

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Reasons as Explanations: A Genuine Alternative to the Guise of the Good Thesis

Permalink

<https://escholarship.org/uc/item/2533v5mv>

Author

McFadden, Meredith

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Reasons as Explanations:
A Genuine Alternative to the Guise of the Good Thesis

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Philosophy

by

Meredith Leigh McFadden

September 2017

Dissertation Committee:

Professor Michael Nelson, Co-Chairperson
Professor John Martin Fischer, Co-Chairperson
Professor Andrews Reath
Professor Coleen MacNamara

Copyright by
Meredith Leigh McFadden
2017

The Dissertation of Meredith Leigh McFadden is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

I'd like to sincerely thank the small army of people that have supported me during the process of writing this dissertation.

I'd of course like to thank my committee: Michael Nelson and Andy Reath for taking part in such rich and helpful reading groups without which I'd be nowhere, John Fischer for the personal investment he has shown in me and the agency crew, and Coleen for going repeatedly above and beyond in her mentorship. Thank you also to Maude for all the cocktails and empowerment, I can't believe my luck in having you as a role model.

The completion of this thing is largely thanks to my cohort's willingness to engage in nonsense, Duke's, pun competitions, dance parties, and a seriously shocking amount of reading groups and workshops. Truly the best cohort to have ever cohorted.

I've got to acknowledge my honorary cohort as well. Thanks, Zac Bachman for sharing your power of always speaking true things. Thank you, Dave Beglin, for what I think in fact are endless conversations, especially those that have drowned out whatever we happened to be watching at the time. Thanks to superhero and Viking warrior Monique Wonderly for being kind and brilliant and putting pressure on the limits of superlatives. Carlo Palombo and Amy Wilson contributed to the quality of mind, life, and thought by existing and by sharing their time and paints, and by promoting lumberjacking and rainbow fires.

I am very grateful to Parker Williams for understanding the importance of small excitements, all things festive, and the curative powers of grocery shopping.

Thank you so much to rock climbers and gamers I met this past year for pushing me when I thought I had run out of volition and reminding me that people are awesome. I am greatly indebted to UCR's dance department, in particular Kelli King, for letting my clumsy self into their fold during the past six years. Your support and community have meant more to me than I can say.

I'd finally like to thank the people I haven't explicitly mentioned here but nevertheless have been of incredible support and help throughout the nine years I've been in graduate school. Staying sane and grounded and productive all at once is no small thing, and above everything it's people that have meant the most throughout this process. If you know who you are, you know how dismayed I am to have not noted you sufficiently here. Family, friends, wonderful people: THANK YOU.

This dissertation is dedicated to my Adam.

ABSTRACT OF THE DISSERTATION

Reasons as Explanations:
A Genuine Alternative to the Guise of the Good Thesis

by

Meredith Leigh McFadden

Doctor of Philosophy, Graduate Program in Philosophy
University of California, Riverside, September 2017
Michael Nelson, Co-Chairperson
John Martin Fischer, Co-Chairperson

When we see a consideration as a reason for action we see said consideration as *supporting*, or *speaking in favor of*, performing the action. It has been difficult to make progress in unpacking the “supporting”, or “speaking in favor of” relation. According to the standard account, the “guise of the good thesis”, we see a consideration as supporting an action insofar as we see it as pointing to the *good* in the action.

In my dissertation, I argue that this is misguided; it leaves us with a distorted and overly virtuous picture of human agency. My overall project is to provide a viable alternative to this long-held view. I put forward and defend an *explanation-based* view of practical reason.

Though some views of practical reasoning have pointed to explanation to illuminate the way reasons support performing actions, the accounts have been woefully underdeveloped. In my dissertation I mine the philosophy of science literature on

explanation in order to apply a pragmatic theory of explanation to the practical domain