our valuing provides motivational force for individuals to act in accordance with with their values.

These include our bodily dispositions to certain kinds of emotional responses. [Manne \(2017\)](#) notes there are some dispositions, what she calls "bodily imperatives" to act, feel, and consider certain types of facts. The presence or absence of these dispositions are psychologically and physiologically contingent facts; they depend on what our minds and bodies are like in certain details that might have been (and, in some cases, might be) otherwise.

Dispositions to feel, per Scheffler's item 2 and item 3, is a fact about the way our minds and bodies are. In the context of Manne's broader metaethical project, these dispositions are explicitly implicated in internalism about reasons and about motivation.¹⁸ The dispositions that produce motivations are the dispositions to feel particular ways. These dispositions motivate individuals to act, to achieve pleasant emotional experiences and bodily states or avoid unpleasant ones. Sometimes discussion of motivation focuses on flattened versions of these states, like pain or aversion or pleasure. Sometimes our emotional states and dispositions are that simple. Sometimes they are not.

¹⁸See [Manne \(2014\)](#).

3. The promoting and protecting what is of value condition

Watching the suffering of another person is not just a painful or uncomfortable experience. It can include a range of other more subtle features and inclinations. Can I make their suffering stop? Is their suffering (in my judgment) unjust or undeserved? Can I relate to a similar experience of that suffering? Does seeing their suffering evoke some of my own past experiences? These subtleties may apply in such a case. In some such cases, there may be an inclination (varied in strength) to provide help; in some cases where helping is impossible, there may be an added level of discomfort with the notion that there is nothing one can do.

Bodily imperatives add a useful layer in discussing reasons for action. They also gesture at an important set of facts that are non-reasons or only reasons under circumstances. Not all of the dispositions that motivate us to act correspond to reasons.

There are some facts that are not reasons, but do incline us to action. Things like the physical associations connected to addiction are tied up in our bodily imperatives, but will not necessarily be reasons. On my analysis, such cravings will fail to be reasons on the basis of items 1 and 3 in many cases, but are a useful way of making sense of our bodily imperatives. If Sam is addicted to the painkiller Ketamine, then Sam will experience certainly bodily imperatives related to the pursuit and consumption of Ketamine. When Sam enters Narcotics Anonymous, or otherwise attempts to cover from his addiction, then it will hold that Sam no longer values Ketamine on the basis of his beliefs about Ketamine and about whether his desires are merited.

When Sam is in NA, the inclinations will not be a reason for him to act on the basis that his values do not accord with the inclination. While

3. The promoting and protecting what is of value condition

he has the bodily inclination to use Ketamine, he no longer has the other components necessary for valuing, and so this undermines the reasons. Some less weighty reason may remain, like the prospect of being rid of his discomfort, but not the reason informed directly by the craving for the drug.

I say above that the discussion of bodily imperatives adds a useful layer to our discussion. It is not just the case that individuals have dispositions to treat certain facts as reasons. Rather, there is a more direct and immediate experience of our values and reasons in the body.

3.4 Valuing as grounding being of value, an argument

Above, I claim that being of value is attitude dependent. This claim is near the core of this dissertation, and at the core of many constructivist theories of reasons. As such, let us take some time to walk through the arguments for attitude-dependence, generally, as well as my version of attitude-dependence in particular.

A quick outline of this section. First, I consider the context of the attitude-dependence claim; attitude-dependence is motivated by a puzzle about the force of reasons to motivate or even compel actions. This is both expository and argumentative, since explaining the puzzle of motivating with justificatory reasons is itself an argument against the existence of attitude-independent reasons

Second, I consider the standard epistemic arguments (Mackie's queerness arguments and the "evolutionary debunking" arguments) against attitude-

3. The promoting and protecting what is of value condition

independent reasons; I illustrate how these epistemic arguments can be fleshed out in terms of the motivational role of reasons.

Third, I sharpen the distinction between merely holding evaluative attitudes and the importance of valuing as a proper subset of evaluative attitudes. Valuing, rather than simply holding evaluative attitudes, is the appropriate way to flesh out an attitude-dependent theory. This makes reference back to the arguments presented in the preceding subsections by focusing on the motivational role, and layers on some broader considerations in psychology and cognitive science.

A core contention in constructivism about reasons is that our reasons for actions are attitude-dependent; in virtue of being attitude-dependent, they turn out to be dependent on certain features of the agents who hold attitudes. This gets into a broader point about the contingency of reasons, that the existence and nature of a reason is contingent on the agent for whom it obtains. Reasons obtain for agents “like us” in certain relevant ways.

The contingency on the nature of agents creates some problems to be addressed in chapter 6. However, because what facts will motivate agents to act depends on the psychological constitution of those agents; this is an expected problem when trying to explore the motivational role of reasons.

3.4.1 Motivating and justifying; the problem of separating the roles

One of the goals of this dissertation, and a major focus of chapters 4 and 5, is to illustrate the way that the various roles of reasons (justificatory, moti-

3. The promoting and protecting what is of value condition

vational, and social) fit together. In presenting both chapters 2 and 3, this dissertation must address why I have chosen not to treat the roles separately and develop a theory that treats justification and motivation as independent and feeding in to the various deliberative and social functions.

There is a conventional philosophical distinction between reasons that justify and reasons that motivate, as in Alvarez (2017). As the distinction goes, while the two may frequently co-occur, this need not be the case. I have considered this at some length in discussing the reasons as evidence thesis; the claim that mere pretexts (in section 2.5) are reasons is a philosophical and practical problem.

Still, this distinction seems important, especially in considering individuals who fail to act in accordance with moral reasons we believe should compel action.

There are facts that motivate without justifying; these are commonly discussed as tendencies and cravings. There are reasons that justify but do not motivate: these are what I discuss in the section 2.5 as pretexts and these are not reasons. As such, I wind up with a much more restrictive analysis of reasons, where reasons must be at the intersection of justifying and motivating (they have to have some other features, too, but the justifying and motivating roles are necessary).

In some cases, the presentation of a fact may fail to motivate action,¹⁹ even when we take the fact to be morally salient and believe it should be binding on all members of the community. I consider the most stark and disturbing iteration of such cases in section 6.3. This differs from the pretext

¹⁹What I have in mind here are cases where there is *no* motivational force. There are circumstances where there may be some motivational force, but it is very weak.

3. The promoting and protecting what is of value condition

cases, where an agent acts consistently with the reason but is not following the reason.

Consider a case common in the literature of the man who abuses his wife.²⁰ One may give justifying reasons to the man, but (the way the case is formulated) this is not sufficient to motivate the man. In the case, realist and realist-adjacent theories suggest justifications are reasons for the man to act, despite those reasons not occupying the motivational role. My approach to making sense of such a case is more complicated, allow us to consider a few alternatives.

On one alternative, the fact presented is a reason the man ought not abuse his wife, even though the fact does not motivate him. This is the view endorsed by attitude-independent theories.

On another alternative, the fact presented is not a reason that the man ought not abuse his wife, because it does not accord with anything in his motivation set. This is a view endorsed by some attitude-dependent theories.

On a third alternative, the fact presented is a reason that the man ought not abuse his wife because it accords with the values in some ideally considered motivation set. While it may not be consistent with the actual motivation set of the man, if he were to engage in reflection, then he would adopt the appropriate valuation into his motivation set.

The first is incompatible with my theory. The third is theoretically compatible, but tricky. Because my theory is dependent on valuing, rather than merely holding any evaluative or favoring attitude, in order for the man

²⁰As far as I can tell, the popularizer of this case in discussion of reasons is Simon Blackburn. Discussed in [Wallace \(2015\)](#).

3. The promoting and protecting what is of value condition

to hold the relevant value he must actually have the relevant emotional dispositions. One could flesh this out through the bodily imperatives to act, which are clearly absent (at least to some extent) in the man who abuses his wife, as he has previously been inclined to act in this morally repugnant way.

One might hold that all agents have a full range of psychological dispositions, if the dispositions are construed sufficiently broadly that they could be realized following psychological changes. That is, the abusive man may have the disposition to feel sympathy in circumstances where his neuro-psychological system is primed and altered in certain ways. But this broad sense of acknowledging the dispositional property is problematic, since we are allowing for wild difference.

This leave us with the second, which is often deeply unsatisfying. I will return to this in chapter 6.

3.4.2 Queerness, debunking, attitude-dependence

The existing literature arguing against attitude-independent theories of value might appear fairly diverse, but that literature involves variations on a theme developed by Mackie.

Mackie's argument from queerness is straightforward enough. As Joyce puts it, "In the generic sense, whenever one argues (A) that morality is centrally committed to some thesis X, and (B) that X is bizarre, ontologically profligate, or just too far-fetched to be taken seriously, etc., then one has presented a kind of Argument from Queerness." (Joyce, 2016, Supplement on Mackie's Arguments for Moral Error theory) Joyce also notes, "it is not

3. The promoting and protecting what is of value condition

entirely clear what Mackie intends” to claim is bizarre or ontologically prof-
ligate. There are a few readings of Mackie that take him to focus on the
claim that moral properties themselves are the subject of the argument; in
appropriate context (given that the rest of Mackie’s book consists of arguing
for a certain constructivist view of moral properties) that would mean some-
thing more like “naturally occurring” moral properties, or moral properties
understood in a certain attitude- or agent-independent way.²¹

Consider, though, an analogy made by Mackie, in distinguishing first
and second order moral skepticisms (Mackie, 1977, p. 15-17). Mackie makes
the analogy between his second order moral skepticism to the historical dis-
cussion over skepticism about properties, particularly color (Mackie, 1977,
p. 19-20).

Robert Boyle and John Locke called colours ‘secondary qualities’,
meaning that colours as they occur in material things consist
simply in patterns of arrangement and movement of minute
particulars on the surfaces of objects, which make them, as
we would now say, reflect light of some frequencies better than
others, and so enable these objects to produce colour sensations
in us, but that colours as we see them do not literally belong to
the surfaces of material things.

Analogously, albeit with gloss, moral properties are not properties of acts,
states, etc. *in themselves*.²² Rather, such properties are, on Mackie’s view, a

²¹Joyce gives more or less this reading in the remainder of the supplement.

²²Both pre-theoretic intuitions and linguistic intuitions may well be committed to the
view that moral properties actually are properties of acts in themselves. Mackie’s point is
that this is not tenable, despite being intuitive.

3. The promoting and protecting what is of value condition

function of a joint relation between certain properties of those things and their dispositions to cause certain reactions in us.

First, an observation novel to my version of constructivism; by exploiting Scheffler's account of valuing, which incorporates our dispositions to feel certain ways, the view considered here is already straightforwardly compatible with Mackie's view of moral properties as secondary qualities.

Second, Mackie's line of thinking here introduces a fairly straightforward and familiar problem in the assessment and evaluation of first and second order properties, of primary and secondary qualities. Perceiving that some object has a property is not clearly a reliable guide to the presence of that property; it is a reliable guide to certain features of the joint relation between the object and our perceptual system, but developing a fuller theory requires bootstrapping and scientific work.

This is where the evolutionary debunking arguments against moral realism come into play. These arguments are sophisticated, but function to flesh out the above worry regarding perception and second order properties.

Evolution includes a broad set of mechanisms by which the human perceptual system (in fact, the perceptual systems of all organisms) developed. Evolution is a process that is sensitive to a range of pressures, including sexual selection, reproductive fitness, survival, genetic drift, and so on. However, there is no reason to believe that any attitude-independent moral facts would apply pressure such that the perceptual system would be sensitive to those moral facts. The physiological and psychological phenomena implicated in perception, cognition, communication, etc. (which is a broad range) do not give us reason to think that they are tracking something exter-

3. The promoting and protecting what is of value condition

nal and, further, it is not clear what such tracking would be or how it would work.

When talking about color, the existence of a successful theory of optics gives some ability to infer physical properties of objects based on our perception of those objects. Based on what we know (via theories of optics) about reflectivity, we can make certain assessments about colors like Vantablack (a material that absorbs light) through bootstrapping. Conversely, we can make inferences about the perceptual systems of other animals based on similar bootstrapping from claims about the material to observations about the perceptual systems of those animals, the relevant similarities and differences of each.

What would such bootstrapping look like for a theory of first order moral properties? This is already familiar in the literature on moral psychology and experimental philosophy.²³ In accordance with that literature, the idea is to develop an account of agent's responses to moral properties. As with the discussion above on color, this literature is principally about the physiology and psychology in virtue of which external properties produce internal states.

²³See, for example, the anthology [Liao \(2016\)](#). Disclosure: I was the research assistant on that project.

3.4.3 Why valuing matters

... and I asked you if that had included love, which is when you said you had to get up early in the morning... Men do not naturally just not love—they learn not to.

Felix, *The Normal Heart*, by Larry

Kramer

The addition in my analysis compared to alternatives is the use of Schefler's account of valuing, rather than some general characterization of favoring evaluative attitudes. Why should the account consider valuing in particular?

Desiring, as the broad term, does not adequately explain the binding and emotional content of reasons. If it includes something like Baker's unmotivated and very weak want for a glass of water, then it will be unlikely to provide a reason unless occurring in conjunction with something stronger. This is why I supplement Baker's want for a glass of water with some condition of want satisfaction, though such a condition of want satisfaction is often implicit in such cases.

Further, there are a range of cases where simple desires are complicated by complex attitudes, and render a range of cases where an individual may have a desire for x but also have second-order desire not to have the desire for x, such that the presence of the first-order desire is inadequate to form a reason. Cases of chemical dependence, for example, are an example where an individual may have a desire for Ketamine but also have a second-order

3. The promoting and protecting what is of value condition

desire not to have a desire for Ketamine. Scheffler's account of valuing illustrates how, by failing to satisfy Scheffler's item 3, the result is a degenerate form of valuing.

While drafting this subsection, the American playwright and HIV/AIDS and gay rights activist Larry Kramer died. Obituaries of Kramer refer to his anger as a response to the HIV/AIDS epidemic, especially during the Reagan administration, when the federal government response was to ignore HIV/AIDS. Larry Kramer was the embodiment of valuing, in contrast to the presence of a simple evaluative or favoring attitude

Obituaries refer to Kramer's "anger" at the government response to the HIV/AIDS crisis. That anger is constitutive of Scheffler's item 2, experiencing "a range of context dependent emotions regarding x."

In the context of Kramer, these emotions involve sorrow at the suffering and death of the LGBTQ+ community in New York City and anger at the government's apparent indifference towards that suffering and death. Further, this is supplemented with clear satisfaction of Scheffler's items 3 and 4. He regarded his anger as appropriate in those circumstances; he regarded considerations of the welfare of those afflicted with HIV/AIDS as reason for action, from political to personal action.

It will not satisfy my point here to illustrate how Kramer satisfies valuing without presenting Kramer's emotional response in contrast to the other evaluative attitudes.

Suppose someone has a belief that the welfare of HIV/AIDS patients is good or valuable; they hold the doxastic attitudes to this effect and even the higher-order doxastic attitude that their first-order beliefs to that effect are warranted. If one thinks that the doxastic attitudes and evaluative attitudes

3. The promoting and protecting what is of value condition

are relevantly different in this context, then we can stipulate the relevant evaluative attitudes too. But leave out Scheffler's item 2, the "susceptibility to experience a range of context dependent emotions..." Does that person value the welfare of HIV/AIDS patients?

Perhaps so, but this satisfaction is less clear. If one is not moved by the suffering and death of those neglected, then this does put a significant asterisk on the extent to which one values the welfare of those so suffering.

Scheffler's valuing is useful in this analysis of reasons precisely because it ties together beliefs, emotions, and dispositions in a way that fleshes out the motivational force. The purpose of using Larry Kramer as an illustration is to show the circumstances in which individuals are disposed not just to act in accordance with what they value, but to experience and express certain emotions. Valuing the welfare of HIV/AIDS patients was not just the grounds for reasons on which Kramer acted throughout his life, but also tied to the subjective experience in virtue of which he was compelled to act, out of anger and sadness.

When philosophers writing about reasons try to give an account of how reasons bind agents to act, there is often an appeal to the motivation set (as in Williams (1979)) or the desires (as in the "Humean" constructivism of Street (2008), *inter alia*).

In my observation, such a treatment will systematically neglect (or at least understate) the varied and rich phenomenological character of the binding force. There are cases where the force of a practical reason will be dull, just characterized by the general pursuit of a desired outcome, but there are other cases where the force is tied up in a more complicated set of emotional responses and dispositions. One upshot of Scheffler's account

3. The promoting and protecting what is of value condition

of valuing is that it makes this possible variation in the motivational force explicit. This helps to explain a core phenomenological element of our reasons, that some practical reasons move us only subtly while others crash over us and may even knock us down.

3.4.4 Restrictiveness of Scheffler's valuing

The combination of the four elements presented by Scheffler creates a very restrictive analysis of valuing, though it is worth noting again here that Scheffler's description is not properly an analysis, but rather a broader account that may be subject to significant variation in the way and degree to which the conditions are satisfied.

The obvious things that are subjects of valuing on Scheffler's analysis are things we properly recognize as values, abstract objects of great personal and interpersonal importance: honesty, compassion, valor, beauty, and the like. However, as in many examples provided above, there are cases of things an agent values that are not properly values, but are things of value.

The purpose of using Scheffler's account of valuing in this dissertation is to illustrate the dispositions to motivate and to consider as reasons, and to provide this as a fleshier account of the motivational role. The purpose of this discussion is to illustrate the relation between the conventional elements recognized by analytic philosophers (like the doxastic elements in Scheffler's item 1) with the phenomenological richness of our experience of reasons and values. To have a reason is not just to have the analytic elements of the belief, but also includes a disposition; further, this disposition consists

3. The promoting and protecting what is of value condition

not just of being inclined to act but also (when conditions are appropriate) to feeling some sort of way.

Motivation is not just a dispositional state, though it includes the disposition to act. Rather, motivation includes phenomenological elements, and these phenomenological elements are part of how we recognize and experience reasons.

Anger, joy, sadness, relief, etc. are emotional responses to the satisfaction or frustration of our attitudes, especially valuing. In some cases, these feelings are subtle; in some cases, the feelings are so overwhelming that they may be the principal feature of our experience, in a given moment. As I wait to hear back from prospective employers, I am anxious and anticipate feeling either enormous joy and relief (should I be offered a job) or anger and sadness (should I not get such an offer); this anticipation of emotional responses is tied up in the way and extent to which I value the ability to continue an academic career, the things I hope to achieve in that career, and the like. It is not merely favoring, or merely the evaluative attitude that it would be good to continue my academic work, but a more intimate, personal, and rich set of emotional states.

Chapter 4

Conjoining the conjuncts

4.1 Conjoining the conjuncts

And the LORD said: 'Behold, they are one people, and they have all one language; and this is what they begin to do; and now nothing will be withholden from them, which they purpose to do. Come, let us go down, and there confound their language, that they may not understand one another's speech.'

Genesis 11:6-7, Old JPS (1917)

translation

Chapters 2 and 3 have provided an analysis of each of the two conjuncts in the analysis of reasons, with arguments for each of the conditions and exposition of their role in the existing literature.

This is necessary because the analysis itself is complicated. In pursuit of depth and precision, the analysis trades away simplicity. The analysis is conjunctive and the conjuncts themselves (as illustrated in the previous chapters) are reasonably complicated and have working parts.

The purpose of this chapter is to illustrate how this analysis functions as a whole, and then to show how the analysis can be leveraged into an even thicker, more satisfying theory of reasons.

As a refresher, the full analysis:

Some fact R is a *pro tanto* reason for an agent S to do an act of type ϕ in a circumstance C iff (1) R is evidence that S ought to ϕ and (2) R implies S ϕ -ing in C promotes or protects something of value.

One contention argued in chapters 2 and 3 is that each of the two conditions are independently necessary. In section 2.5, I argued directly that the reasons as evidence condition requires supplementation, otherwise it over-generates to include pretexts and pretexts are not reasons.

The condition that ϕ -ing must promote or protect something of value has not been defended, and its defense is more complicated. As I unpack the comprehensive analysis in this chapter, one might be tempted to suppose that the reasons as evidence condition can be subsumed into the promoting and protecting condition, that every instance of promoting and protecting something of value is evidence that one ought to ϕ . I provide an argument in section 4.4 that this will not do, that in fact promoting and protecting something of value also over-generates, though in a more subtle way than the extensional arguments in sections 2.5 and 2.6.

The chapter opens by providing some illustration of the combined analysis, showing how the conditions work together in their functional roles. The chapter develops by illustrating both the internal justification and motivation as well as the social dimensions of giving and sharing reasons. Giving and sharing reasons is the more direct object of discussion in section 5.1, but will be useful as an illustration.

4.1.1 Fitting this chapter into the broader project

The purpose of this chapter is to provide a series of direct arguments against reductionist analyses of being a reason. A reductionist analysis is one that takes the justificatory and motivational norms and combines them into a single condition, for example holding (as some realist theories do) that justification fully explains motivational. Some versions of this argument are presented earlier in the dissertation, in sections 2.5 and 2.6.

The major focus of this chapter is on the prospect of subsuming an entire analysis of reasons into the promoting and protecting condition, holding that the reasons as evidence condition can just be a corollary of the promoting and protecting condition and therefore that the promoting and protecting condition is necessary and sufficient. This is argued through two theses.

- There are cases where S ϕ -ing in C will promote or protect something of value, but there is not a reason for S to ϕ in C, because there is some strictly better course of action ϕ' which is covered by all of the *pro tanto* reasons that are ostensible reasons to ϕ . See section 4.4
- There is added explanatory value in maintaining the independence of the reasons as evidence condition. See section 4.5.

The first of those items is intended as a traditional counterexample to a theory of reasons which treats the promoting and protecting condition as necessary and sufficient. The second of those items is a comprehensive argument that suggests the improved explanatory value of the theory is

reason to maintain both conditions. That argument can be summarized as follows.

- A theory of reasons is best if it explains both the justificatory role and the motivational role, and provides an illustration of how those two roles interact.
- Constraints provided in accounting for promoting and protecting do not provide an adequate explanation of the justificatory role.
- It follows from the inadequacy expressed above that some independent expression of the justificatory role is required to satisfy the assessment criteria.
- Because of discussion outlined in chapter 2, the explanation of the justificatory role is best served by maintaining $R \Rightarrow E$.

The purpose of providing a separate illustration of these arguments is to clarify the way that these arguments fit with the broader scope of the dissertation. Hopefully, it is clear that a major purpose of this chapter is to advance the claim that we need *both* the justificatory *and* motivational condition to have an adequate analysis of being a reason. This is a core contention that I also advanced in section 1.4.4.

It is also important to illustrate where these claims, especially the argument for the inadequacy of motivational analyses with constraints, fall in the course of this dissertation and how they feed into the next chapter. At points, the flow of this dissertation may seem more thematic than argumentative, but there is some method to the layout. Thematically, the purpose of

this chapter is to provide connective tissue between the abstract characterizations in chapters 2 and 3 and the fleshier exploration in chapters 5 and 6. Argumentatively, the goal is to lay the groundwork for the major practical argument in favor of my analysis of reasons: An analysis that combines the justificatory and motivational roles is best equipped to explain the diverse social roles of reasons.

- **Thesis:** Reason giving and social interaction (as interpersonal practices) involves considering both justifying and motivating. (This sets up the master argument in chapter 5, which is an exploration of how the combined analysis provides a deeper account of both.)
- Motivating behavior is one reason for (pardon me) the practice of giving reasons, because we want to influence behaviors. (Explored at greater length in section 5.2.1.)
- Giving reasons also involves the normative assessment of acts and the quality of reasons themselves. This is informed by the justificatory (and evidentiary) norms.
- Giving reasons involves intersubjective consideration, where individuals share *both* justificatory norms (which are socially recognized, if not constructed outright) as well as relevantly similar motivation sets, and this provides explanatory grist for social interactions. (Introduced in section 4.5 and handled at greater length in sections 5.4 and 5.5.)

The use of the practice of “reason giving” provides both the thematic and conceptual bridge to the richer social practices taken up in chapters 5 and 6. When we give reasons to each other, we are engaged in a social practice;

some instances of the social practice are minimal enough to allow me to bring out the abstract elements of the analysis of reasons introduced in the preceding chapters of the dissertation, while still being clearly social.

In any case, hopefully this provides some illustration of how this chapter fits into the dissertation, and smooths out some potentially confusing choices in the writing.

4.2 Presenting the analysis and giving reasons

Suppose David is deliberating about whether to enroll his child in Hebrew school. David values his connection to his culture; he wants to impart that value to his child and promote the value of connecting to one's culture generally. This valuation, at least in the context of David's present deliberation, circumstances, and attitudes, has no relevant defeaters.

Enrolling his child in Hebrew school will promote the child's connection to the family's culture and generally improve the community, promoting what is of value more broadly.

As David looks at the website for the Hebrew school affiliated with his congregation, he sees the range of activities that promote both religious holidays but also broader cultural activities. The fact that the Hebrew school has events education and celebrating his own Ashkenazi (eastern European) Jewish culture, as well as middle eastern, north African, Western European, etc. cultural traditions, is a reason to enroll his child in the Hebrew school.

The existence of those activities implies the promotion of the related things of value. The existence of these events is also evidence that David ought to enroll his child in Hebrew school.

In this case, the explanation of how the reasons as evidence thesis functions can directly make use of the story about promoting and protecting. Why is the fact that the Hebrew school hosts these activities a reason to enroll? Because it promotes the connection with cultural identity and this is of value. The fact that there are cultural activities is a *pro tanto* reason for David to enroll his child in Hebrew school.

The broad role of this case is just to provide an illustration for how the analysis may proceed under ordinary conditions.

In this case, the role of the promoting and protecting condition and the reasons as evidence condition are closely related. Under ordinary conditions, the two conditions can be run more or less in tandem in the way illustrated. However, there are going to be cases where the two are not so closely tied together.

Consider the case of Ananya and the fire alarm. The fire alarm going off is evidence that Ananya ought to leave the building. However, there are a range of different things that might be promoted by complying with the fire alarm. She may not know which of the things of value is being promoted. It promotes a general regard for public safety by showing that the fire alarm is generally (admitting of exception) reason-conferring. The alarm is evidence that Ananya ought to evacuate, but that evidence may be informed by a handful of different values and which values are promoted or protected may vary across circumstances.

In the case of an actual fire, the thing being promoted is her own safety. But suppose the instance is a fire drill. In such a case, complying with the fire alarm promotes certain values to the community by providing a test of orderly evacuation procedure and the potential logistical issues,

keeping people calm, etc. Ananya may not know that the fire alarm going off promotes or protects those other things of value, having only the evidence that she ought to evacuate the building, but she can be reasonably sure under most conditions that it will promote or protect something of value, even without clarity about which of the values is being promoted.

There are false alarms; there are cases of failed indication.

My stance is: there are cases where the fallible evidence that an agent ought to ϕ is present, but the evidence is errant and it turns out that S ϕ -ing in C does not promote or protect anything of value. In such cases, there is not such a reason.

An interlocutor may argue that the case of Timmy pulling the fire alarm is a case where *there is* a reason, and take this to be a counterexample. Such a case pivots on the intuition that the fire alarm going off is itself a reason, and remains a reason independently of what causes the alarm to go off. This is consistent with the reasons as evidence analysis advanced by Kearns and Star and satisfies the reasons as evidence condition in my analysis, but is a counterexample by virtue of failing to satisfy the promoting and protecting condition. I argue below that this counterexample only works by fudging whether the case satisfies the promoting and protecting condition.

On the counterexample, one might hold that when Timmy pulls the fire alarm, Ananya exiting the building still promotes compliance with the norms of fire safety such that those things are of value. One might even go further, and say that all cases where there is fallible evidence S ought to ϕ in C entails that S ϕ -ing in C will promote or protect something in value on the basis that it promotes responsiveness to evidence, and responsiveness to evidence is a thing of value. However, this would just be to take a

different route to argue that the case in fact satisfies the promoting and protecting condition (by way of promoting compliance), and is therefore not a counterexample to my theory at all.

Alternatively, in a case where all of those values are moot and the only one at play is Timmy getting out of a test, then the case judgment gets much less clear. Suppose that, as a matter of fact, complying with the alarm when Timmy is trying to get out of a test *does not* promote compliance, fire safety, or anything else. Rather, it is either neutral or worse with regard to all of the implicated senses of promoting and protecting discussed in section 3.3. Under such conditions, do we still want to maintain that the mere presence of the evidence (despite that evidence being errant in this case) is a reason?

If one thinks mere compliance with the alarm as a matter of safety itself provides a reason, then an interlocutor might initially accept the test evasion case. But when we stipulate away the promotion or protection of those values, does that intuition persist? Is it dumbfounded?

Consider this in accordance with the discussion of the only-if factual constraint discussed in section 3.3.4. Ananya holds that the fire alarm going off is reason for her to evacuate and is motivated to do so in accordance with her desire to preserve her personal safety, either directly by getting away from an active fire or by ensuring that the safety systems of the university work. Suppose that this is the only thing she values relative to evacuating the building, the only thing of value at play in this case pushing against staying in the office, ignoring the alarm.

If we grant that no such values are promoted or protected by evacuating (thereby establishing that the case will fail to satisfy the promoting and protecting condition), then the claim that the fire alarm is a reason is dubi-

ous. After all, if Ananya knew the facts, then she would not be motivated to evacuate.

What we are going to wind up with here is something of a variation on the non-motivational reasons cases presented in sections 2.5 and 2.6, where the evidence that one ought to ϕ is present but the motivational force of that reason is gutted. In this case, though, the motivation is gutted by the failure of the evidence to track (in this particular instance) with the facts. Such a reason violates the factual constraints such that the act promotes or protects something being of value. The way to circumvent this case is to hold that maybe there is some weak remnant of value in mere adherence to the evidence even in a case where the evidence failed to track the truth, but this move misses the central premise of the case, which is to limit the scope of the values at play. As such, this counterexample is unsuccessful, because it requires equivocating on the nature of the case; in particular, it equivocates over whether Ananya's evacuating the building promotes or protects something of value, and therefore whether the second conjunct is satisfied.

4.2.1 Justifying behavior and reason giving

Consider David's Hebrew school case, in the context of a social interaction. David is explaining to his partner why he believes they ought to enroll their child in Hebrew school. He explains the reason above, appealing to the value of cultural identity.

In this case, if David presents the reason and things of value as he understands, then that may not be sufficient to motivate his partner, but there

may be other evidence and other things of value, even if those are not motivating to him, which provide motivational grist for his partner. Under these conditions, David may talk about the fun the child will have in the school, making friends and socializing, and learning a second language. David and his partner may have different motivational inclinations with respect to this set of reasons, may be inclined to act differently, but may still come to the shared conclusion that they ought to enroll their child in Hebrew school.

We can run roughly the same story above for David's partner based on different reasons

My point here is that in offering a range of facts that may be reasons for his partner to act jointly in enrolling their child, David is engaged in a practice of giving reasons. These reasons may not be reasons *for David* to enroll their child in Hebrew school, but may well be reasons *for David's partner* to enroll their child in Hebrew school. We can give reasons to others, even if those are not our reasons for action. I will explore this point at greater length as a matter of explanatory upshot in section 5.5 and in discussing counter examples.

In this case, we may suppose there are two facts F and G. F is the aforementioned reason for David. G is some reason for David's partner. In this case, both David and his partner agree on what they should do, even if their actions are responsive to different reasons. They coordinate behavior by acknowledging each others' reasons for acting.

However, they need not even be responsive to different reasons; they may just weight the reasons differently or find different reasons more compelling. It may be that both David and his partner share reasons F and G, but that each is slightly more responsive to one reason or the other.

In the practice of giving reasons, one may foreground either the promoting and protecting condition or the reasons as evidence condition. This is a useful feature both of the analysis and of our ordinary discussion of reasons.

As presented above, David giving reasons for his partner to enroll their child in Hebrew school is treated as an evidentiary exercise. G is a reason for the partner to enroll the child in Hebrew school in that it functions as evidence that the partner ought to enroll the child in Hebrew school. But, as illustrated in section 2.5, this does not imply that the fact will satisfactorily motivate the agent.

Alternatively, we may foreground the promoting and protecting condition by giving an analysis similar to the one provided in the preceding subsection for David. There is something that the partner values, and the partner considers those values in considering G; G implies the promotion or protection of those values, and therefore this motivates the partner to act in the relevant way.

Which condition we foreground in a particular instance will depend on the context. Sometimes it is appropriate to highlight motivation; sometimes it is appropriate to highlight justification. The theory allows for us to consider each based on present needs.

4.3 Interdependent evidence and promoting and protecting cases

On many cases, one might reasonably suppose that the promoting and protection condition is monarch and that the reasons as evidence condition

can be treated as a corollary that is entailed by the promoting and protecting condition.

One characterization of the move is to suppose there is a reason for S to ϕ in C iff S ϕ -ing in C promotes or protects something of value, and that the existence of evidence that S ought to ϕ in C is merely the presence of evidence that S ϕ -ing in C will promote or protect something of value. Instances where S ϕ -ing in C would promote or protect something of value is also generally evidence that S ought to ϕ in C.

Even if we do not know what thing of value is being promoted or protected, as in the fire alarm cases, the existence of the fire alarm is evidence that evacuating promotes or protects some thing of value. It may be the public safety benefit of the fire drills; it may be the preservation of personal safety from an actual fire. The fact that the alarm is going off provides evidence that S ϕ -ing in C promotes or protects something of value.

For the remainder of the chapter, I argue that one should not take the subsuming approach. Instead, we should preserve the reasons as evidence condition as an independent, necessary condition for being a reason.

Still, I understand the temptation towards simplification and want to assess how that temptation is produced from a few different angles.

The first is the straightforward theoretical angle, the idea that an analysis that has only one of the conditions (rather than more) is better, all other things being equal. Such an analysis would be more parsimonious, easier to understand and illustrate, and therefore just better as a matter of our shared values in assessing philosophical theory. I am sympathetic to this argument, but “all other things being equal” is doing a lot of work in that phrasing and my argument in section 4.4 illustrates that all things are not equal.

The challenge is that it is not so clear that a theory which subsumes the reasons as evidence condition into the promoting and protecting condition actually produce adequate explanations. In the fire alarm cases presented above, providing the analysis solely in terms of promoting safety or community values in orderly, safe evacuation fail to capture the role of fire alarm as indicator of a set of values that may be promoted or protected. In some cases, one or more of the things of value may be absent or irrelevant and it can still be the case that the reason obtains. Even if Ananya is unclear about what the relevant value is, she sees the fire alarm as justification for leaving and recognizes that she ought to leave under most conditions where the alarm is going off.

On those cases, because the fact being evidence that Ananya ought to evacuate is independent of the various values being promoted or protected, a theory that focuses only on the latter obscures how the reason works. That makes for an inadequate theory.

The second angle, which is of more interest to those working on theories of reasons, is in holding that one of the justificatory role or motivational role of reasons is prior to the other. On certain reason internalist approaches and constructivist theories, justification is viewed as parasitic on motivation. The fact that there are cultural events at the Hebrew school justifies enrolling is because it motivates David to enroll his child, based on his motivation set, projects, or whatever term we like to delineating the relevant psychological attitudes. Alternatively, albeit in reverse, on some realist understandings of reasons, the fact that the Hebrew school holds such events justifies him

in enrolling his child and this should, at least on a *pro tanto* basis, provide some motivation.¹

Assumptions of priority are dangerous in this analysis precisely because there are cases of separation in each direction. There are cases of justification without motivation. There are also cases of motivation without justification. Of course, individuals in the realist tradition will criticize various constructivist theories on the grounds that it is just apparent that we have motivation without justification, where motivation is degenerate, but to do so is to mischaracterize constructivist theories. No defensible constructivist theory holds that brute motivation is itself adequate for the existence of a reason; e.g. Williams' analysis² requires some condition of possible deliberative reflection and other approaches focused on motivation similarly impose constraints.

Set aside those strawman attacks on constructivism as emphasizing a constraint-free account of reasons.

Simplicity is a persuasive value in the construction of theory, and so there must be strong reasons given not to simplify the theory. I offer two reasons, one in each of the following sections.

Section 4.4 argues that the reasons as evidence condition actually does constrain an extensional difference that is important. The promotion and protection condition would admit of the existence of facts as reasons when they are not evidence that S ought to ϕ , and this would produce an extensional mistake. Admittedly, this argument is not as strong as the other

¹I say "slightly differently" here because of the work "should" is doing. Realist theories acknowledge that while F may justify S ϕ -ing in C, this justification may not in fact motivate S to ϕ in C.

²See Williams (1979).

arguments for extensional difference developed elsewhere in the dissertation. There are also prospective ways of fleshing out my theory which would make it possible to dissolve these cases into the theory without issue.

Section 4.5 argues that there is added explanatory value in maintaining the independence of the reasons as evidence condition, irrespective of whether it can be given as an entailment of the promoting and protecting thesis. This move represents a major pivot within my theory and the defense of that theory: the best argument for the theory of reasons I advance in this dissertation is its explanatory power. By maintaining the justificatory and motivational roles of reasons separately, the theory can directly account for each and incorporate each role into the theory.

I develop the explanatory power at much greater length in discussion the application to the social roles of reasons in chapter 5. While this chapter focuses more on defending the added explanatory value of the reasons as evidence thesis in particular, that is also a subordinate theme in the following chapter, as a part of the full analysis.

Even if one finds the extensional argument unconvincing, the explanatory value argument provides strong reason to avoid simplification.

4.4 Strictly better acts and the reasons as evidence thesis

It may seem intuitive to suppose that if F entails $S \phi$ -ing in C promotes or protects something of value then that implication is in itself sufficient, that in itself is evidence that S ought to ϕ . But this is not so.

Suppose there is a fact F present for an agent in a circumstance, as that agent is determining whether to ϕ . F implies that $S \phi$ -ing in C promotes something of value V . However, F also implies that S doing some alternative act ϕ' in C would promote V at the same level *and also* promote some further thing W .

It would be imprudent to develop a comprehensive schema of the notion of *strictly better*. The idea is roughly something like this: if two things x and y are equivalent in all ways except one or more things of value which favor x , then x is strictly better than y . In this case, ϕ' is (ex hypothesi) strictly better than ϕ . So, while F implies that $S \phi$ -ing in C promotes V , that is not sufficient to make it evidence that S ought to ϕ in C , because there is a strictly better alternative.

Strictly better is a notion borrowed from the world of gaming.³ In chess, a queen is strictly better than a bishop, because a queen can make all of the moves of the bishop. There is no move made by a bishop that cannot be made by a queen; there are moves made by a queen that cannot be made by a bishop.⁴

³“Gaming” here is an American colloquialism that refers broadly to all sorts of board, card, computer, and role-playing games; this is distinct from the more narrow term that refers only to gambling. “Gaming,” in my use, includes but is not limited to, gaming. I owe the notion to broad public discussions by the game designer Mark Rosewater, and to some private conversations with fellow gamers.

⁴Two short observations. First, a queen is not technically strictly better than a rook, because a rook can castle and a queen cannot. Through most of the game, in most circumstances, a queen is strictly better than a rook. In the endgame, well after both players have castled and traded for a while, we might reasonably say that a queen is strictly better. However, there are some circumstances where they are just different. Second, and more starkly, a knight is not strictly worse than a queen. While a knight is considered (rightly) far less valuable than a queen, a knight is just different, in that its movement patterns vary considerably. To be strictly better is to be the same *except* for the respects in which one is better.

The argument is that F is not evidence S ought to ϕ in C, because F is only evidence in favor of the strictly better action. F is evidence that S ought to ϕ' in C. In C, F should not result in our deliberative consideration of ϕ -ing; we should not count F as in favor of ϕ -ing, because the extent to which it should lead us to favor ϕ -ing should just be disregarded in favor of ϕ' -ing.

Suppose we look at all circumstances where S moves a pawn to the back row and is entitled to exchange their pawn for a different piece. Suppose it is a fact that S prefers to use diagonal movements in the endgame; call this fact F. Given F, S replacing his pawn with a bishop in any pawn-exchanging circumstances promotes or protects something of value. The fact is never evidence that S ought to exchange his pawn for a bishop, because the bishop is strictly worse than the queen in all such circumstances. While it may promote or protect something of value, F cannot be evidence for the bishop-exchanging move because the move will always be a deficient version of the queen-exchanging move.

One way to argue this thesis is to hold that when we fix the ordered triple of agent, act, and circumstance, we take it that (for these strictly better cases) the circumstance type entails the possible strictly better action. This is where the chess case is useful, because it is just a feature of the rules governing chess that promoting a pawn always includes in the possibility of promotion to queen (as well as bishop, rook, etc.). Since it is a fact about C that it entails the possibility of ϕ and ϕ' , it follows that in those cases the strictly better alternative is an option in those circumstance types. This suggests that evidence that S ought to ϕ in C will always be undercut by the more decisive course of action ϕ' . This does not quite resolve the case,

but positions the objection on much stronger terms, since it positions ϕ' as possible action in all circumstances of type C.

The major prospective response for those who think that the reasons as evidence thesis is just a corollary is to say that, in these cases, the fact that having the diagonal movement is beneficial to the endgame does still provide evidence that S ought to exchange his pawn for the bishop. There is still a *pro tanto* reason for S to exchange his pawn for a bishop. It is just the case that, because of the rules of chess, he will never have *most* reason to exchange for a bishop, because the queen will always be better.

So, on this construction, it is important to note that ϕ , the strictly worse option, is never the thing that the agent ought to do, but perhaps one might argue that there can still be *pro tanto* reason for S to ϕ . This is an admittedly appealing response; after all, there can be evidence for falsehoods (perhaps even necessary falsehoods). The fact that it is necessarily false (for all possible worlds where the rules of chess are this way) that S ought to exchange his pawn for a bishop does not mean there cannot be evidence that S ought to exchange his pawn for a bishop. For those who want simplification, this is a promising route.

As I noted above, I do not think these strictly better cases are decisive counterexamples to an analysis that subsumes reasons as evidence into the promoting and protection condition. I recognize that some individuals will find it perfectly reasonable to suppose that the strictly better cases have a *pro tanto* reason for S to promote the pawn to a bishop, and see no problem here. I have tried to illustrate some strong positioning to denying that there is a reason to promote the pawn to a bishop in the strictly better case.

There is a separate, more practically oriented reason to maintain that this difference in extension is a mistake (or, at least, not an optimal choice in developing a theory of being a reason). If we want a theory that incorporates an account of reasons as justificatory, then it is important such a theory reasonably set constraints which exclude associations between facts and extraneous acts that are strictly worse alternatives to obvious choices. For any given act type ϕ , there are a range of ways of doing that act which result in strictly worse outcomes. Suppose I have reason to wash the dishes; does it follow that I also have reason to wash the dishes while hopping on one leg? This seems like a strictly worse course of action (though there may be some odd attitudes and circumstances where it is not strictly worse) and it would be odd to suppose the reason that I have for doing the dishes is also a reason to do the dishes hopping on one leg. Again, perhaps this is not as weird an implication to some other philosophers as it appears to me.

But in favor of my view here is the ordinary constraints on justificatory practices. Suppose that my roommate Alican walks in while I am doing dishes, hopping on one leg. He asks me for a justification for this strange practice, and I respond by citing my reasons for doing the dishes. Alican would, reasonable, raise an eyebrow and point out that is very weird.

A suspicious philosopher might think this a linguistic sleight of hand, that the reason Alican's inquiry is not answered is because he is really asking why I am hopping on one leg and *not* why I am doing the dishes, so this should be treated disjunctively. But this is not a sleight of hand; rather, it is a point about the strength of relevance relations in applying to reasons. The facts that are reason to wash the dishes are constrained by the relevance relations such that they are just reasons for me to wash the dishes, and not to

wash the dishes in conjunction with some other act that makes dishwashing more difficult. This is part of the justificatory role that reasons play; it is not some obscure trick of the light. One might expand the extension of being a reason to include some cases, but to do so creates the possibility of missing the way that relevance relations restrict what acts are justified by those reasons. This observation about the role of relevance is a broad point about the role of justification; it also has various practical and moral consequences, as we will see in the following subsection.

4.4.1 Empty-track trolley problems and our values

Consider an alternative case. Imagine the standard switch version of the trolley problem: The agent is a bystander beside a switch that can change the track. The trolley is hurtling down the track towards five people and will kill those people if not averted. The trolley could be redirected down a different track where it will only kill one person.

But now let us modify the experiment slightly and suppose that there is a third setting for the switch. If the agent presses the third setting, then the train will redirect the train down an empty track and bring it to a safe stop. Suppose further that there are no other salient differences between the directions, that this will not even make the passengers on the trolley late by comparison to the alternatives.

On this version of the trolley problem, moving the switch to only hit one person protects human lives, reducing the net lives lost by four. This is a standard improvement and familiar as the consequentialist approach to trolley problems in the first place. Moving the trolley to the track with one

person on it protects something of value, but it seems wrong to say that there is a *pro tanto* reason for the agent to move the train to the track with one person when there is an empty track available. In this circumstance, the empty track is strictly better.

It is false that S ought to move the trolley to the track that will kill one person; depending on one's view of the modality, it may even be necessarily false. My point is not one about necessity, here, except insofar as the act types are baked into the circumstances. When one has the option to harm no one, then one is not permitted to perform a strictly worse action that will harm someone. At least under the parameters of our generally accepted value set, this is just impermissible.

One potential response to these cases is to hold that the fact F is a *pro tanto* reason for S to ϕ in C and F is a *pro tanto* reason for S to ϕ' in C. But the purpose of the strictly better articulation is to illustrate that this just is not the case. There may be an alternative circumstances, call that C', where F is a *pro tanto* reason for S to ϕ in C; in C' there might be other values at play or ϕ' may not be an available course of action. But the point is precisely that those circumstances would be difference and (unsurprisingly) the reasons at play in those circumstances would be different. If we were in the standard switch version of the trolley problem, for example, there there would be a reason to move the train to the track where it would kill one person (rather than five); but in the empty track version, such a move is indefensible despite the fact that moving to the track with one person does reduce the harm (relative to allowing the trolley to run its course and kill five).

The empty track version is the morally salient version of the argument presented in the preceding subsection. Moving the train to the empty track is the strictly better version. The fact that redirecting the train and saving the lives of the five people protects something of value does not provide evidence that the agent ought to redirect the train to the track with one person. That fact only provides evidence that the agent ought to redirect the train to the empty track.

4.4.2 Strictly better cases and the corollary approach

I developed these two thought experiments based on an old experiment in psychology effectively utilized by several philosophers, most notably Chang.⁵ In the experiment, the agent is presented with a choice between three options: a trip to Paris, a trip to Rome, and a trip to Rome (or Paris) with free coffee or some other such perk. On this analysis, it is taken that the last option is a strictly better alternative to one of the two presented. One might expect that S may reasonably choose Paris or Rome with the perk, as Chang discusses that the addition of the perk does not change the preference of Paris to Rome (or vice versa) when there is a pronounced preference. It is weird and arguably outright unreasonable, for the agent to choose Rome without the perk when Rome with the perk is available. This is where the idea of a strictly better option is useful.

I will say first that I understand the drive towards simplification of the analysis when one looks at comparative cases like this, especially those structured like the Paris and Rome vacation and the empty track trolley

⁵See [Chang \(2016\)](#), which also surveys the broad use of comparative decision models in moral and social psychology.

problem. In those cases, one might suppose that the *pro tanto* reason for S to choose the strictly worse choice (like Rome without the perks) remains, even if it is precluded by some constraint on justification in that case. After all, the *pro tanto* reason for S ϕ -ing in C may be overridden by lots of other considerations, including but hardly limited to other reasons and actions plans; that's what it means for the reason to be *pro tanto*, rather than *pro toto*.

The chess case is only very subtly different; in the chess case, the strictly better choice is a feature of the act type (pawn exchanging) such that some acts in the type (exchanging for queens) will always and necessarily be better than the alternative (exchanging for bishops), and will always and necessarily be present. But the difference between those cases where the strictly better choice is set by the circumstances (like the Rome and Paris vacation choice) and the cases where the strictly better choice is set by the act types (like the chess case) is very minor variation. One might be justifiably skeptical that there is anything interesting about the distinction there, subject to one's background assumptions about the metaphysics of act types.

This leaves those in favor of simplification in an apparently strong position. Just hold that the fact is evidence (and a *pro tanto* reason) for the strictly worse choice, but that as a matter of circumstance, it will be evidence for a proposition (that S ought to ϕ in C) that is false *pro toto*. That position gets to have our *pro tanto* reason and our constraint on the rational and reasonable decision making.

So, why doubt this move at all? Because there are two other commitments that I hold independently of but relevant to my theory of reasons that

make me hesitant. For reasons of length and relevance, I will not defend these commitments here; one might reasonably reject them and hold on to everything else I say about reasons in this dissertation. Though, as I noted in the preceding subsection, I think one ought not do so as one jeopardizes parts of the justificatory role.

The two other metaethical comments are to particularism and comparativism about reasons.

The first is a commitment to particularism about reasons, for reasons discussed by [Dancy \(2017\)](#). On this view, some fact being a reason depends on the facts about the particular instance in which the fact is situated, rather than how it might be situated in other circumstances. On such a view, whether some fact F is a reason for S to ϕ in C is not just about the relationship between F and the act type ϕ , or even the fact, agent, and act type. Rather, it must consider the circumstances, including other possible courses of action in those circumstances. As such, whether F is a reason in (for example) the vacation choice case is not independent of the third, strictly-better option.

The second is a commitment to comparativism and the idea that what is of value is itself subject to the circumstances of the choice situation, including what other possible acts and things of value are at play. I indirectly discuss this position and some motivation for it in section [3.2.3](#), as consideration of the moral tragedy cases involves the absence of certain choices and therefore provides (or, perhaps, merely makes more visible) reasons for acts that normally would not have such *pro tanto* reasons. Those reasons only appear in certain choice situations.

It may be possible to reconcile my analysis, these commitments, and the claim that the reasons as evidence thesis is entailed by the promoting and protecting condition. I have no position on this and do not intend to develop one in the near future, though it is a potential direction further down the road, if my analysis continues to develop and there is interest in this question.

4.5 Shared evidence and its explanatory value

The more substantial and stronger argument in favor of treating the two conditions holds that the reasons as evidence thesis should be treated separately because the thesis provides added explanatory value.

At a very broad level, it holds that maintaining the justificatory role independent from the promoting and protecting condition helps to establish the formal and substantive constraints as external to the attitudes of the agent *and* (more importantly for present purposes) illustrate how reasons are shared by several agents even when there is divergence between or uncertainty about the attitudes of those agents.

Both of these are treated at length through the next two chapters, but it is useful to survey the way that the argument works and ties in to maintaining the reasons as evidence thesis as a separate, necessary condition.

Suppose that there is some fact F about ϕ -ing in C , and suppose that S and T are two different agents and each is determining whether to ϕ in C . S is determining whether to ϕ in C ; T is determining whether to ϕ in C . S and T are psychologically different and so what they value may be different. Suppose there are differences (large or subtle) in the values underlying what

motivates S and T to each ϕ in C, but both recognize F as evidence that each ought to ϕ in C and F does entail the relevant promoting and protecting condition for each S and T. In such a case, even if the underlying values that ground S and T being motivated to ϕ in C, F is still a reason for each to ϕ in C. Further, they can share this reason, in discussing and evaluating their justifications and motivations.

There are a few components going on in this case, but I want to suggest first that this is a very common sort of case. There are lots of instances where individuals with different values still share reasons. Consider individuals who take the existence of laws about safe driving to be reasons to drive in the ways prescribed by those laws, like following the speed limit, advertising turning signals, etc. Individuals may have significantly different values in virtue of which the evidence accords with the things they value. S may follow the law because S values the practical protections of public safety the law promotes. T may follow the law out of a broader valuing of adherence to authoritative law generally.

Externalists about reasons have appropriately noted that appealing directly to F is useful for the purposes of discussing justification. Why are both S and T are justified in signaling before turning, even if the underlying internal states in virtue of which each is motivated to signal before turning is importantly different? Because F is evidence that each ought to signal before turning.⁶ One significant area of the reasons as evidence condition

⁶Usefully, when we see these splits in values, there will be extensional differences between the circumstances in which there is a reason for each to signal before turning. On empty roads, there may be a reason for S to signal, given the things of value to S, but not a reason for T to signal. This difference is based on the satisfaction conditions of their attitudes.

is that it provides explanatory grist for how such reasons are shared even when the values differ.

I note all the way back in section 1.2 that we should treat reasons as facts because it allows us to talk about shared reasons. It allows us to talk about how Steve's inclination towards the King's Indian Defense provides a reason for Steve to make certain moves in the chess tournament *and also* for Tonya to make certain (different) moves as Steve's opponent. Similarly and importantly, this applies to S and T above considering the traffic laws in their movements.

I build up the explanatory scaffolding for how these features develop into an account of the social roles of reasons through all of chapter 5. The purpose here is to provide a simplified illustration of the explanatory value added by the inclusion of the reasons as evidence thesis, that the thesis captures something externalists about reasons get right: the shared reference to reasons as facts is integral to the explanation of their role as justifications (and, ultimately, to justifying as social exercise and ground for a set of the social roles).

F is a reason for both S and T to ϕ in C, even if the underlying values are different. Explaining how F is a shared reason, and how it works as such, is crucial to explaining some of the things that reasons do, especially in social contexts. The reasons as evidence thesis helps to make sense of this without having to do the nitty-gritty of mechanics of each of the values of S and T and how those values are related to the same fact, and how deliberation about that same fact, in cases where we are not really interested in the lower-level values. Such cases may include discussing justificatory norms within a

society, where S and T are both adhering to a particular norm for justifying behavior irrespective of how that norm tracks with the lower level values.

In Parfit's *On What Matters*, he uses the metaphor of climbing a mountain and that each of the prominent approaches to moral theory (consequentialism, deontology, virtue ethics) are developing towards the same summit (Parfit, 2011, Sec. 64). This metaphor implies (by design) a substantial claim, that the moral theories "correctly" understood will become coextensive in their judgments about cases. I am skeptical of this substantial claim, but my thought here is a more tentative iteration of Parfit's epochal assessment of the moral state of play, that there are cases where reasons are coincident even though constructed in different ways.

Parfit is emphatically not a constructivist. In fact, he is perhaps the most prominent contemporary example of a realist. The point that he draws in his analysis of the moral landscape, and his metaphorical discussion of mountain climbing, is the notion of convergent reasons. Some reasons are shared by agents even when those agents arrive at the reasons (on my view, arrive by constructing those reasons) in substantially different ways. The difference is tentativeness about the expectation of convergence in important cases. On my view, it is not possible to merely anticipate convergence as an outcome of progress in moral theory (or theory pertaining to any other domain of reasons).

An upshot of constructivism in general, and my analysis in particular, is providing the means to work towards converging reasons in domains where such reasons are desirable (for reasons of social coordination discussed in the following chapter). The upshot of constructivism is that it provides the claim that reasons can be brought about, rather than merely discovered.

4. Conjoining the conjuncts

Because reasons can be constructed, convergent reasons can be brought about in socially conscious ways. It is my view, as I will explore in the next chapter, that my analysis of reasons is especially well situated to construct such reasons because it provides independent accounts of the justifying and motivational roles of reasons.

Chapter 5

Added value and the social analysis of reasons

5.1 Social roles of reasons and explanatory value

If reason is to be realized in the sensuous world, it must be possible for many rational beings to live together as such; and this is permanently possible only if each free being makes it its law to limit its own freedom by the conception of the freedom of all others.

(Fichte, 1889, p. 137)

This chapter illustrates a major advantage and argument in favor of the theory of reasons I have presented in this dissertation. Reasons play a range of roles in our social interactions. We give reasons to communicate about our behavior and to influence the behavior of other members of our communities; we consider the reasons of others in evaluating and coordinating with the behavior of other community members.

Following the pretext arguments advanced in section 2.5, I argued that certain justification-focused accounts of reasons are deficient with respect to their extension. Section 5.2 illustrates this extensional difference and the relevant case judgments as applied to social cases, in contrast to the merely deliberative cases developed in the pretext argument. In continuing to develop the line, section 5.2.1 argues that while expressivist theories of reasons (for example, the theory developed by Gibbard (1990)) do not have the clear extensional issue, they are often explanatorily deficient, or at least

less focused when it comes to developing the constraints on the existence of a reason and on providing a treatment of the various social roles of reasons. Put another way: such theories of reasons do not adequately capture the way in which reasons are used to explain, influence, and coordinate the behaviors of other community members.

The argument goes: because my analysis of reasons accounts for both justificatory and motivational roles of reasons, my analysis provides insight into how each of those two roles is taken up in the various social roles. Some other camps approaching reasons fail to adequately consider one of either the justificatory or motivational role of reasons. Certain expressivist theories reduce the justificatory role to mere constraints on motivational reasons; certain realist theories ignore motivation outright in favor of justification. Arguments against these approaches are considered in section 1.4.4 and chapters 2 to 4. When these competing theories fail to consider both justificatory and motivational functional roles of reasons, they have explanatory deficiencies.

The core argumentative goal of this chapter is to illustrate the explanatory value of a theory of reasons that includes both the justificatory and motivational roles (and makes each explicit). These explanatory improvements provide a reason (pardon me) for adopting my theory of reasons rather than the relevant competitors. At some points in the chapter, I note a relevant similarity between my approach and the iterations of expressivist theories of reasons that develop robust constraints on the existence of a reason. The hybrid theories developed by Schroeder (2009) are a useful example, though less explicit in their social orientation.

The core of the argument, which I hope informs a point of similarity with these expressivist theories, develops through looking at the range of social roles that reasons play. In some circumstances, we use reasons as a way of understanding the psychology of other agents, including predicting and influencing behavior.¹ In other circumstances, we use the reasons as a way of making sense of whether an act was justified or appropriate in a given circumstance. The apparently distinct nature of these psychological and normative explanatory projects collapses together and a theory that distinguishes the motivational and justificatory roles of reasons is better equipped to handle that collapse.

When I talk about the “social roles” of reasons, I am talking about a diverse set of ways in which we use reasons. These include giving reasons generally, discussed in chapter 4. The process of giving reasons includes both the explanatory and influential modes I will cover in sections 5.2.1 and 5.4.1, as well as the broader social considerations of providing public reasons for compliance with social norms, which will be relevant in section 6.3.4. The social roles of reasons also includes the use of reasons to understand and interact with others, which I will discuss in section 5.4.2. The “social roles” of reasons functions as a broad catch-all for the range of uses of reasons that we use in our interactions with others, and in navigating interactions with other peoples and with our communities more broadly.

Much of the course of this chapter looks expository. While the argument is laid out in the following section, the remainder of the chapter provides

¹There is an extensive tradition of reasons discussed in terms of explanation, which [Finlay and Schroeder \(2017\)](#) develop. That tradition develops out of [Davidson \(1963\)](#), though it has now largely been subsumed under the expressivist, psychologically-oriented theories.

an illustration of the ways in which the social roles of reasons are explained straightforwardly using a combination of my own theory of reasons and the basic tools of social ontology (which I expound in section 5.3). The expository sections are themselves contributing to the argument, albeit indirectly; the hope is to illustrate the range of social roles and show how readily my theory handles those social roles. This discussion also helps to set up responses to a range of potential counterexamples to constructivism, raised and addressed in chapter 6.

This chapter also serves to illustrate what has been my principal reason for pursuing this project. The purpose of a theory of reasons is not merely to understand when an agent is justified in acting a particular way, or what motivates agents and communities, or how the anthropology and history of our own reasons might have been different (and what parts might be invariant), and then to stop. Rather, the value of a theory of reasons consists in the enormous importance of reasons to our lived experience and social interactions. This is philosophy with a practical orientation. Developing the social orientation of reasons is important; it informs not just our understanding of practical reasoning, but also of a range of important domains in human life and experience. The expository elements of this chapter develop the connection between abstract characterization of reasons and the familiar, lived uses of reasons, with a broad, forward looking, and practical orientation.

5.2 Extensional improvement on existing theories of reasons

The preceding chapters have already developed part of this extensional case. For example, the pretext argument in section 2.5 holds that, on a personal and interpersonal basis, the attitude-independent, realist theories of reasons get the extension wrong. These attitude-independent, justification focused theories of reasons consider mere justification sufficient, and this means that pretexts (where pretexts are facts that justifying the act but do not motivate it) are reasons. This is a mistake.

In social contexts, we can put an even finer point on this objection to strictly justification oriented theories. If a given fact merely justifies and fails to motivate, then that fact will be wildly deficient with respect to the various social roles of reasons. It will not be recognized as adequate by individuals for whom it lacks motivational force.

The pretext arguments turn on the idea that if an individual is asked “what reason was there for you to ϕ ?” and that individual responds by presenting a pretext for ϕ -ing, then the individual is not cleaving to the convention.² Perhaps the individual is misusing the convention; perhaps the individual is being duplicitous. The question is not merely about the justificatory role. The reason provided by the individual has to also be reflective of the motivation set of the agent to whom the reason applies.

²We see the focus on conventions governing the invocation of reasons developed similarly in Alvarez (2017), where the characterization of reasons as justification and motivation comes about through the social practice of asking for and giving reasons. “Why are you always lying?” is such a request and part of that social practice.

Consider a schematic of potential cases, where T is trying to influence the behavior of S. T may influence S to ϕ in C by presenting a fact to S. Suppose that T presents the fact F to S; F is evidence that S ought to ϕ in C and therefore satisfies the justificatory role. However, F does not have a relation to any value held by S such that F would motivate S to ϕ in C. In these cases, there is a question whether T has presented a reason for S to ϕ in C at all.

Suppose that Mr. Jones is trying to motivate his child Sam to do homework, and notes that there is reason to do homework because if Sam does the homework, then Sam's grades will be better and there will be greater opportunities for getting into a university. Sam does not care about getting into a university, and therefore does not regard this as a motivating reason for doing homework. In offering this reason, Mr. Jones has offered what many of us recognize as a good justification, but what is not a reason *for Sam* to do the homework.³ Rather, it is merely a justification, with no motivational force, as in the mere pretexts cases in section 2.5. It may be a reason for others, whose aims and attitudes are copacetic with the reason; it just is not a reason for Sam.

The intuitions in this case are softer and more conflicted than in the mere pretext cases. However, there is some considerable literature on expressivism that advances this position.⁴ I cannot survey all of that literature here,

³There are a few potential ways of developing this. One is to say (as I have here) that this is just not a reason; another is to say that Mr. Jones as offered Sam a very bad reason for Sam to do homework. There is some slippage between these two points. Deterioration in the quality of a reason may also be related to deterioration in some facts being a reason at all. In some cases, it seems that "being a bad x" and "not being an x" are perfectly consistent, and reasons are a notable instance.

⁴A survey that includes both the original, weaker version and some tempering that I think brings expressivist approaches into compatibility with much of what I have written in this dissertation is presented in [Schroeder \(2013\)](#); [Joyce \(2009\)](#). I hope it is clear that

as it would be imprudent, but the idea is similar to what I have developed above.

5.2.1 Adding explanatory value beyond extension

An alternative, but closely related, line of argument concerns the explanatory power of certain theories of reasons that focus on social interactions including influencing other agents to act. In some accounts, the social practice of giving a reason for an agent to act is to express a particular attitude. Originally, this is developed by [Ayer \(1936\)](#) and [Stevenson \(1944\)](#) as an explanation for non-cognitivist moral discussion, where the expression of a moral property (for example, ϕ is impermissible) is correlated with an expression of disapproval (most famously articulated as “boo ϕ ”). The practice of reason giving is then situated in a desire to influence people to act consistently with our desires; as this approach develops, the understanding of the social practice develops. While I think this is an important part of a theory of reasons, centering the approach results in significant explanatory neglect.

When T presents F to S, it is not merely that T is presenting the content of F on its own, but is doing so with the intention of influencing S to ϕ in C. In that context, stating that F is not just exhausted by the truth conditions of F but includes the influence of F on S ϕ -ing (or not) in C, the non-cognitive theories get this element right. To treat it merely as justificatory without considering the motivational impact is to miss a core element

more sophisticated, contemporary versions of expressivism implement constraints that create relevantly similar extensions to my own analysis of being a reason.

of the phenomenon, but to treat it as motivating without recognizing the closely related role in justifying is also to miss the point.

In this respect, my approach attempts to capture something that historical expressivists like Ayer and Stevenson are after, and which develops through the contemporary (and more sophisticated) iterations of expressivism. Expressivists want to consider T presenting F to S as an act of influence and in terms of the expression of T's motivations and desires. I want to preserve this, but the result of my theory is that it imposes a stricter success condition on giving reasons than the expressivist analyses do.

Exerting influence and expressing attitudes are sufficient for expressivists like Ayer and Stevenson, but the refinements put in place by more recent hybrid theories help to also establish the justificatory role. The expressivist theories get the importance of influencing others and expressing preferences, but need a bit more flesh to get the justificatory elements which I view as necessary.

Contrary to the expressivist theory, I use the justificatory condition for constraining being a reason. In order for T to present a fact that is actually a reason for S to ϕ in C, it cannot merely be a fact that aligns with S's motivation set. That fact must also satisfy the justificatory condition.

This part of my argument is more intuitive. There are lots of ways of influencing behavior which do not actually involve justifying the behavior. One of the familiar elements of conventional social psychology (and related domains) is to address instances of influence without justification. Much of this literature implicitly appeals to our norms of justification in advancing a normative judgment about the practices of influencing without justifying (that it is wrong to use F to motivate S to ϕ if F does not also justify S ϕ -ing);

in some cases, such influence is pragmatic and appropriate, but in other cases it poses serious moral and political worries.

When my father was a little boy, his mother was trying to get him to put on a jacket. She offered all sorts of reasons to him that justified putting on a jacket but could not motivate him to do so. Finally, my aunt (my father's older sister) told him, "mom is cold, so you have to wear a jacket." This coincided with my father's motivation set, not wanting his mother to be cold and uncomfortable. Of course, grandma being cold does not justify dad putting on the jacket; it succeeds in motivating but fails to justify. Did my aunt offer my father a reason? On certain unconstrained expressivist accounts, or even some unconstrained constructivist accounts, she did.

On my constructivist account, she did not.

I have the justificatory role as a constraint. It is not the case that grandma being cold is evidence dad ought to put on the jacket. The fact that putting on the jacket is not justified by someone else being cold means that this is not an instance of offering a reason. However, it is an effective case of motivating behavior. Sometimes we motivate behavior by presenting things that look like, or are wrongly thought to be, reasons. This is an unsurprising, common enough case.

5.3 Basic considerations in social ontology and explanatory value

The previous section deals with extensional difference and why the more restrictive analysis I provide in this dissertation is better than some of the

competing theories of reasons. Building out the explanatory account of difference, explaining how my theory represent an improvement in regards to explanatory power, especially for the social role of reasons, requires providing some general philosophical background on social ontology. Some of this social ontology discussion may seem vestigial, but it also sees significant use in addressing counterexamples in chapter 6. My goal is to be as brief as possible here, but there is significant ground to cover.

The idea of the explanatory growth is that accounting for both justification and motivation allows for making sense of the plurality of social roles that reasons play, and how those roles interact when dealing with influencing, coordinating, and explaining the behavior of other community members. This explanation occurs within the realm of social facts, and is reinforced by institutions, social norms, and other things pivotal to social ontology. As such, providing some account of those social facts, institutions, norms, etc. is instrumentally indispensable at this stage.

On my analysis, reasons are a proper subset of facts. The set of reasons includes both “natural” facts and social facts. The metaphysics of social facts are relevantly different from that of “natural” facts.⁵

Naturally, “natural” facts are viewed as independent of the attitudes of agents. That light travels at 299,792,458 meters per second in a vacuum is true regardless of the attitudes of any agents.⁶ Not so for social facts. Social facts generally depend on attitudes of and within the societies in which those facts are situated.

⁵See [Epstein \(2018\)](#) for a survey.

⁶There are reasonable discussions about the socially constructed domains of study of “natural” facts, e.g. the claim that while the speed of light is attitude-independent, the units of measurement are socially constructed. See [Hacking et al. \(1999\)](#); [Pickering \(1999\)](#).

5. Added value and the social analysis of reasons

Social facts include those that obtain in virtue of regulative and constitutive conventions,⁷ as well as the mechanisms by which those conventions are in force.⁸ A *regulative convention* pertains to patterns of action in cases where an agent might do otherwise. A constitutive convention gives conditions for what counts as performing an act of a certain type; that moving a piece counts as a chess move or that saying a set of words in a given context counts as promising is governed by the constitutive conventions of playing chess and promising (respectively).

Regulative conventions are important to the social roles of reasons, especially in exercising influence, but they are also deeply familiar and important to understanding core moves in constructivism about reasons. On the constructive analysis, various norms regarding what facts justify behavior, and what sorts of norms are in force, play a pivotal role in making sense of some putative counterexamples to constructivism in section 6.3.4. For this reason, it is useful to focus on the ontology of constitutive conventions and say more about how they are grounded.

Following Searle (2006, 1995), facts about the office of the President of the United States are grounded by the community collectively believing certain things about the office. Joe Biden is the President of the United States in virtue of a set of constitutive conventions relevant to the office, how it is acquired, etc., that are laid out in the United States Constitution. These constitutive conventions only work so long as, collectively, people believe them.

⁷See Gilbert (1990).

⁸See Brandt (1967).

Alternatively, there are cases without such grounding beliefs, but where the behavior of members of the community grounds a social fact.⁹ The existence of a neighborhood, for example, may be grounded independently of the beliefs of the community; it may be governed by the behaviors of the individuals who tend to live, work, and shop within the boundaries of that neighborhood. The approach focusing on behavioral coordination helps to establish social facts in economics, or the existence of certain powers of community members that are not the object of any explicit constitutive convention or active, collective belief.

Following [Brandt \(1967\)](#), there is a useful distinction between norms (both constitutive and regulative) that are *in force* and those that are not in force. A society may have regulatory norms governing the democratic process, but those norms are ignored by those in power and the population generally such that the norms are not actually playing a causal role in determining who leads the country. In those circumstances, the norms are not in force. A norm being *in force* just means that the norm is adhered to within the society and that there are enforcement mechanisms (which will vary based on the norm and circumstance) that address violations of said norm.

There are a handful of conflicting theories of social ontology. I am grateful to [Vooy and Dick \(2019\)](#) for their discussion of this ongoing debate around the metaphysics of money. I will not go into great depth here, because I think the range of positions at issue are not exclusive and the correct

⁹These are not robustly attitude-dependent; that is, they still depend on the members of the community having certain attitudes. However, these facts are independent of the narrow set of attitudes about the fact. Agents believing that the dollar is worth more than the yen is not necessary for the dollar being worth more than the yen; however, agents have to have some beliefs about currency for the relative values of those currencies to hold.

5. Added value and the social analysis of reasons

approach is just pluralism about these theories, but a brief survey is in order.

(Table 5.1 provides a quick reference.)

Theory	Social facts depend on...	advocates
Collective intentionality	Collective attitudes	Searle (1995, 2006)
Intention and enforcement	Attitudes and force	Brandt (1967)
Social Functionalism	Playing a specified role	Guala and Hindriks (2015)
Hybrid Intentional View	Combination of the above	Vooy's and Dick (2019)
Pluralism	Any of the above	me

Table 5.1: Table of social ontological theories

Some theories in social ontology hold that social facts are grounded by collective intentional states. The community collectively believes that p , and this grounds the social fact F . The fact that the conventions for the electoral college are laid out in the Constitution is not, itself, sufficient to ground the electoral college and the creation of the office of the President of the United States. Rather, it is further that the community recognizes and believes in these conventions; not all members believe this, but enough do that the norms are still in force.

Some theories hold that social ontology is not grounded in collective intentional states, but instead is grounded by functions which usually (but not always) correlate with these mental states.¹⁰ The fact that there is a border between the United States and Canada is not a matter of the collective attitudes of members of those communities, but a matter of the separation created, the policing of the border, the handling of cases in different jurisdictions, etc. That there is a border obtains because of these functions. If people had attitudes to this effect, but the functions deteriorated, then there

¹⁰The most famous advocates of this view are Hindriks and Guala. See [Hindriks and Guala \(2019\)](#); [Guala and Hindriks \(2015\)](#).

might not be a border anymore or the border might move, as it sometimes does during active conflicts.

Some theories, including those advanced by [Vooyoys and Dick \(2019\)](#), propose hybrid analyses for cases, where both the functional and collective intentional dimensions play a role.

I think one should be a pluralist about these explanations. Part of the reason for my pluralism is that the appropriateness of one account or another may depend on the kind of analysis an individual is developing. Economic analysis, for example, often wants to consider functionalism, because very few (if any) members of a society have beliefs about the value of currencies and the value of a currency exchange or stock markets are better reflected by the functions of exchange. By contrast, notions of political and institutional authority (as developed by [Raz \(2009, 1990, 1985\)](#), *inter alia*) are grounded in practical reasons and considerations of collective intentions.

I do have thoughts and positions on this that follow out of constructivism and the theory of reasons, but this would not be prudent in this dissertation. Instead, let us focus on some of the elements of social ontology that are important for this dissertation.

First, collective intentionality does not require unanimity. Collective intentional states are states held by a group that are directed at the same object; these may include both doxastic and evaluative attitudes.¹¹ In some cases, the collective intentional state held by a group is governed by constitutive norms of decision making within the community; “the United States believes that Saddam Hussein has weapons of mass destruction” is established not by the unanimous beliefs of members of the United States, but by

¹¹See [Gilbert \(2002\)](#).

constitutive conventions that establish the intelligence apparatus and designated George W. Bush as “the decider.” This helps to distinguish between the norms of that in force within a community, irrespective of whether they are held unanimously.

This observation is important in developing the understanding of social influence in sections 5.2.1 and 5.4.2, and will play an especially prominent role in developing the responses to prospective counterexamples in chapter 6.

Second, some social facts may obtain in some minimal sense but not be in force; conventionally, these are called “dead letters.” States may have laws on the books, such that those laws are still part of the social facts, but they are not in force (and may actually not even be permissible subjects of enforcement, as a result of other social facts) because they do not have any influence whatsoever in a society.

Third, collective intentional states may vary from individual intentional states in some important respects, but they still have some familiar features. Gilbert (Gilbert, 2002, p. 133-136) notes some respects in which there may be collective feelings, which gives us some framework for talking about Scheffler’s element of valuing that deals with dispositions to feel, see item 2. Communities can be disposed to feel, collectively, in ways that set up similar elements in talking about collective intentional instances of valuing. This is considered less immediately in this dissertation and is instead more a feature of the future-looking discussion in chapter 7.

This dissertation cannot plausibly develop a robust account of collective intentional states, but piggybacking on the standard approaches developed by Searle, Gilbert, *inter alia* makes this account straightforward enough.

The complex intentional states, like valuing, that give rise to the existence of reasons can obtain for groups of agents, and reasons develop for those groups of agents in the same way.

5.4 The robust explanatory value in accounting for the social roles

One major argument for advancing the theory of reasons developed in this dissertation is that it provides a range of explanations for social phenomena that make use of reasons. By providing the more thorough account of reasons in terms of their functional roles, the analysis also allows for using the distinctions and interrelations in those roles to develop explanations of a range of social phenomena, including the use of reasons as means of influence, the social assessment of the reasons for action, and the formulation of new reasons by societies in order to influence the behavior of members of the community.

The argument provided in section 5.2 holds that the analysis of reasons provided in this dissertation is better because it maintains constraints that prevents any fact that influences behavior from being a reason and it prevents non-motivating justifications from being a reason, and both of these constraints help to maintain the appropriate class of reasons that occupy the social roles of reasons. As a straightforward matter, this is just about getting the extension of the analysis right; however, as I indicated at the beginning of this chapter, there is a second argument in favor of the analysis.

This second argument holds that we should prefer my analysis of reasons on the basis that my analysis provides a better explanatory framework for the social roles of reasons. There are two ways of setting up this argument. The first and more straightforward way to set this argument up is to note that by including both the justificatory and motivational roles in our explanation of how some reason works, especially in social contexts, provides for explaining two dimensions of reasons which sometimes do not overlap. In some cases, the explanation of how a fact justifies S ϕ -ing in C will differ significantly from how that same fact motivates S to ϕ in C, and it can be useful to differentiate and account for both.

Consider a case where S is considering whether S ought to keep a promise to T. S recognizes that there are a set of conventions around keeping promises in virtue of which S is justified in keeping the promise to T; these include both the constitutive conventions of what it is to make a promise and the regulatory conventions on keeping and breaking promises.¹² However, it might be the case that S is motivated to keep the promise not because of the general belief in the institutions, but because the relationship between S and T requires trust and S values the trust that T puts in S, even without consideration of the social conventions. In this case, there are a set of facts about trust and its importance in society, generally and in the particular relationship between S and T, and the explanation of the reason for S keeping the promise to T can be discussed in rich detail by recognizing both the

¹²The standard line of development here is from Austin's work on speech acts in [Austin \(1975\)](#) like "I promise to ϕ " and the development of social conventions. This also includes the regulatory conventions that it is wrong, except in extreme circumstances, to break a promise. While some readily distinguish between the constitutive and regulative norms, the promising case illustrates that this is not so straightforward.

justificatory and motivational roles and accounting for both as explanation of S's behavior.

One need not focus on extensional differences to get this explanatory improvement. The actual range of the explanation is more significant than just tying together the justificatory and motivational roles of reasons. If that were the limit of the explanatory improvement, then I would just have offered the extensional cases as part of the preceding chapter (on combining the two elements) and then proceed with that line.

More significantly, there are three broad, non-exclusive categories of social roles that reasons play and which are more comprehensively explained by my analysis of reasons. These three sets of social roles are (first) the explanatory and predictive role of behavior, (second) the influence and coordination of behavior among members of the community, and (finally) the constitutive norms in virtue of which certain act types are created by a combination of the collective reasons and the basic scaffolding of social ontology presented in the previous section.

5.4.1 Explaining and predicting behavior, the basic social roles

It is an important feature of human social life, and even personal psychological life, that we can predict and explain the behaviors of other agents. The fact that we can predict and explain the behaviors of others pivots on reasons, understanding both the reasons that there are for individuals to act, as well as the reasons an individual is likely to adopt given what we know about their psychology. An upshot of my theory of reasons is that it

provides further substantial development for how we do the predicting and explaining.

Such predictions are of course fallible, but familiar and integral to our navigation of social spaces. Our relationships with other agents that we interact with are often predicated in part on our ability to predict certain sorts of behaviors. If a professor fails to turn in letters of recommendation for graduate students, then we understand this will be a reason for frustration among the graduate students and that professor's colleagues. We understand the frustration, because we know the fact, common attitudes of the graduate students and colleagues, the norms in virtue of which that frustration is justified, and so on. We can develop the intuitive story of how we predict and explain the frustrated response through the theory of reasons.

There is an extensive literature on social cognition and its role in social life;¹³ I will not rehash all of it here, though it is important that my own theory of reasons is especially well positioned to inform that literature. My theory operates as a tie between the philosophical theories of justification and motivation, the cognitive psychology of social interaction, and social psychology and sociology. The relation to the phenomena of explaining and predicting behavior is the first substantial explanatory benefit of my theory of reasons.

The most straightforward way in which the explanatory and predictive roles of reasons are satisfied comes through a direct understanding of the

¹³This is an enormous, sweeping gesture at the literature on social cognition that develops out of Fiske (1993) and into various projects on the roles of attention and perception in social psychology. See Tomasello and Carpenter (2007); Mundy and Jarrold (2010); Tomasello et al. (1995) for illustrative cases, and/or Fiske and Taylor (2013) for an advancement of the tendrils and implications of this psychological literature.

5. Added value and the social analysis of reasons

internalist approach to reasons. If I understand that an individual has certain contents in their motivation set, then I can explain and predict the behaviors of that individual. This need not apply to reasons; it can apply to non-reason instances of motivation just as well.

Suppose I know that Aaron dislikes his job and continues to do the job in order to support his family. I can reasonably predict that, if given an opportunity to leave his job in favor of something that more closely aligns with his values, either by making him happier or allowing him to better support his family or some-such thing, Aaron would leave his current job. This prediction might lead those who have the relevant information to behave differently; perhaps it would lead his manager to offer better pay or benefits to incentivize Aaron to stay, or perhaps it leads some other potential employer to try and poach Aaron. This sort of recognition is integral to our social life.

This is true whether it is forward-looking (predicting Aaron's potential departure from his job) or backward-looking (understanding why he left). This should be familiar to the internalist about motivations and is no doubt one of the strong motivations for adopting a theory of reasons that takes the motivation set as relevant.

The challenge with the motivation set is that, by design, it requires the postulation of internal states. In many cases of reasons, it may be the case that postulation of internal states is not especially relevant. Interestingly, one of the major recent developments in cognitive psychology (especially

as applied to social psychology) is the use of attentional exercises to explain various forms of belief and even reasoning ascription.¹⁴

Suppose that Terrence sees Ronda looking at an especially putrid dumpster as she walks down the street. Terrence might reasonably suppose that Ronda will cross the street to limit her exposure to the smell of the dumpster. He can predict her behavior (or, if this is after the fact, explain her behavior) based on tracking her attention, her eye-movement. Historically, some psychological and philosophical theorists have supposed that this involves some level of attitude ascription to Ronda, as one might see in the case of Aaron refusing the job; on this evaluation, we could talk about Terrence understanding Ronda's behavior as an instance of "mind reading." But more recent literature in cognitive psychology challenges the necessity of this attitude ascription in the process of predicting behaviors.

Lots of animals, including human infants, engage in the sort of assessment of behavior based on looking at patterns of visual attention among conspecifics, even while flouting apparent attitude-ascription. Whether predicting behavior based on apparent patterns of visual attention and predicting behavior based on ascribing attitudes are coincident is an empirical question, and the empirical answer seems to be that they are not coincident, even if they typically co-occur for neurotypical adult humans. The fact that

¹⁴There is an enormous amount to say about these phenomena, which I believe have been largely neglected by philosophers outside of a very narrow body of work in the philosophy of mind and cognitive science. Some philosophers, e.g. Prinz (2012, 2011), have developed the potential role of attention in consciousness and visual perception; this still neglects the significant role of attention in social interaction, including the assessment of reasons, ascription of internal states to other agents, etc. For a survey of this phenomenon, see Moore et al. (2014); Mundy (2016). This literature is a part of why my concern with reasons as facts, as discussed in section 1.2, is central; accounting for common points of reference along with differences in the minds of agents helps to both understand the common and divergent elements of cognition, including social and moral cognition.

the two are separable provides for us to recognize a split in the two processes that we recognize in these social psychological phenomena.

5.4.2 Influencing and coordinating behavior, to build on predicting and explaining

The preceding sections provide some discussion of the role of influencing behavior, on the basis that predicting and explaining are often used as a way to influence. When one knows that, if S knows that F, then S will be inclined to ϕ , it is possible to influence S to ϕ (or not) through this awareness. However, much of this is simply consistent with the standard notion of influence discussed in the simplified expressivist views and section 5.2.1. The more challenging project is in developing this notion of influence alongside the justificatory role of reasons, as this helps to understand whether the influencing includes the offering of a reason, based on whether the justificatory role is satisfied.

The latter, the provision of a justificatory framework for assessment of behavior, is familiar to philosophers. Still, perhaps a superficial gloss is useful here. In chapter 2, I note that the reasons as evidence thesis helps to clarify the relation between the fact and the ordered triple of agent, act type, and circumstance. The fact must stand in a justifying relation to the ordered triple in order to function as a reason. R is a reason only if R is evidence that S ought to ϕ in C. Constructivism holds that this relation is itself not attitude-independent, though it has attitude-independent elements. Still, these relations are familiar to us in the context of ordinary acts of convincing others.

5. Added value and the social analysis of reasons

When trying to convince someone that they should do dishes if they find dishes in the sink (even if the finder not responsible for those dishes), one might appeal to the sink as a community space, to the likelihood that dishes will stink up the rest of the community space if not dealt with quickly, and so on. We, as members of a shared community, recognize these considerations as appropriate justifications (though only *pro tanto* justifications) for the behavior, and try to influence others through appeal to that shared recognition. We may not succeed in influencing others, but the idea is that there are these further constraints on what counts as a reason for action.

Philosophers across various subdisciplines are concerned with determining, assessing, and reforming these various constraints on what is appropriate, providing reasons for changing (or not) some of our existing norms governing reasons. This approach to influencing is a significant part of first-order work in moral and political philosophy, and perhaps even other philosophical domains where substantive proposals are supposed to motivate individuals to change their actions and beliefs. But (to be crude) let us avoid shoving this part of constructivism up itself.

The general idea here is that the justificatory roles of reasons explains these constraints on what counts as a reason, recognized by a community, and the motivational considerations help to explain the force of persuasion on individuals' courses of action. By combining the motivational and justificatory functional roles, we get a better understanding of how each of these functions (pardon me) function. This also allows for making sense of interplay and conflict between the two roles.

For example, consider a set of rhetorical moves whereby an individual appeals to facts to influence a group to behave in a particular way. Suppose

5. Added value and the social analysis of reasons

that these rhetorical moves are incredibly effective, that they work at mobilize people to vote for the individual to lead the country. However, as a matter of the actual norms of justification, by way of some constraints established, make it the case that the appeal to these facts does not justify the behavior of installing this person as leader. Given recent American history, this should not be so hard to imagine. Suppose that the motivational effect is strong, but the justification absent. My theory of reasons provides us the ability to talk about why the facts set out by this individual are not reasons; that is, the theory gets the extension right and further explains why the particular facts in that case do not fall under the extension of being a reason. This is significant added explanatory value. Further, this distinction allows for appropriate study of the relationship between effective motivation and appropriate justification.

This is not just about the use of reasons to influence the behavior of other agents. Facts influence our own behavior, and the prediction and explanation of behavior discussed in the previous section can make it the case that we behave differently to coordinate our behavior with the behaviors of other agents.

Suppose that I know there is a checkpoint that local police set up following games of American football at the university, to deal with drunk fans and frat boys. My predictions of the behavior of the police department may influence me to take a different route home, to avoid this issue. If I know that there is another event is likely to result in similar debauchery from the frats on campus and that the police department (knowing this) will set up that checkpoint, then I may preemptively respond to this reason and the

5. Added value and the social analysis of reasons

expectation of the behavior of the police department by adjusting my route home, even without direct knowledge that the police checkpoint is set up.

In this case, the fact that the police department has reason to act influences me to alter my behavior. Knowing that there is a reason for the department to set up a checkpoint leads me to anticipate the likelihood and adjust accordingly. This is usefully explained by my knowledge of the facts and by my knowledge of how others respond to those facts.

Coordination is important because it applies to both a range of explicit and implicit norms governing communities of individuals. There are explicit rules of the road, based on expectations about driver behavior and what sorts of restrictions the community want to impose to balance public safety and effective transportation. There are implicit rules of the road that help individuals make predictions about the behavior of other drivers and coordinate their behavior accordingly. Maintaining several car-lengths of distance on the highway is important, and is informed by specific explicit norms, but the fluctuation of that distance based on lane (maintain more distance on the right lane, to anticipate merging) and on observation of other drivers' behavior (extra distance for those who display erratic behavior).

Explaining how these practices of coordination develop is important to the social sciences, as well as to improving the effectiveness of communities' attempts to improve their own rules governing such coordination. Understanding what incentives and disincentives help individuals coordinate behavior is important, and this requires understanding the appropriate justificatory frameworks within the community as a whole and the motivations of individuals and the community.

5. Added value and the social analysis of reasons

I am writing this during the COVID-19 pandemic, and so the question of how actual behavioral coordination works is a serious and consequential matter. Maintaining compliance with public health guidelines turns out to be complicated; there is clear evidence that individuals ought to wear masks and large groups of individuals for whom this evidence seems to motivate them to do precisely the opposite. Maintaining effective public health policy requires considering both the conditions of justifying compliance and the way to motivate actual compliance, or to create norms that coordinate the behavior of members of the community in ways that mitigate risk. Unfortunately, facts that justify coordinating behavior and perhaps even align with some things the agents in question value do not provide adequate motivation. There are questions about whether such agents are acting rationally or reasonably, even they are even reason responsive with respect to these issues. These practical problems are easier to tackle when we acknowledge the psychological and normative dimensions.

Considering both the justificatory and motivational roles of reasons help to illustrate the relative complexity of these cases and provide grounds for developing explicit rules and other mechanisms of influence that appropriately coordinate behavior and promote the desired public health outcomes. Being able to consider both under the theory of reasons is a striking advantage and necessary for getting into the application of reasons to domains of social coordination.

5.4.3 Explaining and leveraging constitutive conventions

Constitutive conventions are the set of social facts that make it the case that S ϕ -ing also counts as S ψ -ing. The classic example is chess moves. The rules of chess are constitutive conventions that make moving the pawn from D2 to D4 count as a chess move and moving that same pawn from D2 to D5 instead not a chess move.¹⁵

Constitutive conventions play a further role in understanding reasons (on any theory, not just mine) by creating new act types. There are constitutive conventions governing chess moves, making promises, holding public offices, and so on. These conventions govern when the performance of a particular act constitutes the performance of a given social act, when saying “I swear to tell the whole truth” constitutes taking an oath (in a courtroom, for example, and not on the set of a television show). The background of these social acts are relevant to reasons in that they create act types for which they may be reasons and the conventions themselves create further reasons for action.¹⁶

Consider the case of taking an oath to tell the truth in a court of law. It is not merely that one says the words and therefore the testimony given is under that oath, but further that the community that recognizes that oath

¹⁵There is sometimes dispute over whether an illegal move is either (a) a chess move that is illegal or (b) not a chess move at all. There are some interesting puzzles around this, for example when Inarkiev plays an illegal move against Carlson in the world chess championships, and it is Carlson who forfeits the match on a technicality (by continuing to play after the illegal move). This is where the relation between regulative and constitutive conventions can get blurry.

¹⁶In Daniel Handler’s (pseudonymously authored) books *A Series of Unfortunate Events*, the wicked Count Olaf attempts to trick the community into recognizing a marriage between him and his young ward Violet, by putting on a play in which the two are married. This hinges on the mistake that not any instance of stating marriage vows is sufficient to constitute marriage. Context matters; and the performance of those vows in a staged play is not an appropriate context. Otherwise lots of actors would be polygamous.

holds stricter rules on lying when one is under oath. Perjury is different from merely lying; there are special regulatory norms that the community puts in place to discourage lying under oath, regulatory norms that specially apply to acts created by the constitutive norms governing testimony under oath.

The existence of the act types is one dimension and layered on top of those special act types are further reasons that apply to those special act types. We create these special pockets of our social world with especially important kinds of influence to ensure that there are times when certain values necessary to the community are subject to more rigorous enforcement and agents recognize stronger reasons. When testifying, one is in a pocket of our social world where honesty is especially valuable and the consequences for dishonesty are commensurately much more severe.

Constitutive conventions, like regulatory conventions, are a way for communities to influence and coordinate behavior. It builds up these layers of coordination with the more robust explanation of how the social world is brought about, not just through bringing about social facts, but making those facts reasons for members of the community and the community as a whole.

5.5 Explanatory value, contra alternative theories

My theory of reasons is not the only theory of reasons that account for the social roles played by reasons. Rather, the goal is to illustrate how a

theory of reasons may provide such an account; theories with the available tools can do what I have done here. The more subtle point is that the way in which theories of reasons provide the explanation of the social role is by taking the broader, methodological approach to reasons developed in this dissertation: they combine elements of motivational internalism with constraints on justification.

Consider the sorts of tempered varieties of expressivism developed in [Schroeder \(2013\)](#). Some elements of this expressivism propose constraints on what sorts of things can be said, and those constraints are not so relevant to my theory here ([Schroeder, 2013](#), p. 283-286), but there are other elements that Schroeder proposes for tempering expressivism that provide the justification conditions. Schroeder develops an expressivist account of the meanings of certain terms by holding that some terms (though not necessarily all terms) have expressive and non-cognitive components as well as descriptive components.¹⁷ These hybrid expressivist theories temper the meaning of some of the salient moral terms with constraints on appropriate use.

Notably, I have not provided an account of the meaning of moral terms in this dissertation, nor do I plan to do so. The purpose of invoking the hybrid expressivist theories here is to note that the theories embrace both the motivational internalism and some external constraints in order to provide the motivational role (which has long been the focus of internalism) without losing the justificatory role.

On this argument, hybrid expressivism fares better than the alternatives, e.g. Parfit-style realism that admits of non-motivating reasons and better

¹⁷See ([Schroeder, 2013](#), p. 287) and [Schroeder \(2009\)](#).

5. Added value and the social analysis of reasons

than expressivist theories that lack appropriate constraints. The improvement of my theory over the more proximate competitors of such hybrid expressivism is in making these distinct roles and functions explicit in order to ensure that explanations of the functional roles of reasons can play clearer roles in each case, rather than foregrounding motivation and backgrounding justification as those expressivist theories do.

Chapter 6

Counterexamples to the proposed analysis of reasons

6.1 Counterexamples to my proposed theory of reasons

I collect church collapses,
recreationally... If [God's] up
there, he just loves it, Officer
Starling. Typhoid and swans - it
all comes from the same place.

Hannibal Lecter in Thomas
Harris' novel *Silence of the Lambs*

The theory I set out in this dissertation is new and so there are no objections that directly focus on the theory. However, there are objections to standard constructivist theories of reasons; some of these objections might develop a response to my theory, given the theory's commitment to constructivism about reasons and values. There are also standing objections to elements of my theory borrows from other theories in the literature, like the reasons as evidence thesis, though I have tried to handle these as they come up.¹ While my theory does not outright solve all of these objections, for reasons I make clear in developing the objections and responses, the theory does provide tools for a more thorough handling of these objections. One implication of this chapter is that the combination of the reasons as evidence and promoting and protecting conditions leaves the theory better equipped to handle many standard objections to constructivism.

¹Section 2.4 considers some objections to the use of the reasons as evidence thesis as a necessary condition, as well as direct objections to the thesis itself. Section 3.3.4 considers objections to constraints on valuing. Section 4.4 considers a possible objection to my own theory of reasons based on the redundancy of the reasons as evidence thesis. As such, I do not consider these objections in this chapter.

The core of the objections to constructivism can be broken out into two parts. First, there are some attitude-independent reasons and constructivism cannot account for those reasons. As such, constructivism fails to provide an adequate explanation. Second, constructivism results in some facts being reasons when they should not be, and therefore constructivism fails by over-generating. Section 6.2 of this chapter deals with the strongest arguments for the former. Sections 6.3 and 6.4 deal with strong arguments for the latter.

I will not presume that my responses to these objections are decisive or exhaustive. The objections are well entrenched in an extensive literature, and there are a range of versions of these objections.² Some versions of the objections are subtly different in ways that require a different response. I have done my best to present what I take to be the most challenging versions of the objections here.³ This chapter consists of what I view as the best counterexamples.

6.2 Entailment of normative facts

One putative set of counterexamples to constructivism is the existence of descriptive facts which entail some normative judgment. On constructivism, the existence of a reason for an agent to perform an act of a given type is contingent; it is contingent on a range of facts, including the attitudes of agents (including the agent acting).

²I provide lengthier discussion of these variations in section 7.4.

³Some further versions of objections are considered in section 7.4, with short explanations of why I regard those particular versions as deficient.

6. Counterexamples to the proposed analysis of reasons

On this class of counterexample, there is a description of some fact F such that F entails a normative judgment about the actions of the agent. The argument then proceeds that such a fact *just is* a reason, because it entails (at a minimum) the justificatory condition for action in such a strong way that it compels the agent to act, at least on a *pro tanto* basis.⁴

Perhaps the most widely used counterexample to this effect is Parfit's "agony argument." (Parfit, 2011, p. 82) On the agony argument, the fact that $S \phi$ -ing in C would prevent S from being in agony is a *pro tanto* reason for S to ϕ in C , irrespective of S 's attitudes.⁵ The fact that S failing to ϕ in C would result in agony is sufficient to entail a normative judgment that S ought to ϕ in C .⁶ Sobel writes,⁷

Parfit asks "Who could possibly deny that the nature of agony gives us reasons to want to be happy?" He claims that the Agony argument shows that "such claims *must* be denied by those who accept subjective theories about reasons." The Agony Argument

⁴One prospective class of these conditions would be thick concepts (Väyrynen, 2019, Sec. 2). Thick concepts have a descriptive dimension and conceptually entail a normative dimension. However, given disputes about the existence and status of thick concepts (Väyrynen, 2019, Secs. 3-4), they are not an especially useful example. They represent a proper subset of potential candidates, but there are alternative cases where a fact or property seems to entail a normative judgment that are stronger. I focus on those stronger cases instead.

⁵Parfit does not actually use the language of "*pro tanto*" reasons in his writing, which can create some ambiguity in various instances. There is an open question as to whether Parfit means *pro tanto* or decisive reason in talking about agony. My view is that the most charitable interpretation is that he means *pro tanto* reasons, as this would be a much less demanding view. The reading of Parfit on which these reasons are decisive entails that there are such *pro tanto* reasons, but also requires a further claim about weight. As such, I just focus on the version concerned with *pro tanto* reasons.

⁶For direct responses to Parfit's version and Enoch's variation, see Sobel (2011, 2009). Sobel's version of subjectivism provides good reason for suspicion about the entailment of this normative judgment.

⁷This is (Sobel, 2011, p. 53); quotations are from (Parfit, 2011, p. 66), with emphasis in Parfit.

6. Counterexamples to the proposed analysis of reasons

maintains that we have reasons to want to get future feelings that we will like and to avoid future agony even when we have no current desires which would be served by doing so.

On Parfit's view, it is a feature of "the nature of" happiness and agony in virtue of which we have reasons to act in certain ways. Parfit spends a great deal of time specifying the narrow sense of happiness and agony such that happiness and agony entail the existence of a *pro tanto* reason to act in ways that influence our happiness and agony. These conceptual distinctions are discussed by Parfit (Parfit, 2011, p. 62-63) and David Sobel (Sobel, 2011, p. 55-61), but such distinctions are not necessary for my purposes here.

What is relevant to this immediate issue is the distinction advanced between what Parfit calls "object-given" and "subject-given" reasons.⁸

Object-Given reasons: "... are certain facts that give us reasons both to have certain desires and aims, and to do whatever might achieve these aims. These reasons are given by facts about the objects of these desires or aims..." (Parfit, 2011, p. 57)

Subject-Given Reasons: "... [are] provided by, or depend upon, certain facts about what would fulfill or achieve our present desires or aims." (Parfit, 2011, p. 57)

Parfit holds that subjectivist theories (which include constructivist theories) simply cannot account for object-given reasons of the kind he describes in the agony argument. This just seems unlikely for a few very general reasons.

The first is that agony itself, both intuitively and as characterized by Parfit, is a subjective state. Whether something counts as agony depends on

⁸See (Parfit, 2011, p. 57) and (Sobel, 2011, p. 63).

6. Counterexamples to the proposed analysis of reasons

the responses of the agent anyway. As a result, some “subjectivist” theories (inclusively understood to be any theory that takes reasons to be dependent on some attitudes of agents to whom those reasons apply) have argued that one should subordinate these judgments about reasons to avoid agony to the attitudes of the agents, basically biting the bullet.

In some cases, those philosophers bite the bullet by appealing to idiosyncrasies in the subjective experience of agony or the meta-hedonic states pertaining to the evaluation of agony.⁹ In some cases, this is done through the denial of the objective and subjective distinction as provided by Parfit.¹⁰ In some cases, biting the bullet is achieved through varying the ends or aims of the agent.¹¹ This provides a range of varied responses that can be combined or tailored based on the particular “subjectivist” theory one has in mind. At a general theoretical level, my own theory is compatible with versions of each of these.¹²

The second and more serious issue with the agony argument is that it requires far too strong a commitment. One could hold that there are elements which tie descriptions of our subjective experiences (like happiness or agony) to normative assessments, while also holding that those elements are parasitic on subjectivity and (therefore) not independent of our attitudes. As such, I want to consider an alternative examples which I think provides a more compelling counterexample to my own theory.

⁹See [Street \(2009\)](#).

¹⁰See ([Sobel, 2011](#), p. 64).

¹¹See ([Korsgaard and Korsgaard, 1996](#), “Skepticism about Practical Reasons”) and ([Schmidtz, 1994](#), p. 226-251).

¹²I provide a survey of potential responses in section 7.4.

6.2.1 The Empty Track versions of the Trolley Problem

Let us return to the empty track version of the trolley problem, previously discussed in section 4.4.1. As a refresher, the empty track version posits a standard switch trolley problem (the default track has five workers; the auxiliary track has one bystander; the agent is at a switch), but proposes that there is a second auxiliary track that is empty and on which the train can be brought safely to a halt.¹³

In the standard switch trolley problem, one might suggest there is reasonable disagreement about whether the description of the state is evidence one ought to pull the switch, thereby (at least on the reasons as evidence condition) constituting a reason to pull the switch. In the empty track problem, the intuition that there is a *pro tanto* reason to direct the train down the empty track (in fact, the intuition is that there is a *decisive* reason to do so) is sufficiently strong that it just seems entailed by the description.

The purpose of this counterexample is to suggest that there is a straightforward description of the case that just seems to entail a normative judgment about the case. There are some substantive background claims built into the judgment, which people in our moral community broadly share, but which may themselves be contingent. I discuss these claims below and argue that they are not properly “object-given” in a way that poses a problem for constructivists.

The fact that there is an empty track under those conditions *just is* a reason for the agent to pull the switch to direct the train down the empty

¹³This version of the thought experiment is my own development. I am grateful to S. Matthew Liao, Sharon Street, and the late Derek Parfit for their input in its development and help in exploring its repercussions. While I use the case for somewhat different purposes in section 4.4.1, this was the issue the case was developed to explore.

track. Further, what the agent values or does not value is irrelevant; whatever attitudes the agent has are irrelevant. The reason obtains no matter how we vary the psychological contents of the agent, and if the agent does not see the reason when aware of the empty track, then that is a defect in the agent.

The goal of the case is to provoke the intuition that anyone who would deny the existence of a reason like “flipping the switch to the empty track would save lives” is just making a mistake. Whatever attitudes they hold, whatever their motivations, are irrelevant. They are just making a mistake.

6.2.2 Contingency and the constructivist pivot

A promising response proceeds as follows. The fact that flipping the switch would save lives is a reason for the agent to flip the switch under the empty track case, but not as a matter of strict entailment. Rather, it is a reason because of an underlying value that we share, and expect (or even demand) other agents share. Because we share the value, we are disposed to treat the claim that lives have value as an attitude-independent fact.

It is not the case that the fact F entails that S ought to ϕ in C , on either a *pro tanto* or *pro toto* basis, in a way that is independent of the attitudes of agents; rather, it just looks that way from our vantage point because it is broadly shared.¹⁴ So a constructivist theory rejects the claim that lives being of value is an attitude-independent fact, and thus holds that the reason generated is not attitude-independent. On a constructivist account, we take it that F entails S ought to ϕ in C as a matter of the attitudes that the

¹⁴We take for granted that the lives at stake in the trolley problem are things of value and satisfy the conditions laid out in section 3.2. Even though this valuing is not unanimous, it is ubiquitous enough that we often just take that value for granted, the way that I do with winning and not losing chess in the toy examples throughout chapter 3.

6. Counterexamples to the proposed analysis of reasons

community broadly shares. We have an expectation that there are some points of shared value across members of the community, and the empty track variation of the trolley problem gets at one of those broadly shared values. However, that the value is shared is neither a matter of necessity nor attitude independent. Rather, the fact that the value is shared is contingent on a set of shared attitudes and conventions.

Constructivist theories can hold that there are alternative worlds where human lives may not have a positive value, based on the evaluative attitudes of agents in those worlds.¹⁵ Suppose that in those worlds, as a matter of course, every agent in the world shares indifference for human life. Under those circumstances, the constructivist then argues, it would be the case that the preservation of human life is not a *pro tanto* reason to flip the switch in those worlds.

This response addresses the first set of counterexamples to constructivism, namely that there are cases where some reasons seem to be attitude-independent, but that constructivism can account for them by maintaining that they are not *actually* attitude-independent, but just seem that way as a result of the contingent circumstances rendering them psychologically and socially ubiquitous.¹⁶

Take this response in tandem with the existing arguments advanced by constructivism against the possibility of attitude-independent reasons. I

¹⁵Benatar's famous arguments for anti-natalism indicate some possible evaluative attitudes might impact the valuation of life in the actual world (Benatar (2015, 2012)). This would of course depend on Benatar's arguments satisfying the constraints of rationality outlined in section 3.3.2. Even if Benatar's position is defensible (and it may not be), it may not entail differences in reasons in the empty track case, though one could tease out those entailments.

¹⁶Such values may not be unanimous, but they are common and generally actively endorsed and enforced by the society.

6. Counterexamples to the proposed analysis of reasons

presented these arguments in section 3.4. If one finds those arguments compelling, then this constructivist story of how we get reasons that *seem* attitude-independent is a useful addendum on the phenomenology of some reasons.

If this response holds, then the under-generation worries of constructivist theories of reasons are addressed. It is not the case that a constructivist theory will fail to account for the reasons in the empty trolley case; rather, it just posits a different explanation for how the apparently attitude-independent reason works and its appearance. In my experience, this is the lesser of the two sets of objections to constructivism. Much of the criticism of constructivism recognizes that constructivist theories are capable of building out these explanations for reasons and therefore avoiding the under-generation worries like those expressed above.

There are constraints on reasons in the actual world, in our own social circumstances, which have been and are constructed by existing regulative norms within our communities. These constraints are contingent, and depend on facts about our world. If our world were different, then the constraints would be different. Constructivism is well equipped at recognizing, explaining, and even building such constraints. The issue then becomes the non-necessity of those constraints, but this is a less distressing issue.

However, it is precisely the prospect that constructivism can generate these reasons and that the reasons are attitude-dependent that creates the over-generation worries I address the following sections. Those worries are more serious, and pose substantial worries that constructivist theories of reasons do not adequately constrain the explosion of reasons an agent might have for all sorts of weird and even repugnant acts.

6.3 Ideally coherent eccentrics

There are a range of cases throughout the philosophical literature that may broadly be characterized as instances of “ideally coherent eccentrics.” The core conceit of ideally coherent eccentric (ICE) cases is that there is an agent who has attitudes which produce a set of practical reasons for action that are “eccentric” to the point that one might be inclined to deny such reasons.

Among these cases are Rawls’ grass-counter (Rawls, 1971, p. 432), Gibbard’s ideally coherent anorexic (Gibbard, 1990, p. 171), and Parfit’s Future Tuesday Indifference (Parfit, 1984, p. 124).¹⁷ These various cases represent a handful of different potential objections to constructivism. In the case of Parfit, the objection is that the existence of the evaluative attitude and the ideal coherence are not sufficient to make it the case that there is a reason for the eccentric individual to act. In some other cases, like those addressed by Rawls, the objection is that the attitudes are of the wrong sort to form a reason. In other cases still, like those considered by Gibbard, the appeal is to a brute judgment that surely there cannot be a reason and therefore something must be wrong with the attitude-dependence claim. Notably, Rawls and Gibbard are considering potential objections to their own positions.

I have already addressed the first two objections indirectly in the course of unpacking the constructivism about values.

The objections raised by Parfit run headlong into the argument against attitude-independent reasons and problems of access addressed in section 3.4.2, since the bridging principle one would need in order to get Parfit’s objection to hold is exactly what Mackie *inter alia* attack.

¹⁷Street (2009) provides a survey of a range of these cases, including all of these examples and the ideally coherent Caligula case I use throughout the chapter.

6. Counterexamples to the proposed analysis of reasons

The objections considered by Rawls (and other so-called “Kantian” constructivists) are difficult for my theory to address directly, because I already impose fairly rigid constraints on what it means for an evaluative attitude to be of the right sort to ground the existence of a reason. Because the account of valuing I lay out in section 3.2 is very restrictive, it largely already covers worries laid out by Rawls. My use of Scheffler on valuing already covers certain dispositions to emotional responses that serve to sidestep some of Rawls’ worries, while the conditions of justification address some other worries. It would be imprudent to do the entirety of this exercise here, but perhaps this is a good subject for a future paper, should my theory of reasons find its way into the broader constructivist discourse.

The challenge with many of these cases is that they do not adequately provide stakes for the counterexample to constructivism. If a man wants to count grass and it genuinely satisfies his coherently considered aims, then what are the consequences of allowing him to proceed with his eccentricity? One might bite the bullet on the cases of Rawls’ grass counter and Parfit’s Future Tuesday Indifference with a shrug and an appeal to pluralism about practical reasons, without feeling as though there is something wrong. The approach developed by Street (2009) is to require some additional specification of the psychological states in virtue of which the ideally coherent eccentric holds the eccentric attitudes. In the anorexic case, it may make us uncomfortable to bite the bullet on the basis that we believe this person fairs worse, but we might still take a small-l-liberal approach and allow that individuals should be the arbiter of their own welfare, when that deliberation is reasoned and coherent.

6. Counterexamples to the proposed analysis of reasons

The challenge is that not all cases are like this. The most worrisome version of the ideally coherent eccentric cases is Gibbard's ideally coherent Caligula (Gibbard, 1999, p. 145). As Street writes (Street, 2009, p. 292),

One might be willing to grant that a weird enough character could have most normative reason to starve herself to death for the sake of a 'trim' figure, or to choose an agonizing operation on a Tuesday over a mildly painful one the next day. But one might think the situation is entirely different when it comes to ICEs [ideally coherent eccentrics] who see no reason not to inflict suffering on others.

If there is an adequate solution to the case of ideally coherent eccentrics, then that solution must be able to address the most pressing, high stakes examples.¹⁸ A solution that only gives results on the silly, low stakes cases is of little value if that solution cannot handle Caligula. Caligula is a case with actual stakes.

Structurally, these ideally coherent eccentrics cases are similar; they differ somewhat in the lessons one is expected to draw and in the way they flesh out the details of the case, but the form can be summarized as follows: There is some combination of an act ϕ and circumstance C such that we judge, for any agent S, that S ϕ -ing in C is irrational, unreasonable, or otherwise incompatible with the existence of a *pro tanto* reason. However, there is some possible ideally coherent agent A whose attitudes are such

¹⁸There is an alternative way of formulating these as undergeneration worries, where there is "no reason not to ϕ ." But these are less of a worry because, as discussed in section 5.4.2, reasons that act as disincentive often are constructed.

6. Counterexamples to the proposed analysis of reasons

that A believes, coherently and following reflection in A's values, that some fact F is reason for A to ϕ in C.¹⁹

Under these circumstances, we have an intuition that S ϕ -ing in C cannot be supported by a normative reason and an ideally coherent agent A whose values are such that A believes there is a reason for A to ϕ in C. The challenge is whether the constructivist can tell a story that reconciles these two parts.

In theory, theories of reasons that reject "subjectivism"²⁰ can simply block this claim by straightforwardly saying that in some cases there is just an attitude-independent, substantive constraint on being a reason such that there cannot be a reason for A to ϕ in C. As such, while A coherently believes there is a reason, A is mistaken because of this substantive constraint. So these non-"subjectivist" theories of reasons can just grant the circumstance and hold that while A believes there is a reason for A to ϕ in C, A is just mistaken.

The issue is not straightforward for constructivist theories of reasons. On constructivism, agents can construct their own reasons based on their attitudes; there are constraints on this construction, as discussed in section 3.3.2 and in further development below. When dealing with an ideally coherent agent whose attitudes are such that they want to construct a reason to ϕ in C, then there is a question as to what condition is stopping them. In order to preserve the intuition that there is no reason for A to ϕ in C, and

¹⁹Obviously, the formulations of ideally coherent eccentric cases vary in regimentation subject to author. All I am doing here is applying my own regimentation to the case.

²⁰I put "subjectivism" scare quotes here. I am focused on constructivist responses, and constructivism is regarded as a branch of "subjectivism" when the ideally coherent eccentric cases are presented. This "subjectivism" is considered a broader set that includes things like Gibbard's expressivism.

6. Counterexamples to the proposed analysis of reasons

the result that A is mistaken in believing that there is such a reason, one needs such a blocking condition.

Table 6.1: Quick reference for constraints discussed in the dissertation

Constraints	Discussion in the dissertation
Justification as necessary condition	chapter 2
Promoting and Protecting constraints	section 3.3
Factual/“only-if” constraint	section 3.3.4
Formal constraints	section 3.3.2
Constraints on valuing	section 3.2.2
Strictly better and comparative constraints	section 4.4.1

The alternative is to bite the bullet and suggest that, in fact, the intuitive finding in the ideally coherent eccentric cases is mistaken and that A is correct in believing that A has a reason to ϕ in C. As I intimate above, there are some cases of ideally coherent eccentrics where this bullet seems tasty enough. For grass counters and Future Tuesday Indifference, the consequences only impact the agent making the decision and so one might be inclined to just default to pluralism about reasons. In the case of Caligula, though, this is a bitter conclusion.

6.3.1 Straightforward practical reason construction

The concern about whether there is a reason for A to ϕ in C arises partly out of the fact that in virtue of having certain attitudes and ideal coherence, it seems as though A can just make it the case that there is practical reason for A to ϕ in C. Suppose that agent A wants to satisfy the aims that A has (something that seems trivially true, based on what it is to have aims); these aims include A’s positive evaluative attitude towards suffering.

6. Counterexamples to the proposed analysis of reasons

On some constructivist stories that allow positive evaluative attitudes to play this role in constructing practical reasons, we could leave it here. In the case of my own version of constructivism, we can further stipulate that the agent's attitudes include all of the relevant components of valuing established in section 3.2, including the beliefs, dispositions towards certain emotions, and dispositions to count as reasons. *Ex hypothesi*, A is also ideally coherent, and we might further allow that the standard defeaters for valuing bringing about some thing being of value do not apply. With all of this in play, instances of A ϕ -ing in C where such an act promotes or protects suffering of the kind A values satisfies the second condition of my analysis.

What about the reasons as evidence condition? Consider the interdependence of the conditions discussed in section 4.3. Since A ϕ -ing in C promotes or protects something of value and there are no other courses of action that promote or protect that thing of value in C, then the promotion or protection of that thing of value is evidence that A ought to ϕ in C, on a *pro tanto* basis.

While the reasons as evidence condition is present, it is unlikely to be sufficient to stop the existence of a reason in this case. Supposing that preference satisfaction is regarded as sufficient for the existence of a *pro tanto* reason under the justificatory norms, then the relevant values can produce a *pro tanto* reason. The claim that preference satisfaction is sufficient is often just taken for granted in these cases, as it grounds acts that promote the practical preferences of an agent. As a result, the reasons as evidence condition does not undercut the practical reasons.

6. Counterexamples to the proposed analysis of reasons

In addressing the ideally coherent Caligula, one might hold that there are substantive constraints in virtue of which the ideally coherent Caligula cannot form such a practical reason.

Consider a Kantian constructive approach like the one laid out by Rawls:

... this kind of view [Kantian constructivism] sets up a certain procedure of construction which answers to certain reasonable requirements, and within this procedure persons characterized as rational agents of construction specify, through their agreements, the first principles of justice. (Rawls, 1980, p. 516)

Rawls notes in a following parenthetical that he is intentional in using the word “reasonable,” and later elaborates that these impose substantive restrictions. On this view, the Kantian constructivist might hold that while being ideally coherent, if Caligula still hangs on to his joy in harming in the face of the construction of justice (and other related moral conceptions), then Caligula is being unreasonable, and should be excluded.

The problem with this approach is that it ignores the *pro tanto* status of the reasons for Caligula. While we recognize that moral reasons are frequently so heavy that they render practical reasons irrelevant in deliberation, that does not mean the practical reasons cease to exist. The satisfaction of my preference gives me reason to shoplift; the fact that my moral reasons not to shoplift far outweigh those preferences and their related reasons does not mean that the latter disappear. While the variation in weight has some limiting effect, that effect is limited.

6.3.2 The Socratic Gambit

One potential approach available in the case of the ideally coherent Caligula is to interrogate the elements of Caligula's valuing, in particular the belief that suffering is good or worth wanting and the disposition to treat the promotion (rather than prevention) of suffering as a reason for action. Interrogating the grounds for which agents believe or act is, on some understandings of disciplinary history, the founding act of philosophy. I call this the Socratic Gambit because it is characteristic of Socrates that interrogation of the beliefs of others leads to the modification of those beliefs.²¹

One could suppose that in the case of any ideally coherent eccentric (including Caligula), that one could engage and through rational introspection help the eccentric individual realize some substantive error in their evaluation or change their construction. One need not adopt any of Socrates or Plato's metaphysical commitments to a singular form of goodness or justice to hold such a position; rather one could hold that there is a pluralism and constructivism, but with substantive constraints that will exclude the ideally coherent eccentrics from having such reasons when engaged in rational deliberation, including Caligula.

In taking this position, one adopts the view that by engaging in such an exercise, the deliberative process in virtue of which Caligula values suffering can be changed. We see this in cases of direct interrogation of the reasons of other agents; under some circumstances, we influence the behaviors (incipient and future) of other agents, as discussed in section 5.4.2.

²¹It has been independently suggested that the root is in Plato, rather than Plato's character of Socrates. Given Plato's commitment to authoritarianism and realism, I am uncomfortable with invoking him under these circumstances.

6. Counterexamples to the proposed analysis of reasons

There is a barrier built into the case of the ideally coherent Caligula that creates a wrinkle. Given that Caligula is ideally coherent, on what grounds do we exercise such influence? Is it possible to exercise such influence?

By holding *ex hypothesi* that Caligula is ideally coherent, it will not do to suppose that there is some substantive error of reasoning such that Caligula's values can be undermined by demonstrating incoherence. Further, there is no reason to suppose that there is some factual basis that is necessary for Caligula valuing suffering such that Caligula's valuation will fail to give rise to the value. This leaves us with limited options to influence Caligula.

One might suppose that there is some substantive restriction on taking promoting the suffering of others as a reason to act. That restriction can be grounded relative to attitudes in one of three ways. First, it could be attitude-independent; however, this is off the table for constructivists, because of the prior commitment to attitude-dependence. Second, it could be dependent on the attitudes of each agent, but this would not get us anywhere with Caligula because of Caligula's differences in attitudes. Third, it could be dependent on the attitudes of various agents taken together.

The third exists somewhat in the constructivist literature already. In particular, substantive constraints are raised by so-called "Kantian" constructivists as limits on what reasons an agent can have. Usually, this is applied just to the moral domain.²² Holding that there is a local, contingent constraint limits the scope of the ideally coherent Caligula problem. It grants that there is no *necessary* constraint which precludes the existence of a Caligula, but holds that there is a constraint in the actual world and the communities that exist in that world. This is a partial answer to the

²²See [Korsgaard and Korsgaard \(1996\)](#); [Rawls \(1980\)](#).

6. Counterexamples to the proposed analysis of reasons

counterexample; the other part is biting the bullet on the logical possibility of an ideally coherent Caligula in other worlds, or other contingent circumstances.

6.3.3 Constructing other reasons

The most significant available course of action addressing the ideally coherent Caligula is that, given constructivism, communities are capable of creating reasons that influence the ideally coherent Caligula not to torture. Basically, because some subset of reasons are socially constructed (rather than constructed by the individual agent), this process of socially constructing reasons can be leveraged to influence behaviors.

This works on two levels. The first is in forming social facts that act as practical reason for compliance; there are certain regulative norms created within the society that are practical reason for the ideally coherent Caligula not to torture. These regulative norms provide practical reasons for individuals for refrain from certain actions (or even to perform certain actions, depending on context). These include the broad regulative assessment by the community, that it is wrong to contribute to the suffering of others. This first dimension is often discussed by Kantian constructivists (like O'Neill and Rawls) in terms of public reason; "Public reason... requires that our moral or political principles be justifiable to, or reasonably acceptable to, all those persons to whom the principles are meant to apply." (Quong, 2018, Sec. 1) In the case of the ideally coherent Caligula, a case of justification can

6. Counterexamples to the proposed analysis of reasons

still be provided to Caligula, even if he is not motivated by the public reason provided.²³

The second has to do with second-order regulative norms that are socially constructed to manage the reasons as evidence condition. These are intra-communal norms governing what counts as evidence some agent ought to act.²⁴ This is more complicated; I treat this at greater length in section 6.3.4.

The idea is a familiar one. There are practical reasons for individuals to avoid paying taxes through various means, but tax law, the expectation of audit and penalty create practical reasons not to engage in certain kinds of tax evasion or avoidance. In some cases, these social facts are matters of law; in some cases, there is a social stigma attached to tax avoidance.²⁵

In tax avoidance, there is a clear *pro tanto* reason to avoid paying tax; one has more money. However, the hope with the existing enforcement mechanisms is that individuals will recognize the possible penalties (whether legal or social) as strong disincentive.

Consider the same approach to the ideally coherent Caligula. We create disincentive not just for Caligula, but for everyone, to engage in torture and murder and the like. The threat of incarceration or other severe consequence

²³In lots of instances of public reason, there may be individuals who disagree with the reasons for which the community institutes regulatory norms. This is common. Prohibitions on certain criminal acts, for example, have both elements.

²⁴I largely ignore the question of whether these norms are constitutive or regulative. My suspicion is that some are constitutive and some are regulative. This distinction is important for parts of this dissertation, but not so much for this part.

²⁵Ask British comedian Jimmy Carr, whose legal tax avoidance became a major national news story and is still the subject of mockery almost a decade on. While his behavior may have been strictly legal, the consequences of public awareness and the impact on his career were quite serious and resulted in paying back the tax he had avoided.

6. Counterexamples to the proposed analysis of reasons

is supposed to act as such disincentive, as countervailing practical reason pushing back on the practical reason the agent has.

A respondent might try to circumvent this issue by supposing that Caligula could simply not care about any of these consequences. However, this is a short-lived pivot. The society can simply leverage any and everything that Caligula does care about as potential object of sanction.

This substantially addresses the practical problem, at least insofar as it provides motivation for Caligula subject to what Caligula values. Threatening those other values, making it the case that protecting his other values involves not murdering and torturing, is a useful tactic for motivating social compliance. This provides a constructivist route to answering the counter example. Conveniently, it also illustrates how constructed reasons (even those acknowledged by non-constructivist theories) can play a role in justifying and motivating desirable behavior in eccentric agents, thereby creating reasons for compliance with the norms.

Unfortunately, I suspect that some readers will find this of limited value in addressing the counter example.

The ideally coherent Caligula could be understood as a brute counterexample where we present an agent (Caligula) and that agent's *pro tanto* reasons to act in an undesirable way, where the problem is the actual inclination to act in that way. This objection can be addressed by the creation of countervailing reasons that prevent the inclination, on the basis of providing decisive *pro tanto* reason not to act in this way. However, in formulating the case of the ideally coherent Caligula above, I note that the problem in the case is not merely Caligula's inclination to ϕ , given the *pro tanto* reasons

6. Counterexamples to the proposed analysis of reasons

related to suffering, but the *existence* of the *pro tanto* reason for Caligula to ϕ in C.

My approach does stop the inclination, and may ensure that there is never a *pro toto* reason for Caligula to ϕ in C. This approach does not stop the existence of the *pro tanto* reason for S to ϕ in C; the approach only pushes back by creating reasons for S not to ϕ in C.

My approach to addressing Caligula is about motivating compliance and (if that fails) doing damage control. It is not about changing the underlying structure of the practical reason for Caligula to torture, but about trying to address the possibility that Caligula will torture. Is this adequate? I suspect judgments will be mixed. Some may hold that a Caligula who has a practical reason to torture but (as a matter of the countervailing reasons) cannot have a *pro toto* reason to torture is a sufficiently declawed version of the thought experiment. Others may hold that the shift from the existence of the *pro tanto* reason to the existence of the *pro toto* reason is a sort of philosophical sleight-of-hand. Both of these positions strike me as reasonable.

There is a further upshot of exploring this point: it illustrates the tie-in between historical discussions of rationality and reason (on the one hand) and social order and law (on the other hand). Section 5.1 is a pivot from the abstract constructivist theories of reasons towards that connection, but the importance of that connection to constructivist theories of reasons is illustrated in handling cases like the ideally coherent Caligula. The reason that Kantian-inclined scholars are interested in public reason and Humean-inclined scholars are interested in the anthropological and sociological development of cultural norms is intimately related to this issue. The ways in which we connect to and influence each other to stop us from acting in ac-

cordance with our worst inclinations is a requirement of good constructivist discourse.

6.3.4 Potential reasons as evidence constraints

A second constructivist approach is for communities to develop norms that apply to the reasons as evidence condition, such that there is a substantive constraint on what there is reason for an individual to do. There are constraints on the construction of reasons, as I discussed in section 3.3.2, but the nature and extent of these constraints is subject to some debate.

The core idea in the case of Caligula is something like this: there are a subset of things that are such that S ϕ -ing in C promoting or protecting them is not acceptable as evidence that S ought to ϕ in C. On my analysis, Caligula may satisfy the promoting and protecting condition for being a reason, but will fail to satisfy the reasons as evidence thesis, because promoting the agony of others is precluded from being evidence S ought to ϕ in C.

Even for someone sympathetic to this response, there are a few worries about this. First is that it rests on adopting a constructivist view of norms governing evidence in the reasons as evidence condition. Second is that it depends on contingent facts about the community. The second is addressed directly in section 6.4; I will get to addressing the first later in this subsection.

The basic idea of this response to the ideally coherent Caligula proceeds by first granting the possibility that there are no substantive, attitude-independent constraints on reasons. In section 3.3.2, I present the idea that some facts drawn from the logic of rational choice impose formal constraints on reasons. A formal constraint only considers properties of the

6. Counterexamples to the proposed analysis of reasons

facts and attitudes relative to other facts and attitudes, like the ordering of preferences. Substantive constraints are those that consider the facts about the agent, circumstances, and act.

Consider another chess example. Steve is considering possible moves. Steve is in the midgame and playing black. Steve believes that he has strong reason to move his pawn on f7 to build a particular pawn structure. However, his opponent's queen is on h5. Given the state of the board, Steve moving his f7 pawn is illegal, because the move would put his king in check (by the queen on h5). As such, there is a substantive constraint on what there is reason to do; that constraint prohibits a reason for that move.

Just to clean up this demonstration, one might say that the fact about the desired pawn structure is a *pro tanto* reason for Steve to move the pawn in that situation, but it is defeated by some other consideration. On the promoting or protecting condition, this analysis works; promoting the valued pawn structure is a reason to make the move (assuming no false beliefs are grounding the valuation of that pawn structure). On the reasons as evidence condition, it does not, because no fact can justify making a move which is not permitted, at least within the rules of chess. (The rules for chess hustlers and cheats are slightly different.) The constitutive conventions of chess are prohibitive, such that the act type is just impossible because it would violate the rules. It is not the case that the pawn structure is evidence Steve ought to ϕ in C, as ϕ -ing in C is outright prohibited by the rules of the game.

This substantive constraint is imposed by the constitutive norms of the game of chess; as such, these substantive constraints are socially constructed.

6. Counterexamples to the proposed analysis of reasons

This creates a new subset of potentially worrying cases, namely that there are those who benefit by gaming or outright ignoring the norms governing evidence one ought to ϕ and are able to do so without any consequences. Broadly, these can be characterized as free-riders with respect to reasons, social norms, or other features than an individual may defy.²⁶ The general idea of such cases is that there are social arrangements where individuals can exploit the rules through non-compliance without fear of consequences.

The appropriate response to such free-rider cases is to note that such cases actually shift away from the substantive Caligula objection. They acknowledge that there is no reason as a result of constraints on construction, but rather raise shift the burden back to the issue of compliance. Even if it is the case that, according to the constraints on construction, Caligula has no reason but stands out on the basis of a desire to act anyway. This is an issue with pressuring the agent to comply, and circles back to the above discussion of creating countervailing practical reasons.

I say above that imposing substantive constraints on the reasons as evidence thesis first means excising the notion that there are attitude-independent substantive constraints. This is not strictly true, but rather only instrumentally necessary for dealing with the Caligula case. If we held that there were attitude-independent substantive constraints in that case, then we could just apply those constraints to Caligula and be done with it. (This is the world realists want.) In the process, we would have given up the thoroughgoing attitude-dependence thesis required for constructivism.

²⁶This may include Plato's Ring of Gyges discussion in *The Republic* (Jowett et al., 1888, 2:359a–2:360d) and other historical variations.

6. Counterexamples to the proposed analysis of reasons

The constraints on what counts as evidence that S ought to ϕ , and therefore what satisfies the reasons as evidence condition, is partly constituted by norms governing what is relevant in justifying a course of action. Instead of supposing that there are attitude-independent “rules of evidence,”²⁷ we hold that as in a courtroom, it is a matter of constitutive and regulative norms that hold what does and does not count as evidence one ought to ϕ in C. As such, enforcing these norms allows for the imposition of some substantive restrictions on what there is reason to do. Some facts are regarded as irrelevant because the constitutive norms governing what counts as evidence are not satisfied.

Suppose that as a matter of the justificatory norms, the fact that Caligula ϕ -ing in C causes suffering is not recognized by the community as providing a justification for Caligula to ϕ in C. Rather, it is treated like the fact that Hilda is cold is a reason for David to put on a jacket; the two things are not appropriately related such that Hilda being cold is not evidence that David ought to put on a jacket. On this view, then it would be the case that while Caligula’s views satisfied the promoting or protecting condition, it would be the case that as a matter of norms, this would not give rise to a reason. The reasons as evidence condition is necessary; it would never be satisfied by Caligula on this matter.

Some people may reasonably be squeamish at the idea that the “rules of evidence” are socially constructed in this way. There is a reasonable

²⁷I am putting the phrase “rules of evidence” in scare quotes here, because it is quasi-metaphorical. There are explicit rules of evidence in courts of law; there are not explicit rules of evidence in most areas ordinary contexts, including the broad framework of epistemology. The notion of “rules of evidence” is just meant to invoke whatever higher-order facts there are such that a fact F is evidence for a proposition p. The rules by which we can find that F is evidence that p are the “rules of evidence.”

6. Counterexamples to the proposed analysis of reasons

argument to be made that there are some attitude-independent elements to the rules of relevance as discussed in [Hare \(1978\)](#). In the context of epistemic reasons, for example, there are notions of relevance in scientific explanation that one might argue are attitude-independent.²⁸ I cannot get into the merits of these arguments at the present time; there is a lot of baggage there and it falls well outside of the scope of this dissertation. However, even if we allow that there are these attitude-independent constraints on relevance, and that we may be able to introduce the attitude-dependent constraints to leverage substantive constraints on reasons, that does not undercut the prospect that some constraints are socially constructed (as in the chess case, where the constitutive norms are a necessary condition).

6.3.5 Agents without societies

There are two worries stemming from the approach to the ideally coherent Caligula developed in the preceding subsection. Both are ultimately reducible to the worry that the contingency of social norms leaves any potential response to Caligula at the mercy of the society (or lack thereof) in which Caligula is situated.

The first response is to worry about what one might say if Caligula is not situated in a society at all, and therefore has no benefit of these social norms in responding to Caligula.

The second response is to worry about what might happen if the society just did not have such social norms or was constituted mostly by Caligula-

²⁸[Strevens \(2008\)](#) discusses such an approach that centers on difference making, which is core to several such accounts of explanation.

6. Counterexamples to the proposed analysis of reasons

like figures. I will dedicate significant discussion to this second point in section 6.4, but the first also merits treatment.

How do we consider agents without societies? What norms govern the “rules of evidence” or any constraints on justification for agents outside of societies?

The formal constraints on justification will still apply, but this is not sufficient to block out the ideally coherent Caligula. Given the stipulation that Caligula is ideally coherent, then this removes the possibility that formal constraints will do the work of blocking the formation of reasons *ex hypothesi*.

One straightforward way to consider the case of the ideally coherent Caligula outside of a society is just to suppose that such an entity lives independently of any other agents. Perhaps there are other agents around, but they are not organized into a community in the relevant way. What one arrives at here is a sort of “state of nature” variation where there are no norms or mechanisms of enforcement in place to drive Caligula’s behavior.

Unfortunately, there is no satisfying answer available that would simply hold that, even in such a non-social situation, Caligula would be unable to construct such reasons. One point under these circumstances is that even Caligula should recognize the possibility to external enforcement mechanisms; the inclination to eliminate threats is pervasive and may be recognized as reason by other agents. We kill or otherwise neutralize threats in the state of nature, when possible. If Caligula is a predator, harming others in the surrounding area, then this behavior incurs risk of retaliation. Unlike the justificatory constraints, this does not preclude the *pro tanto* reason for

performing acts that cause suffering; it just provides a countervailing consideration, given that Caligula's motivation set includes self-preservation.²⁹

On the committal constructivist position, there can be no internal constraints that prevent the formation of reasons for Caligula. When we remove the influence of other agents as members of a shared society, then we remove the prospect that such influence may block the formation of some reasons. What we are left with is the prospect that certain countervailing considerations provide reasons that may limit the formation of a *pro toto* reason.

6.4 Abhorrent societies and further worries

So let me ask you a question about this brave new world of yours. When you've killed all the bad guys, and when it's all perfect and just and fair; when you have finally got it exactly the way you want it, what are you going to do with the people like you? The troublemakers. How are you going to protect your glorious revolution from the next one?

–The Twelfth Doctor (Peter Capaldi)

Doctor Who 09.08, “The Zygon Inversion”

One objection to constructivism about reasons (and many other varieties of constructivism) is that it entails some version of relativism.³⁰ This is

²⁹There are some potential discussions about the mathematical grounding of cooperation, in virtue of which certain practical principles of cooperation and social interaction among conspecifics may be necessary, but these are subject to significant variation in circumstances. The doctoral work of my friend and colleague Archie Fields III considers some of these questions.

³⁰See (Baghramian and Carter, 2020, Sec. 4) and Gowans (2019).

6. Counterexamples to the proposed analysis of reasons

often just flatly taken as an objection, though it may just suffice as a simple observation. Flatly, constructivism does entail some version of cultural and moral relativism.³¹

There are two ways to go about setting up this issue. The first is as a brute response to constructivism. The second is as a follow-up to the ideally coherent Caligula case. In the end of the preceding section, I laid the groundwork for this second approach. For a moment, let us bracket the ideally coherent eccentric cases and just look at the first approach.

The objection to cultural relativism as a consequence is that it results in an uncomfortable and perhaps counter-intuitive claim that there are instances where individuals within a community have reason to engage in morally repugnant behavior on the basis of the accepted norms within that community. [Boghossian \(2007\)](#) discussed infanticide as such an example, but one might equally choose other historical practices like genocide, slavery, rape, etc. Human history is ripe with cases where social facts sanctioned and even encouraged a range of behaviors that are morally repugnant.

Frequently, this implication of moral or cultural relativism is used as a sort of broad, cudgeling objection to various views on which what counts as a moral or immoral act in a society is depend on the social facts or attitudes within that society. These examples often take the form of incredulity and repugnance; “how could one believe that such-and-such a behavior is morally permissible? The mere acceptance in the society doesn’t tell us that the behavior is acceptable.”

³¹Constructivism also entails some version of relativism about practical reasons, which one might argue gives rise to these versions of cultural and moral relativism. Set that issue aside.

6. Counterexamples to the proposed analysis of reasons

This is deeply intuitive and forceful, even without a clear argumentative structure. Surely it cannot be right that, were American slave holders or Nazis coherent, that they would be justified in enforcing slavery or genocide through their legal order. Surely it must be that making such laws is morally wrong, independently of the attitudes of any agent.

For our purposes, though, one might develop a very minimal argumentative structure that does not require any consideration of what it is for a reason or act to be moral.

The theory of reasons presented here entails the thesis that for some cases of a reason R , the existence of R for S to ϕ in C depends on the social and cultural situation for R . The worry goes that there will be instances where there is an R for S to ϕ in C , where ϕ -ing in C is morally repugnant. We can borrow the Caligula case for this example; we could also just talk about the range of repugnant behaviors in human history for which the cultural norms of various communities seemed to create a practical reason.

There are a few ways to attack this issue. The first is to attempt to formulate substantive constraints that would apply to whole communities. This is not so feasible for constructivists, because constructivists hold that some fact being a reason is contingent on attitudes.³² As we saw in the ideally coherent Caligula case, the best available course is to make use of the social elements of that dependence, but that is not available, since the social ele-

³²Rawls (1993, 1999) attempts to develop an international approach to setting boundaries which would apply even given the differences in social values across communities. However, this project has much more serious problems than Rawls' approach to negotiating agreements from the original position amongst individuals who will share a community. It turns out, just scaling up the methods of small-l-liberalism is not entirely adequate and that cross-cultural divergences in values create serious problems. I am grateful to John Baker for drawing my attention to this part of Rawls' work in our conversation about the problems of cultural relativism.

6. Counterexamples to the proposed analysis of reasons

ments are already established. We could try to cast a broad substantive constraint, but existing constructivist arguments against attitude-independent theories of reasons make that tricky.

The second is to advance some arguments for setting certain political norms. These political norms would themselves be contingent; some societies do not and will not accept values like pluralism and non-malevolence. The purpose of this second move is not to suppose that all societies are bound to some form of liberal democracy, but rather to hold that the truth of constructivism about reasons provides strong *pro tanto* reason for some version of liberal democracy.

For a group of individuals, it may be the case that there is reason for one member to ϕ in C while there is not a reason for another to ϕ in C. Individuals may have variation in what there is *pro tanto* and *pro toto* reason to do even in relevantly similar circumstances.³³ Given this variation in reasons, the community should allow individuals to act in accordance with those reasons all other things being equal.

The idea is simple enough; suppose there are different practical reasons for what Ann and Ron are planning to pursue in their careers, because of differences between their attitudes. All other things equal, the community should allow Ann and Ron to take these different paths. The community and individual members may try to influence Ann and Ron based on other values, but individuals should be allowed to act in accordance with their own interests. This is a central position of small-l-liberal political philosophy,

³³This version of pluralism about practical reasons is entailed by constructivism about reasons; there are non-constructivist theories of reasons which may also hold this claim.

6. Counterexamples to the proposed analysis of reasons

even as the limitations on this freedom are subject to ongoing practical and moral dispute.

The core of the move is that our metaethical theory can inform political theories, and even provide arguments in favor of certain positions in political and moral philosophy. This may seem a weird move based on direction of fit, but I submit that it is not that strange to suppose that normative theory requires consideration of descriptive theory (including descriptive metaethics) and that this is a normal state for lots of moral philosophy.

With this in mind, consider a few approaches below which may work as ways of addressing the abhorrent societies.

6.4.1 Social subgroups and societies-within-society

Societies are not homogeneous groups of actors. Societies contain structures and different complex power relations and societies within societies. The United States is a society that includes a range of formal societies as part of its constitution, like the judiciary, as well as formal and informal societies that are not recognized as part of the larger society. So while the United States may have norms, there are various subordinate groups within the United States that may have their own norms and are constituted differently.

For example, a mafia family may have regulative norms that explicit contradict the explicit legal norms of the United States. The United States holds that perjury is bad and that it is punishable with time in prison. The mafia family holds that perjury is good and that perjury is rewarded (and fail to do perjury is punishable). There are two sets of problems that come out of cases like this.

6. Counterexamples to the proposed analysis of reasons

The first set of problems are really about compliance. Tony Soprano is a member of two societies with conflicting norms. He is a member of the United States and he is a member of the mafia. When the anti-perjury norms of the United States and the pro-perjury norms of the mafia provide conflicting reasons when Tony has the opportunity to commit perjury, then Tony has to engage in practical deliberation about whether to commit perjury. This becomes a problem of compliance, trying to compel members of both the United States and the mafia to comply with our preferred norm (the anti-perjury norms, I hope). This is not such an interesting problem, as it just reiterates the compliance problems raised above, in the discussion of individual agents like Caligula.

This is not so interesting as a version, since all moral and political theories have compliance problems. Within any given system, some portion of the agents within that system will violate the relevant norms. This is a problem for everyone, on more-or-less equal terms.

The second set of problems relate to the ways in which some societies are parasitic on other societies. Some societies only exist in virtue of other societies. The existence of the American judiciary is a necessary condition for the existence of the American Bar Association, the Federalist Society, and other such organizations. These societies often include norms that are parasitic on norms of the broader society; for example, if becoming a member of the judiciary means having certain qualifications and one of the explicit goals of the Federalist Society is to install its members in the judiciary, then the Federalist Society is likely to (and does) have norms in favor of acquiring the relevant qualifications.

6. Counterexamples to the proposed analysis of reasons

The challenge with these cases is more subtle than the mere compliance puzzle. Suppose that the mafia adopts norms which are not outright non-compliance with the norms of American society, but in some way modify the norms of American society. Instead of bribing a judge to break the law, they influence legislators to change the law without doing anything that violates the law. Lots of organizations have explicit or implicit interests in substantively changing the norms of the broader community in which they are situated; some may act legally in doing so. In these cases, there is a more serious problem. My hope is that the discussion below constitutes a partial answer, though the reality of such organizations is that there are going to be instances where organizations attempt to influence the norms of societies in ways that we oppose.

6.4.2 Open society and small-l-liberalism

In the early- and mid-20th century, political philosophy had to reckon with a long history of genocide, chattel slavery, and other such institutional forces in societies present and historical. Such political philosophy was confronted with actual societies that were abhorrent and (at least practically, though not ideally) coherent. In trying to adapt a political ideology that would have both justificatory and motivational force when engaging such societies, the political philosophical literature came up with a few potential courses of argument.

These are meant to provide substantive reasons for abhorrent societies to stop their abhorrent practices and institutions, based on general, shared understandings of what is of value. Ultimately, the most successful solutions

6. Counterexamples to the proposed analysis of reasons

proposed (in my view) turn on a hope that abhorrent societies share some concept of impartiality, or at least have a relevantly similar conception such that the argument force carries over. However, this is only effective, on a constructivist theory, insofar as it accords with what the abhorrent society values. In this respect, the proposal for abhorrent societies is much like the proposal for the ideally coherent Caligula wherein some countervailing reasons are provided that we hope will provide a stronger counterweight; the idea is that such reasons will count against existing *pro tanto* reasons such that what the abhorrent society has most reason to do is refrain from those abhorrent acts, institutions, norms, etc.

The standard approach developed as a response to worries about moral relativism come out of the political philosophy literature, as the proliferation of anthropological and sociological literature produced acknowledgment of descriptive relativism in anglophone philosophy and the “superior western civilization” myths collapsed under the weight of moral guilt at colonialism, fascism, and other such ideological elements in world history.

The core idea of these analyses is an attempt to balance certain moral claims against a broad pluralism about moral and practical reasons. There were some acts and institutions which were regarded as impermissible on some demarcation criteria³⁴ and then everything except those acts and institutions was regarded as permissible (at least subject to the rules of the legal order) if established within the appropriate norms of that social order. As such, societies and individuals were allowed to govern themselves within certain constraints, and societies and individuals were allowed to

³⁴The discussion in Popper (1957) considers a range of possible criteria for demarcation, including whether or not authoritarianism is outright prohibited.

6. Counterexamples to the proposed analysis of reasons

engage in first-order disagreement about what was morally permissible and impermissible, and set norms on their own terms.

The open society proposal can be summarized as a short argumentative series from the metaethical to the moral to the political. In the metaethical domain, it turns out that norms are socially constructed and so are subject to reasonable variation at both the community and personal level; pluralism about both practical and moral reasons is true. This influencing the moral domain by holding that individuals and communities ought to be allowed to act in accordance with their reasons (within certain parameters). As I noted above, this follows partly out of a historical context of concerns about colonialism and forced assimilation, where the history of violence (including forced assimilation) of a range of groups (indigenous people, Jews, etc.) were justified by a thoroughgoing monism about what one ought to do. That monism entailed some variety of religious adherence, certain lifestyles, subjugation by government, and so on. The moral repugnance at this history, especially in the period following two World Wars, is hopefully one with which we can empathize.

The political move is then to indicate that the goal of such societies should be to allow individuals to participate in the formulation of norms for that society while ensuring that the rights of minority communities to act in accordance with their interests and values is preserved. The society is discursively and practically “open” in the sense that individual members are able to advocate for reforms in the norms governing society and are allowed to act in accordance with their own, varied practical reasons. The adoption of an open society requires certain first-order principles of governance, namely that reforms in societies which restrict the rights of individuals to

6. Counterexamples to the proposed analysis of reasons

act in accordance with their interests are to be regarded as strong suspicion and disfavor, but others (features like tax structure, state religion, etc.) are left to the negotiation of the community.

This move is a core thesis of Walzer's *Spheres of Justice*,

... that the principles of justice are themselves pluralistic in form; that different social goods ought to be distributed for different reasons, in accordance with different procedures, by different agents; and that all these differences derive from different understandings of the social goods themselves—the inevitable product of historical and cultural particularism.” (Walzer, 2008, p. 6) In Walzer's case, it is explicitly grounded by both the metaethical constructivist thesis. He begins the paragraph with the radical and unambiguously constructivist claim, “Justice is a human construction, and it is doubtful that it can be made in only one way.” (Walzer, 2008, p. 5)

Walzer lays out the argument that the conditions for a society being just are in part a function of the appropriate agreement of the members of that society. However, the development of such a society requires the participation and consideration of the interests of all members of the society capable of participating. Walzer's argument is that the construction of a just society has a constraint that requires at least some consideration of the interests of all or most members of that society.

The objection proposed by critics on the basis of cultural or moral relativism may then proceed to criticism of Walzer's claim on the basis that there is no substantive constraint which precludes a society from the creation of

6. Counterexamples to the proposed analysis of reasons

a permanent, unrepresented subordinate class. (As a matter of historical fact, a disturbing number of human societies had and have such subjugated classes.) If justice is a construct and there are no substantive constraints on the content of that construct, then why should we privilege the societies where there is equal participation amongst agents over those where there is a subjugated class? Surely these are merely differences in construction and (as there are no substantive constraints) there is no metaethical reason to exclude one or the other.

Because, like the Empire, constructivism strikes back.

The point that Walzer and other constructivists advancing small-l-liberal democracy have in mind is not that this provides a substantive constraint, but that it provides a strong *pro tanto* reason for constructing and understanding justice in the more inclusive way.

Rawls, in discussing Kantian constructivism, puts this forward as a hope, rather than something that is required to successfully construct a notion of justice. “Our hope is that there is a common desire for agreement, as well as a sufficient sharing of certain underlying notions and implicitly held principles, so that the effort to reach an understanding has some foothold.” (Rawls, 1980, p. 518) There is not a claim that there will be such shared notions or principles, or that there are some principles that we should all share.

To justify a Kantian conception within a democratic society is not merely to reason correctly from given premises, or even from publicly shared and mutually recognized premises. The real task is to discover and formulate the deeper bases of agreement which one hopes are embedded in common sense, or even

6. Counterexamples to the proposed analysis of reasons

originate and fashion starting points for common understanding...(Rawls, 1980, p. 518)

Appealing to these small-l-liberal values and allowing a broad range of voices in social participation allows for more individuals to represent and advocate for their own interests, while simultaneously recognizing the variation in the reasons of individuals. While some individuals may not themselves value this in themselves, they may still recognize the value system as instrumental in ensuring their own autonomy and pursuit of what they do value. Even as constraints may be laid out and frustrate some of those pursuits, the value of this conception of justice is in providing a *pro tanto* reason to everyone, on the basis of allowing their pursuit of desires without subjecting the permissibility of such pursuit to authority.

6.4.3 Revolutionary approaches and equal consideration of interests

The open society approach to addressing these societies is sometimes regarded as too moderate and too permissive of certain forms of abhorrence that may be protected within an open society. The alternative is an approach to preserving pluralism while simultaneously imposing a strict prohibition on diminishing the interests of those who may not have the power to protect their own interests.

Rather than allowing the society to be open, the argument takes the broad metaethical and moral themes of the open society argument presented above, but then pivots to a stronger first-order conclusion, holding that a society has a positive obligation to protect the interests of groups

6. Counterexamples to the proposed analysis of reasons

who are less capable of protecting their own interests. The result is that the society is less open, because it restricts certain forms of discourse and action that transgress against the interests of more vulnerable communities.

This is a somewhat more radical approach to the one that comes out of the small-l-liberal views on public participation. Strictly speaking, it is not incompatible with those liberal approaches, though it does create points of tension. The general idea is to focus on advancing the interests of disadvantaged groups, with the expectation that advantaged groups (by virtue of being advantaged) will be perfectly capable of preserving their own interests. Rawls' assessment from behind the veil of ignorance is an attempt to clarify this principle. Like Walzer, Rawls takes the normative structures to be constructed, but holds that there are strong reasons to construct them in egalitarian ways. The veil of ignorance is an exercise in applying impartiality and recognizing the welfare considerations of other agents, in particular disadvantaged agents.

There is an enormous literature on Rawls, and I will not presume to rehash it here. My goal here is just to contextualize Rawls as a potential set of radical proposals for substantive principles that address abhorrent communities. One should not presume Rawls' conception of justice is necessary, that it is one that societies must adopt as a matter of substantive requirement (or cannot violate as a matter of substantive constraint); rather, it simply provides a strong *pro tanto* reason for those societies to be less abhorrent.

There are a range of arguments for impartiality within a society. Both the open society and radical approaches to addressing abhorrent societies ultimately turn on arguments for particular conceptions of impartiality,

6. Counterexamples to the proposed analysis of reasons

with the hope that such societies will adopt such a conception as a part of the framework and thereby eventually move towards less abhorrent norms. This is optimistic, and perhaps even Panglossian, as an assessment of the ways in which societies might change.

In both of these proposals for substantive *pro tanto* reasons, one might reasonably wonder whether the facts provided can satisfy the motivational role for the communities. In the case of the open society approach, there is a direct self-interest consideration in virtue of which a society, or at least a significant portion of its membership, may be appropriately motivated.

6.4.4 Relative weakness of these responses in contingent circumstances

There is a fairly serious problem for those worried about such case, which is that I have not excluded the existence of *pro tanto* reasons from abhorrent societies. As a matter of fact, there can still be such *pro tanto* reasons. Not only are such societies conceptually possible; some versions of them may well be actual. There is a relevantly similar issues with the Caligula case, which is that a Caligula in a society that does not institute constraints on reasons (or outside of societies altogether) can still have such *pro tanto* reasons to engage in the relevant abhorrent behavior.

When we get to this point, this is where constructivists do have to be honest about the bullet. These constructivist theories have only limited tools. Societies create reasons that the societies then bring to bear on members of communities; societies that want to create and enforce reasons for members to act in abhorrent ways are going to be able to effectively construct those

6. Counterexamples to the proposed analysis of reasons

reasons. Reasons of commerce, for example, related to abhorrent practices arguably obtain in our present society. The existence of moral prohibitions on human trafficking provide a countervailing *pro tanto* reason for agents not to engage in such trafficking, but there is a serious question whether there is an adequate counter-weight of the practical, financial reasons for those who engage in human trafficking.

Consider these cases in terms of the extension of being a reason. Suppose the ideally coherent Caligula and abhorrent society cases are instances where an interlocutor argues that there just is not a reason for the repugnant behavior in question. On certain realist theories that are appropriately restrictive, and exclude even practical reasons for such abhorrent actions, there may be the comforting corollary that there are no such reasons in these cases. But the question remains as to whether the judgment that Caligula and the abhorrent society is grounded in the best reflective judgment of the case or merely our revulsion at its consequences.

Sometimes the appropriate response to certain facts is repugnance, but this does not provide us any reason to doubt the truth of those facts. Obviously, this is more complicated in the cases of moral judgments, where our reactions to potential case judgments are (at least on many accounts of how moral judgment work) a source of insight into the truth of those moral judgments. The facts about crimes against humanity throughout the world provoke a judgment of repugnance; the facts that individuals in positions of political power stand-by also provokes a judgment of repugnance, along with a set of moral evaluations about the character and actions of those standing by. The repugnance of the facts about crimes against humanity does not give us reason to doubt the existence of those facts; however, the

6. Counterexamples to the proposed analysis of reasons

facts about those standing by may well lead us to doubt certain claims about whether the existing structure of power is just, authoritative, or any other positive moral assessment.

Which set of judgments is at play in the case of talking about Caligula and the abhorrent society? Is the existence of reasons for the abhorrent society to subjugate members just a straightforward fact that is not effected by its repugnance? Or does the repugnance substantially alter what to believe about the judgment?

I certainly do not want to believe in the possibility of abhorrent societies, though (in my experience) what I want to believe is rarely relevant to what there is. And sometimes the fact that I don't want to believe something is an indicator that there exist good independent reasons to believe that thing.

Chapter 7

Making reasons matter more

7.1 Revisiting the course of the dissertation

The dissertation develops and defends an analysis of being a reason that is designed to be more restrictive than the available constructivist and other attitude-dependent alternatives.¹ The available constructivist approaches to reasons over-generate the extension of being a reason, leaving them susceptible to counterexamples, and so the upshot of my analysis is that it looks to exclude those prospective counterexamples.

My analysis focuses on the intersection of two existing analyses: the reasons as evidence analysis and the constructivist analysis based on the relationship of an act to things of value (where being of value is understood in constructivist, attitude-dependent terms). The focus on the intersection prevents the various objections to constructivism through the use of the reasons as evidence condition, thereby avoiding a range of popular and strong attacks on constructivist theories of reasons. The dissertation also considers some constructivist arguments against the reasons as evidence thesis, in order to show that the reasons as evidence thesis itself also over-generates.

The analysis I advance goes:

¹One might note that I only briefly consider extensional issues for attitude-independent theories of reasons, and do not focus on these as arguments against attitude-independence as such. Rather, when I consider (for example) extensional arguments against the reasons as evidence biconditional in sections 2.5 and 2.6, I am focusing on a different issue with that biconditional analysis. I am not focusing on the biconditional analysis' attitude-independence. This is because attitude-independent theories of reason tend to involve curve-fitting with respect to intuitions about the extension of being a reason. While I consider arguments against attitude-independence in section 3.4.2, the extensional arguments are not at play. This is because attitude-independent theories of reasons can just suggest that the attitude-independent facts in virtue of which some fact F is a reason just are the facts that generate the appropriate extension. Unless such a move conflicts with other commitments of the attitude-independent theory, the attitude-independent theory can adjust for any variation of the extension.

Some fact R is a *pro tanto* reason for an agent S to do an act of type ϕ in a circumstance C iff (1) R is evidence that S ought to ϕ and (2) R entails S ϕ -ing in C promotes or protects something of value.

The purpose of this analysis is to develop an account that makes explicit and makes use of both the justificatory and motivational roles. As I discuss in section 4.3, it is possible that one might be able to combine these conditions with a bit of modification and come away with a more elegant analysis that includes only one condition. This could be achieved by taking the motivational role condition (the promoting and protecting condition) and specifying the constraints on being evidence such that those constraints are part of how we understand “promoting or protecting something of value.” While this is plausible, I caution against it in chapters 4 and 5.

The reason that I caution against it is because to do so would be to miss a major contribution of the analysis itself: the analysis spells out the distinct roles that reasons play (justification and motivation) and tries to illustrate the way in which each of the roles contributes to some fact being a reason. To subsume the justificatory role as an entailment of the motivational role (and the constraints developed in section 3.3.2) would be consistent with some available approaches to internalism about reasons and motivation, but it would lose the methodological element that shows this interrelation, especially as the analysis develops social considerations in chapters 4 and 5.

The broader methodological commitment of this dissertation is to functionalism about reasons, as discussed in section 1.4. Some fact is a reason in virtue of the fact that it functions as a reason. Functioning as a reason

involves and includes motivating agents (or having the ability to motivate agents under some conditions), but it also includes justifying the act. Justifying is a part of the analysis of reasons; it is among the roles reasons play in our personal and social lives.

Bringing the way in which some fact justifies an act into the foreground is important; to treat justifying as merely motivating under appropriate constraints misses the importance of justification. It misses justification as an independent practice and subject to norms of assessment that add value to discussions of reasons, by making clear that lots of our motivations are not reasons. While some theories of reasons do just build the relevant kinds of constraints into a motivational theory of reasons, and may be able to achieve extensional equivalence this way, these methodological commitments lead me to argue we ought not do so.

We recognize that the immediate inclination to use a narcotic may not be a reason to use that narcotic, given the reflective attitudes of the agent; the urge to scratch at a fresh tattoo is not a reason to scratch the tattoo. There are constraints to how a motivation gives rise to a reason, and I develop those in discussing the motivational role in chapter 3, but many of those constraints become explicit in the justificatory role. Even if we could subsume the constraints imposed on justification into the motivational role between mere motivation (on the one hand) and motivation that gives rise to reason (on the other), to do so risks losing sight of the importance of justification. Scratching at fresh ink is not justified, even with the immediate inclination, the expectation that it might bring the itching to cessation, and so on.

The social roles of reasons discussed in chapter 5 represent further important explananda for reasons talk.² The personal roles of reasons in justification and motivation contribute to the social roles of reasons. I do not include social roles in the analysis, in order to preserve some parsimony and because I hope that explaining how the justificatory and motivational roles of reasons taken in conjunction give rise to these social roles is more instructive and elegant.

My goal in setting out to develop this dissertation was circumscribed by a set of success conditions. A successful theory of reasons should explain, at a basic level, the different functional roles that reasons play. While accounts like those surveyed by Alvarez (2017) provide insight into each of the roles, they often do so by focusing on a single role, and therefore wind up deficient both with respect to over-generation (because mere justification and mere motivation will over-generate) *and* with respect to the relations between the roles. The fact that some fact motivates S to ϕ in C may obtain independently of the constraints on justification, but understanding what would provide motivation while providing justification has broader use. A compulsion for a drink is psychologically interesting; a compulsion combined with an understanding of the rational constraints in virtue of which that compulsion grounds the fact that the agent ought to drink has much greater philosophical interest.

²Alvarez (2017) provides the impetus for recognizing the independence of the explanatory social role, and my analysis tries to carry the ball into further elements of social practice from there.

7.2 Making reasons matter more

Reasons talk has become a major focus of the literature in metaethics.³ This reason for this (pardon me) focus on reasons and the resilience of reasons literature is established by the diversity of roles described throughout this dissertation. It is not just that reasons are either justificatory or motivational, or that they play social roles, but that they do each of these things and so provide broad interest to a range of philosophical domains.⁴ Reasons are applicable to a range of areas of philosophical interest; because of this wide use, reasons also serve as a point of common reference when connecting these areas of philosophical interest.

If one is interested in political questions of legitimacy, then practical reasons, the Kantian idea of public reason, and various justificatory conditions for authority are appropriate; this line is advanced in [Raz \(1999\)](#) and [O’neill \(1996\)](#). In section [5.4.2](#), I note the applicability of these considerations to matters of public health during the present pandemic; the questions of authoritative public reasons add a further normative layer on top of the social scientific and policy making enterprises of coordinating behavior.

If one is interested in the differentiation between mere inclination or brute motivation, then the imposition of the constraints, the justificatory role of reasons, and conditions of social appraisal are useful to the discussion. These issues are discussed in philosophical context by [Broome \(2013\)](#)

³This section takes its name from the seminal [Fodor \(1989\)](#), which unpacks the functionalist approach to mental properties and which was a major methodological influence in this dissertation.

⁴A version of this argument is developed independently in [Schroeder \(2021\)](#). While Schroeder advances a version closer to my own argument than anything else in the literature, I think this reasons first view is a salient feature of lots of the contemporary metaethics literature since the late 1970s, especially from Parfit, Scanlon, Raz, and O’Neill. Schroeder is making the reasons first view more explicit.

as well as in the various interdisciplinary literature around economics, social psychology, and ethics. The portions of this dissertation dealing with constraints, including sections 3.1 and 3.3.2, provide a space to interface with this extensive, interdisciplinary literature.

If one is interested in the communication and coordination of behavioral dispositions across communities, then the justificatory and motivational roles both come into play, and can be applied alongside discussions in philosophy of language and accounts of norms. There are a range of attempts to develop this, from the expressivist literature of Gibbard (1990) and Schroeder (2009) to the morality driven approaches of Copp (2007).

If one is looking to connect communication and coordination of those behavioral dispositions to the questions of moral and political application, as I raise in chapter 5, then reasons provide a way of relating the descriptive dimension of the former to the normative dimensions of the latter. A lot of contemporary philosophy involves building up connections between established domains, and reasons are a useful bridge between and across such domains.

One of my first introductions to serious 20th century philosophy was Sellars' "Philosophy and the Scientific Image of Man." Sellars' broad pragmatic assessment is useful to frame up these programs and the value of reasons as central theme and potential connective tissue.

The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term. Under 'things in the broadest possible sense' I include such radically different

items as not only ‘cabbages and kings’, but numbers and duties, possibilities and finger snaps, aesthetic experience and death. To achieve success in philosophy would be, to use a contemporary turn of phrase, to ‘know one’s way around’ with respect to all these things, not in that unreflective way in which the centipede of the story knew its way around before it faced the question, ‘how do I walk?’, but in that reflective way which means that no intellectual holds are barred. (Sellars, 1963, Sec. 1)

Reasons are a part of the way a lot of things (including normative and descriptive features) hang together.

Some modifications of the discussion of reasons, and even modification of my analysis, may be necessary depending on the domain of use. When one is talking about, for example, the sorts of exchanges between individuals that communicate and influence behavior, there is less of a role for the formal constraints on rationality; those constraints are not really something that most people communicating by giving reasons are actively considering, and so one need not get into the technical issues around acyclicity (for example) when we recognize that the canons of rationality may not be rigidly or explicitly applied in that context.

Consider an example of a non-reason like the case where David is told “Mom is cold, so you should wear a jacket.” Our intuitive discussion of why this is not a reason (despite having motivational force for David) does not need to require deep discussion of the technical details governing relevance. In a context where the constraints on justification are the focus, then one should include such discussion of relevance and factual defeaters, as I do

in section 3.3.4; in a context where one is focusing on communication and influence, those considerations can take a back seat, though they are still present.

There is still value in recognizing the constraints on the construction of reasons, especially in instances where individuals recognize that the constraints render some justifications illegitimate; when Smith says that he has stopped drinking dairy because Mercury is in retrograde, then one might reasonably challenge whether this counts as a reason given the irrelevance of star signs to dietary choices. The point is just that the fine-grained understanding of such constraints may not be necessary, even where the general recognition that there are such constraints in play is necessary and useful. Sometimes, a coarse-grained treatment is all that is required.

One function of this dissertation is a push back on the general view that the functional roles of reasons are separable and that talking about justification without motivation (or *vice versa*) is the optimal philosophical approach to the conceptual analysis of what it is to be a reason. The idea that reasons can be separated as justificatory and motivational (while the sets clearly overlap) is a mistake, since the explanatory value of reasons lies in significant part in the ways in which the justifying and motivational condition hang together and produce the various social roles.

This strikes a tricky balance. I argue for the unity of the roles of reasons while acknowledging that different projects with different goals may see fit not to emphasize or discuss different functional roles. There is a tension between the conceptual claim that to be a reason is to be at the intersection of the functional roles (on the one hand) and acknowledging that some projects may stray out of that intersection in order to discuss particular

features of reasons and some reason-adjacent-facts (on the other). This is not an impossible balance; it just requires some practical consideration of the projects' needs.

Different projects have different requirements, and the fluctuation of these requirements will result in different points of emphasis. As psychologically oriented projects like Davidson (1963) were focused less on the formal constraints of justification, it makes sense that such a project would focus primarily on motivation.

7.3 Glimpses Beyond

Reasons are of philosophical interest, and this dissertation is of value, because of the broad range of areas in which reasons play a significant role.⁵ This leaves a great deal of space for the ideas in this dissertation to run.

Given the broad range of applications of the reasons literature, there are a handful of ways in which one might proceed to develop the project. One might proceed by introducing technical refinements to constructivist theories, giving more details on the constraints developed in the previous section. One might push on potential limiting cases to each of the conjuncts, the reasons as evidence and promoting and protecting conditions. One might push for the broader social, moral, and political applications of this theory as the ideas develop in chapters 5 and 6. My own inclination is towards the latter, but this is driven by my own interests in the social, moral, and political domains, rather than decision theory or epistemology. As such, this section focuses on those future projects.

⁵This phrase is borrowed from the closing discussion of Quine (1982).

The first potential direction is individuating an account of the *domains* of reasons. It is common to talk about some reasons as practical and other reasons as moral; some philosophical discussions also include reference to prudential, epistemic, legal, and other domains of reasons. One of the useful things about the account that I provide is that it makes explicit a range of dials that might help to clarify what differentiates (for example) a practical reason from a moral reason. Is it the particular values that are at play? Is it that the reason takes as its objects certain kinds of considerations? There is already a considerable literature on the mark of the moral in discussion of reasons, and one upshot is that my theory helps to specify some of the potential areas where reasons may differ, such that there is differentiation in domains.

My hope is that developing an account of the domains of reasons would be straightforward, since the regimentation of assessing those different areas is already a part of this dissertation. (Of course, I may later find out that the regimentation in section [1.1.1](#) is not adequate to handle certain issues; this is always a risk.) As such, the theory may be useful in navigating some of the disputes in the existing literature.

There are two slightly more ambitious long term projects that I hope to pursue in continuing to develop this project, and tying in with other areas of my philosophical interest.

The second potential direction is to tie the existing theories of reasons with existing work in other disciplines, especially in the social science (taken broadly to include cognitive science, psychology, etc.). There have been some attempts to draw this connection; as I note above, some philosophers (e.g. Mark Schroeder) are interested in the relation between reasons and

communication as a means to understanding social coordination. This is the sort of ambitious, interdisciplinary project that I have in mind, albeit with a different angle.

The idea is something like this. Understanding contemporary issues in cognitive science and psychology involves consideration of intersubjective assessments of reasons. This gets tied up in the contemporary philosophy of mind literature, as there are still ongoing discussions of how mental content, attention, and other philosophical issues in mind work, but even without clear resolutions to these problems, one can make significant progress on the tie-ins. In chapter 5 I make some general gestures at Tomasello and other programs in cognitive psychology during the late '90s and early '00s that approach social cognition by focusing on the attention patterns of organisms, as well as the ability of organisms to anticipate behavior. In the case of Tomasello, this focused on developmental and comparative psychology (the comparative work largely limited to primates) as a way to assess potential relevant similarities in the assessment of social interaction.

The discussion of reasons as facts in section 1.2 focuses on a particular feature of this, that it is useful to understand when individuals are attending to the same object (joint attention) and when they recognize that the other agent is also attending to that object (shared intention). This is a part of the literature in developmental psychology; theorists like [Tomasello et al. \(1995\)](#); [Gopnik et al. \(1994\)](#), *inter alia* have attempted to evaluate how the various cognitive processes tied up in assessing the reasons of others work. This is a useful psychological program that is fundamentally about the intersubjective assessment of reasons for action. Chapter 5 develops the analysis of this program in broader, philosophical terms by focusing on how the

understanding of intersubjective assessment of reasons and fundamental literature in social ontology gives rise to social structures (which is an important philosophical explanandum), but I think the psychological project can also contribute a lot of value by helping to flesh out our understanding.

There is a sophisticated relationship between the facts and the attitudes of agents, and that these attitudes are formed in part by the psychological and neurological features of the agent. Accounting for many of the elements expounded in the analysis of valuing involves accounting for how the psychological and neurological processes work; when agents are disposed to feel certain things in response to facts (or perceived facts) this implicates a scientific story. Further, there's a scientific story of how we recognize these sorts of dispositions and feelings in others (or fail to recognize them, in some cases).

These provide the empirical backbone for fleshing out the motivational role; it also helps to circumscribe some of the potential neurological and psychological explanations for shared values (and atypical values held by eccentric agents). The ability to recognize the dispositions and feelings of others helps to work into the issues of behavioral coordination; the recognition of joint attention and collection intention also helps in developing these cases. The existing literature on neurology and psychology helps to bring together a pretty substantial range of philosophical issues.

Further, this approach allows for tying that more robust literature together with the social, intersubjective appraisal of reasons for actions and the ways in which this informs social activity. There is a range of interdisciplinary work that involves the discussion of how societies organize, assess, and coordinate the behavior of members; all of this is ultimately grounded

in out reasons talk across a few philosophical domains, including the reasons of the societies to impose certain norms and the reasons of individuals that may be assessed as justifying (or not) the agent acting in such a way. There is already some substantial and interesting work in philosophy on how discourse on practical reasons can inform this sort of social activity and coordination, and the hope is to use my theory to address that.

Earlier in the dissertation, I note the discussion of Raz's and O'Neill's work on the relationship between practical reasons, public reason, and theories of law and authority. I note above the role of contemporary projects in communication and social coordination by philosophers like Schroeder, which include application to linguistics. These discussions all hang together through the use of reasons, but may often be viewed as disparate because the former set of issues (as in [Raz \(1999\)](#); [O'Neill \(2015\)](#)) is on justification and authority while the latter set of issues (as in [Schroeder \(2021\)](#)) is on motivation and influence. My hope is that these can be brought together using the moves and ideas advanced in this dissertation.

This is an ambitious project, but one that has a lot of potential upside and broad philosophical and interdisciplinary interest.

The third potential direction concerns the political implications that develop throughout the end of chapter 6. This is similar to the second direction, in that it hopes to tie together some broad themes across disciplines (including some different subdisciplines of philosophy). As is somewhat clear in the discussion of Chapters 5 and 6, the view adopted by many constructivists is that constructivism plays at least an indirect role in motivating certain liberal democratic positions, which is why the work of O'Neill,

Rawls, Raz, and Walzer seem to contribute to a broad set of liberal political philosophies.

The idea is best explored through some of the thematic moves in Rawls' theory of justice, which includes a fairly explicit tension between implications. (I will return to that tension below.) Rawls does not cash this out in terms of constructivism, really, though there is a lot of constructivist language and (because we know from other writing that Rawls is a constructivist) the condition seems appropriate.

The point that Rawls makes in the early development of *A Theory of Justice* is that there are conditions of deliberation under which the values and preferences of individuals result in an inclination towards substantial political and moral claims, especially claims about justice. Individuals considering what is fair under certain conditions, namely some version of impartiality brought about by the veil of ignorance exercise, are inclined to hold a substantive requirement on an act being just that it is (for example) fair to those affected by the act, responsive to deserts, etc. These sorts of substantive considerations allow for the construction of some shared conception of justice (and other morally and politically salient conceptions) that can act as a starting point for collective discussion about the range of our society. The conception is constructed, because it is the product of deliberative exercise. Still, it is widely shared enough that it can act as a substantial constraint on the reasons a society has for acting.

For example, it is wrong for an individual to receive a benefit that is not in response to some justifying consideration of their position. An agent who is given an advantage because he is white (for example) is benefiting unjustly; an agent who is given an advantage because of hard work (by contrast) may

be viewed by some as having earned that benefit. Alternatively, there may be justificatory conditions which an agent deserves in virtue of some other consideration, like the additional accommodation given to compensate for existing disadvantages. Rawls' consideration in these cases is that the development and application of the notion of justice under such circumstances is a social and political construction, really the development of these norms is an act of construction by the community.

Rawls' approach to justice is constructivist; Rawls argues that our shared judgments and considerations through reflective equilibrium should result in adopting certain substantive restrictions on the appropriate conception of justice and on what we recognize as reasons both for individual action and government/collective action. Rawls' approach acknowledges that there is a range of reasonable disagreement even if one adopts the substantial constraints brought about by a shared conception of justice. Suppose we hold that (for example) pleasure drawn from wanton torture is excluded as a justification for action, and therefore that the ideally coherent Caligula cannot justifiably act in accordance with his values of suffering; this does not entail that there is a single, monomaniacal conception of justice which imposes strict constraints. There may be certain kinds of practical reasons which are acceptable on some conceptions of justice and not on other conceptions, while the shared conception does not exclude either.

Consider a substantial dispute over the limits of corporate responsibility for bad actors using their products. Suppose some company C makes a product that is widely used; many of its users benefit from the product without causing any harm and a small percentage of the users do harm with the product. There is no way for the company to change the product to

prevent harm without also reducing the benefit to some users. Two policy-makers might reasonably disagree about whether they should pass a policy that would compel the company to change the product. One might argue that the reduction of harm is paramount; the other might argue that the company should be allowed to exercise their discretion. This is a reasonable dispute, and familiar to those who engage in substantial discussions about complicated policy issues, though some views may be less reasonable.

The conceptual constraints on what actions are justified is underdetermined in many cases. A more abstract way of framing this, consistent with this: for some circumstances C , there is an act type ϕ and a non-identical act type ψ such that it is permissible for an agent in C to ϕ and permissible for an agent in C to ψ . Under these circumstances, one wants to maximize the ability of agents to act according to their own preferences. Consider the case of menu options on a wedding invite; it is permissible for me to RSVP for fish and it is permissible for me to RSVP for chicken. Under these circumstances, I should have the ability to act in accordance with my preference for fish over chicken (or *vice versa*). By maximizing choice in these circumstances, one generally maximizes the outcomes for individuals; this assumes that individuals are (at least generally) the most informed about what act produces the best outcome for them. There may be some dispute about the limits of this generalization, as there are circumstances where individual agents are perhaps not in the best position to assess the outcomes, but it functions well enough as a general principle. This is broadly an argument for the implementation of autonomy; the standard view advanced by Rawls and other liberal philosophers is that autonomy should be permissible except in cases where there is strong countervailing interests. Sometimes they use

the term “liberty” to talk about this, or to describe the social feature rather than the agent-level autonomy, but I have found that “liberty” has a range of vexed conceptions and so try to avoid that language when I can.

The balance in these contexts is between the constraints on reasons imposed by the substantial arguments for conceptions of justice (and related salient conceptions) and by other socially constructed constraints on justification (on the one hand) and the general inclination to grant latitude to agents in their construction of practical reasons on the basis that this maximizes their autonomy and outcomes (on the other hand). This is an ongoing debate, and the development of such constraints is a serious project in political philosophy, as is trying to work through a framework for the application of the ideas in political philosophy to contemporary political issues.

The prospects for this dissertation’s use are extensive. They range from applications in decision theory and abstract approaches to reasons to the applied domains of the social sciences, ethics, and political philosophy. The purpose of this section is to highlight this range, as well as provide some insights into how such connections may develop, and how reasons may be used to develop bridges across such projects.

7.4 A note on value

Reasons are core constituents of a range of philosophical subdisciplines, from the philosophy of the social sciences to moral and political philosophy to abstract projects on the nature of rationality. As such, the hope is that this dissertation provides a clear and practically useful guide to those who want

to make use of my theory of reasons; further, the hope is that developing a theory of reasons deliberately oriented to the diversity of roles that reasons play in our lives provides greater value to those looking for a plausible and tractable theory.

A major argument in this dissertation, really *the* major argument for the theory of reasons I advance through this dissertation, is that a theory of reasons is at its best when that theory can be taken up and used in other domains, including the social and political. There should not be such a pronounced distinction between the ostensibly abstract areas of philosophy and the range of potential applications. Spelling out the applications of abstract philosophical work is methodologically valuable for those who are coming to philosophy from the outside and may struggle to see the potential value and use without some help and encouragement. It is also politically essential in a culture that is wrestling with the role of academia and abstract generally, and philosophical work in particular.

Bibliography

Alvarez, Maria. 2017. Reasons for action: Justification, motivation, explanation. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2017 ed.

Andreou, Chrisoula. 2020. Dynamic Choice. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2020 ed.

Anscombe, Gertrude Elizabeth Margaret. 1957. *Intention*. Cambridge, MA: Harvard University Press.

Austin, John Langshaw. 1975. *How to do things with words*. Oxford: Oxford University Press.

Ayer, Alfred Jules. 1936. *Language, Truth, and Logic*. London: Gollancz.

Baghramian, Maria and J. Adam Carter. 2020. Relativism. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, fall 2020 ed.

Bagnoli, Carla. 2017. Constructivism in metaethics. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2017 ed.

- Bell, John L. 2021. The Axiom of Choice. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2021 ed.
- Benatar, David. 2012. Every conceivable harm: A further defence of anti-natalism. *South African journal of philosophy* 31(1): 128–164.
- Benatar, David. 2015. The misanthropic argument for anti-natalism. *Permissible progeny: The morality of procreation and parenting* 34–64.
- Bentham, Jeremy. 1789. *An Introduction to the Principles of Morals and Legislation*.
- Boghossian, Paul. 2007. *Fear of knowledge: Against relativism and constructivism*. Oxford: Clarendon Press.
- Boudon, Raymond. 1998. Limitations of rational choice theory. *American Journal of sociology* 104(3): 817–828.
- Boudon, Raymond. 2003. Beyond rational choice theory. *Annual review of sociology* 29(1): 1–21.
- Brandt, Richard. 1967. Some merits of one form of rule-consequentialism. *University of Colorado Studies in Philosophy* 3: 39–65.
- Briggs, R. A. 2019. Normative Theories of Rational Choice: Expected Utility. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, fall 2019 ed.
- Broome, John. 2013. *Rationality through reasoning*. New York: John Wiley & Sons.

- Brower, Bruce W. 1993. Dispositional ethical realism. *Ethics* 103(2): 221–249.
- Brunero, John. 2009. Reasons and evidence one ought. *Ethics* 119(3): 538–545.
- Capps, John. 2019. The Pragmatic Theory of Truth. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2019 ed.
- Chang, Ruth. 2016. Comparativism: The grounds of rational choice. *Weighing reasons* 213–40.
- Coleman, James S and Thomas J Fararo. 1992. Rational choice theory. *New York: Sage*.
- Copp, David. 2007. *Morality in a natural world: selected essays in metaethics*. Cambridge: Cambridge University Press.
- Dancy, Jonathan. 2000. *Practical reality*. Oxford: Clarendon Press.
- Dancy, Jonathan. 2017. Moral Particularism. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2017 ed.
- Davidson, Donald. 1963. Actions, reasons, and causes. *The journal of philosophy* 60(23): 685–700.
- Enoch, David. 2015. Against public reason. *Oxford studies in political philosophy* 1(20): 112–142.

- Epstein, Brian. 2018. Social ontology. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2018 ed.
- Fichte, Johann Gottlieb. 1889. The science of rights, translated by ae kroeger. *Philadelphia: JB Lippencott* .
- Finlay, Stephen and Mark Schroeder. 2017. Reasons for Action: Internal vs. External. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, fall 2017 ed.
- Fiske, Susan T. 1993. Social cognition and social perception. *Annual review of psychology* 44(1): 155–194.
- Fiske, Susan T and Shelley E Taylor. 2013. *Social cognition: From brains to culture*. Thousand Oaks, CA: Sage.
- Fodor, Jerry A. 1989. Making mind matter more. *Philosophical Topics* 17(1): 59–79.
- Gauthier, David. 1963. *Practical Reasoning: The Structure and Foundations of Prudential and Moral Arguments and Their Exemplification in Discourse*. Oxford: Clarendon Press.
- Gibbard, Allan. 1990. *Wise choices, apt feelings: A theory of normative judgment*. Cambridge, MA: Harvard University Press.
- Gibbard, Allan. 1999. Morality as consistency in living: Korsgaard's kantian lectures.
- Gilbert, Margaret. 1990. Rationality, coordination, and convention. *Synthese* 84: 1–21.

- Gilbert, Margaret. 2002. Collective Guilt and Collective Guilt Feelings. *Journal of Ethics* 6: 115–143.
- Gopnik, Alison, Virginia Slaughter, and Andrew Meltzoff. 1994. Changing your views: How understanding visual perception can lead to a new theory of the mind. *Children's early understanding of mind: Origins and development* 157–181.
- Gowans, Chris. 2019. Moral Relativism. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2019 ed.
- Gregory, Alex. 2016. Normative reasons as good bases. *Philosophical Studies* 173(9): 2291–2310.
- Guala, Francesco and Frank Hindriks. 2015. A unified social ontology. *The Philosophical Quarterly* 65(259): 177–201.
- Hacking, Ian, Jan Hacking, et al. 1999. *The social construction of what?* Cambridge, MA: Harvard university press.
- Hansson, Sven Ove and Till Grüne-Yanoff. 2018. Preferences. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2018 ed.
- Hare, RM. 1978. Relevance. In *Values and morals*, 73–90. Springer.
- Hindriks, Frank and Francesco Guala. 2019. The functions of institutions: etiology and teleology. *Synthese* 1–17.

- Horsten, Leon. 2019. Philosophy of Mathematics. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, spring 2019 ed.
- James, William. 1895. The meaning of truth. *Psychological Review* 2(2): 105–124.
- Jowett, Benjamin et al. 1888. *The republic of Plato*. Macmillan.
- Joyce, Richard. 2009. Expressivism, motivation internalism, and hume. In *Hume on motivation and virtue*, 30–56. Springer.
- Joyce, Richard. 2016. Moral anti-realism. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2016 ed.
- Kearns, Stephen and Daniel Star. 2008. Reasons: Explanations or Evidence? *Ethics* 119(1): 31–56.
- Kearns, Stephen and Daniel Star. 2009. Reasons as evidence. In *Oxford Studies in Metaethics*, ed. Russ Schaeffer-Landau, vol. 4, 215–242.
- Kearns, Stephen and Daniel Star. 2013. Weighing reasons. *Journal of Moral Philosophy* 10(1): 70–86.
- Kelly, Thomas. 2014. Can evidence be permissive? *M. Steup, John Turri, and E. Sosa, Contemporary Debates in Epistemology, 2nd Edition*. Oxford: Wiley-Blackwell.
- Kelly, Thomas. 2016. Evidence. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2016 ed.

- Korsgaard, Christine. 2003. *The dependence of values on humanity*, 63–85. Oxford University Press.
- Korsgaard, Christine M. 1983. Two distinctions in goodness. *The philosophical review* 92(2): 169–195.
- Korsgaard, Christine M and Christine Marion Korsgaard. 1996. *Creating the kingdom of ends*. Cambridge: Cambridge University Press.
- Korsgaard, Christine M, Christine Marion Korsgaard, Christine Marion Korsgaard, Gerald Allan Cohen, Raymond Geuss, Thomas Nagel, and Bernard Williams. 1996. *The sources of normativity*. Cambridge University Press.
- Liao, S Matthew. 2016. *Moral brains: The neuroscience of morality*. Oxford: Oxford University Press.
- List, Christian. 2013. Social Choice Theory. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, winter 2013 ed.
- MacBride, Fraser. 2020. Truthmakers. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, spring 2020 ed.
- Mackie, John. 1977. *Ethics: Inventing right and wrong*. London: Penguin UK.
- Manne, Kate. 2014. Internalism about reasons: sad but true? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 167(1): 89–117.

- Manne, Kate. 2017. Locating morality. *Oxford Studies in Metaethics* 12 12: 1.
- Mill, John Stuart. 1863. *Utilitarianism*.
- Moore, Chris, Philip J Dunham, and Phil Dunham. 2014. *Joint attention: Its origins and role in development*. Psychology Press.
- Mundy, Peter and William Jarrold. 2010. Infant joint attention, neural networks and social cognition. *Neural Networks* 23(8-9): 985–997.
- Mundy, Peter C. 2016. *Autism and joint attention: Development, neuroscience, and clinical fundamentals*. New York: Guilford Publications.
- Nagel, Thomas. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.
- Nagel, Thomas. 1972. War and massacre. *Philosophy & Public Affairs* 123–144.
- O’neill, Onora. 1986. I. the public use of reason. *Political Theory* 14(4): 523–551.
- O’neill, Onora. 1996. *Towards justice and virtue: A constructive account of practical reasoning*. Cambridge: Cambridge University Press.
- O’Neill, Onora. 2015. *Constructing authorities*. Cambridge University Press.
- Orend, Brian. 2001. Just and lawful conduct in war: reflections on michael walzer. *Law and Philosophy* 1–30.
- Parfit, Derek. 1984. *Reasons and persons*. Oxford: Oxford University Press.

- Parfit, Derek. 2011. *On What Matters*, vol. 1. Oxford: Oxford University Press.
- Pickering, Andrew. 1999. *Constructing quarks: A sociological history of particle physics*. Chicago: University of Chicago Press.
- Popper, Karl R. 1957. *The open society and its enemies: the spell of Plato*. New York: Routledge & Kegan Paul.
- Prinz, Jesse. 2011. Is attention necessary and sufficient for consciousness. *Attention: Philosophical and psychological essays* 174–203.
- Prinz, Jesse. 2012. *The conscious brain*. Oxford: Oxford University Press.
- Quine, Willard Van Orman. 1982. *Methods of logic*. Cambridge, MA: Harvard University Press.
- Quong, Jonathan. 2018. Public Reason. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, spring 2018 ed.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls, John. 1974. Some reasons for the maximin criterion. *The American Economic Review* 64(2): 141–146.
- Rawls, John. 1980. Kantian constructivism in moral theory. *The journal of philosophy* 77(9): 515–572.
- Rawls, John. 1989. The domain of the political and overlapping consensus. *NYUL Rev.* 64: 233.

- Rawls, John. 1993. The law of peoples. *Critical Inquiry* 20(1): 36–68.
- Rawls, John. 1997. The idea of public reason revisited. *The University of Chicago Law Review* 64(3): 765–807.
- Rawls, John. 1999. *The law of peoples: with, the idea of public reason revisited*. Cambridge, MA: Harvard University Press.
- Raz, Joseph. 1985. Authority and justification. *Philosophy & Public Affairs* 3–29.
- Raz, Joseph. 1990. *Authority*, vol. 1. New York: NYU Press.
- Raz, Joseph. 1999. *Practical reason and norms*. Oxford: Oxford University Press.
- Raz, Joseph. 2009. *Between authority and interpretation: On the theory of law and practical reason*. Oxford: Oxford University Press.
- Scanlon, Timothy. 2014. *Being realistic about reasons*. Oxford: Oxford University Press.
- Schafer, Glenn. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.
- Schafer, Karl. 2014. Constructivism and three forms of perspective-dependence in metaethics 1. *Philosophy and Phenomenological Research* 89(1): 68–101.
- Scheffler, Samuel. 2010. *Equality and tradition: questions of value in moral and political theory*. Oxford: Oxford University Press.
- Schmidtz, David. 1994. Choosing ends. *Ethics* 104(2): 226–251.

- Schroeder, Mark. 2009. Hybrid expressivism: Virtues and vices. *Ethics* 119(2): 257–309.
- Schroeder, Mark. 2013. Tempered expressivism. *Oxford Studies in Metaethics, Volume 8* 283.
- Schroeder, Mark. 2021. *Reasons First*. Oxford: Oxford University Press.
- Schroeder, Tim. 2017. Desire. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2017 ed.
- Searle, John R. 1995. *The Construction of Social Reality*. New York: Simon and Schuster.
- Searle, John R. 2006. Social Ontology: Some Basic Principles. *Papers* 80: 51–71.
- Sellars, Wilfrid. 1963. Philosophy and the scientific image of man. *Science, perception and reality* 2: 35–78.
- Shafer-Landau, Russ. 2003. *Moral realism: A defence*. Oxford University Press on Demand.
- Slote, Michael. 2015. Saucers of mud: Why sympathy and empathy require altruism. *Etica & Politica* 12–26.
- Smith, Michael, David Lewis, and Mark Johnston. 1989. Dispositional theories of value. In *Aristotelian Society Supplementary Volume*, vol. 63, 89–174. Oxford University Press.
- Sobel, David. 2009. Subjectivism and idealization. *Ethics* 119(2): 336–352.

- Sobel, David. 2011. Parfit's case against subjectivism. *Oxford Studies in Metaethics* 6: 52–78.
- Stein, Joshua. 2019. Exchanging for reasons, right and wrong. *The Journal of Value Inquiry* 53(2): 213–223.
- Stevenson, Charles Leslie. 1944. *Ethics and Language*. New Haven, CT: Yale University Press.
- Street, Sharon. 2008. Constructivism about Reasons. In *Oxford Studies in Metaethics, vol. 3*, ed. Russ Schafer-Landau, chap. 8, 207–246. Oxford: Oxford University Press.
- Street, Sharon. 2009. In defense of future tuesday indifference: Ideally coherent eccentrics and the contingency of what matters. *Philosophical Issues* 19: 273–298.
- Street, Sharon. 2017. *Nothing “Really” Matters, but that’s not what matters*, chap. Nothing “Really” Matters, but that’s not what matters, 121–148. Oxford: Oxford University Press.
- Strevens, Michael. 2008. *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael and Malinda Carpenter. 2007. Shared intentionality. *Developmental science* 10(1): 121–125.
- Tomasello, Michael et al. 1995. Joint attention as social cognition. *Joint attention: Its origins and role in development* 103130.

- Traub, Stefan, Christian Seidl, Ulrich Schmidt, and Maria Vittoria Levati. 2005. Friedman, Rawls, Boulding—or somebody else? an experimental investigation of distributive justice. *Social Choice and Welfare* 24(2): 283–309.
- Vooyo, Sarah and David G Dick. 2019. Money and mental contents. *Synthese* 1–16.
- Väyrynen, Pekka. 2019. Thick Ethical Concepts. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, summer 2019 ed.
- Waldron, Jeremy. 1986. John Rawls and the social minimum. *Journal of Applied Philosophy* 3(1): 21–33.
- Wallace, R Jay. 2015. Blackburn on reasons. *Passions and Projections: Themes from the Philosophy of Simon Blackburn* 246.
- Walzer, Michael. 2008. *Spheres of justice: A defense of pluralism and equality*. New York: Basic books.
- Walzer, Michael. 2015. *Just and unjust wars: A moral argument with historical illustrations*. Basic books.
- White, Roger. 2014. Evidence cannot be permissive. *Contemporary debates in epistemology* 2: 312–23.
- Whiting, Daniel. 2018. Right in some respects: reasons as evidence. *Philosophical studies* 175(9): 2191–2208.
- Williams, Bernard. 1979. Internal and External Reasons. In *Rational Action*, ed. Ross Harrison, 101–113. Cambridge: Cambridge University Press.

Williams, Bernard. 1995. *Making sense of humanity: and other philosophical papers 1982-1993*. Cambridge: Cambridge University Press.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

Appendix

Survey of potential variations and responses to ideally coherent eccentrics

There are a handful of potential variations on the ideally coherent eccentrics cases in the literature, some of which are not labeled as such or which are articulated in the context of other arguments. While my dissertation includes a discussion of the most interesting and valuable cases, it is useful to illustrate the broader range of these cases and provide some response to those cases. In this appendix I explain the available responses on my theory, as well as a short justification for why they are not included in the discussion in the dissertation.

As I note in chapter 6, the dissertation takes a pluralist approach to addressing these problems. In particular, the handling of ideally coherent eccentrics focuses on a combination of potential tools developed in the dissertation: (a) constraints imposed by the reason as evidence condition, (b) constraints imposed by the elements of valuing, and (c) other practical considerations providing countervailing reasons. The discussion in section 6.3 focuses on the ideally coherent Caligula, as I regard it as the most interesting and well-motivated of the ideally coherent eccentrics.

Here is a short list of the relevant cases:

Empty Track Trolley Problems See section 6.2.1.

Rawls' grass counter There is a coherent agent (Rawls suggests a brilliant mathematician) whose project in life is counting the blades of grass on his lawn. He holds this value as a matter of his attitudes, rather than of any deficient beliefs. (Rawls, 1971, p. 171)

ideally coherent anorexic There is an agent who is ideally coherent, but values an emaciated appearance even above other things like personal health. (Gibbard, 1990, p. 171)

Parfit's future Tuesday indifference There is an agent who is coherent but does not value his pain on future Tuesdays, and so is inclined to plan in ways that result in greater pain just because the pain will occur on a future Tuesday. (Parfit, 1984, p. 124)

Anscombe's saucers of mud Some agent(s) value the existence of saucers of mud, as an unmotivated and basic desire. Discussed in Anscombe (1957).

ideally coherent Caligula See section 6.3.

ideally coherent David Benatar There is an agent who believes that lives are actually of disvalue (rather than of value) and so, in addition to holding an anti-natalist view, holds inverted judgments in a range of cases. Based on Benatar (2015).

Even this list, supplemented by my own analysis, is not comprehensive, but the core idea is to isolate a broader range than is discussed in the actual chapter. One might see this as developing a range of potential theses that reject constructivism (either about values or about reasons) in a variety of ways.

Before proceeding into the cases that are unique to this appendix, it is worth getting the two items discussed in the dissertation out of the way.

Suppose that an individual is confronted with a situation like the standard switch version of the trolley problem, with the subtle variation that one

might redirect the train to a safe stop without causing any harm. Is it permissible to do anything *other* than direct the train to a safe stop? Section 6.2.1 considers the possibility that there are direct normative entailments to the presentation of certain cases. It follows, for example, from the presentation of the empty track version of the trolley problem that an agent A ought to ϕ , irrespective of the particular values of the agent. This is developed throughout the dissertation, as the empty track account provides a useful illustration of potential normative entailments from attitude-independent facts.

Suppose there is a person whose idiosyncratic attitudes result in practical reasons to do severe harm to others, that this is the way in which the person achieves personal pleasure and satisfaction. Section 6.3 considers the possibility that there are substantive, attitude-independent constraints on what can be of value. On this view, there are restrictions on the way in which reasons can be constructed, such that for some potential things x that an agent may value, even x valuing either cannot make x of value or x cannot bring about a reason. This is the major application of the ideally coherent Caligula cases.

There are some other approaches to this question that I do not address, partly because I do not regard them as very interesting or do not find them to be compelling as objections to constructivism. The purpose of this survey is to provide some survey and to note their lack of inclusion with some short address.

Parfit's agony argument and Anscombe's saucers of mud pivot on an argument from incredulity. Arguments from incredulity pivot on the imaginative nature of the case. Michael Slote writes, "The simple desire for a

Case	Puzzle	Response
Empty Track	section 4.4.1	section 6.2.1
Ideally Coherent Caligula	(Gibbard, 1990, p. 171)	section 6.3
Agony argument	(Parfit, 1984, p. 124)	Incredulity (below)
Saucers of mud	Anscombe (1957)	Incredulity (below)
Grass Counter	(Rawls, 1971, p. 171)	First-order norms (below)
Anti-natalism	Benatar (2015)	First-order norms (below)

Table 1: Quick reference Cases, puzzles, and responses

saucer of mud is not something we can readily understand, and I say the same about the simple desire to hurt people, sheer malice. And if one can always wonder whether the neurologically damaged psychopath might not have a basic and unaccountable for hatred of others, one can wonder too whether such a psychopath or someone else with a brain abnormality might not just have an unaccountable-for desire for mud.”(Slote, 2015, p. 18) But the nature of the contrast class that Slote sets up illustrates how the move works. Anscombe’s unmotivated desires for saucers of mud are hard to understand because they are *weird*; the fact that Slote then spends much of the paper arguing that we are wrong in the immediate response that in fact we do understand (at least at some level) the malice of violent, misanthropic sociopaths. While Anscombe’s case of saucers of mud are weird, Slote’s case of the malignant, violent sociopath is distressingly familiar.

The arguments from incredulity presented in the saucers of mud case indicate that because we cannot make sense of our desire for p, that renders the desire illegitimate. On this point, Nagel (1970) provides the general response, that sometimes we just have unmotivated desires. In some cases, we can understand the etiology or causal story behind those desires, but an individual who likes pop music may have no notion about the pentatonic

structure of familiarity of the I-IV-V chord progression or anything like this. The reason that Anscombe's saucers of mud case works (and Stole's misanthropic, violent sociopath case does not, in spite of the structural similarity) is because of the unfamiliarity, the sheer weirdness, of the example. [Street \(2009\)](#) notes that the mere strangeness of these cases is partly a result of the lack of detail, the lack of explanation of the facts in virtue of which the agent arrives at the idiosyncratic attitude.

If we had a subset of the population who had an unmotivated desire for a saucer of mud, then we might try to develop a causal story of that desire emerges (and Stole and Anscombe both frame it in the coarse language of mental illness) but this would render the case more familiar. These arguments from incredulity are not merely unsound; they are unhelpful and even actively obfuscating. One can still respond in any of the range of ways developed by constructivists: one can accept divergence in values as legitimate, as in [section 6.2.2](#); one can appeal to constraints on reasonableness, as in [sections 3.3.2](#) and [6.3.4](#); one can appeal to potential errors in factual conditions, as in [section 3.3.4](#). I discuss all of these throughout the dissertation. Pick one and a short, swift swing at Anscombe's example demolishes the argument.

Rawls provides a somewhat different framing regarding the grass counter case. While the grass counter is an eccentric, in the conventional formulation of the ideally coherent eccentrics cases, there are significant worries about non-contribution to society. Suppose that the grass counter is a skilled mathematician or engineer, counting grass rather than contributing to the community. While Rawls is (I think rightly) quite open about the sorts of aims an individual can pursue while still contributing to society, the

purpose of the grass counter is to illustrate a case where there is minimal contribution to the community.

The challenge for Rawls has to do with making sense of how a social arrangement negotiated from the original position handles someone who is an eccentric and whose eccentricity results in freeloading. In the course of developing the issues with free rider cases in discussing counterexamples, in both sections 6.3 and 6.4, I note that this is principally a compliance problem. While Rawls' grass counter does add a layer by putting the ideally coherent eccentric in the situation of the free loader problem, this is really not so interesting a version. The prospect of actively harmful free riders is more significant than the grass counter case, where we are merely talking about non-participation in productive work, but the solution is ultimately the same. The solution is to note the role of the social elements and the ways in which these address (albeit imperfectly) compliance problems.

In presenting the Empty Track version of the trolley problem, I try to advance what I regard as the most straightforward, small-c-conservative version of the claim that certain descriptions of states entail normative judgments. Ultimately, I think this argument requires falling back to a pre-existing assumption about the modal status of values implicated in the normative judgment, like the positive values of human life. In that context, I briefly allude to Benatar's anti-natalism, which holds that in fact lives do not have a positive value, but rather a negative one. Benatar's argument largely proceeds against a metaethical backdrop of realism about values and a commitment to the real negative value of lives; as such, I think Benatar's argument is unconvincing on the basis of the same assumptions constructivists reject, as outlined in section 3.4.2.

The reason that I intimate about Benatar is as a way of illustrating what a conceptual argument against the value of life may look like in a first-order context. I do not think the argument is successful, because it relies on a realist schema that I think is false. But if one takes the constructivist approach of regarding such conceptual arguments as an exercise in the construction of values, then Benatar's account is illustrative. Benatar's argument is an illustration of how an ideally coherent eccentric might look if that eccentric did not hold the relevant valuing attitudes towards life.

Benatar is at least practically coherent; Benatar is also missing the doxastic commitments to the valuing of life, though it turns out that he probably isn't missing all of the valuing conditions. The Benatar case is interesting in that it allows for first-order discussion of values and the constraints governing how we talk about idiosyncratic values. I avoid this because it takes us into a first-order domain.

The major concern in talking about these issues is the divergence in values, which is an anthropological fact across and within communities. It is [Boghossian \(2007\)](#) who really gets to the core of this issue in his discussing worries about constructivism and the outer limits.

There are two distinct approaches available to constructivists, and the two distinct approaches are usually characteristic of the break in "Humean" and "Kantian" constructivisms. "Humean" constructivists tend to appeal to the doctrine of anthropological non-normativity, that it is an incoherent exercise to impose external norms on internal communities. By contrast, "Kantian" constructivists tend to focus on one of two approaches: the appeal to universal constraints or the appeal to public reason (where "public" means universally accessible).

In my view, and hopefully as illustrated through section 3.3.2 and chapter 6, is that these approaches are too restrictive when treated independently, and that the better approach is to apply a diverse range of potential tools available to constructivist theories in combination in order to provide thorough, appropriate answers to these potential worries.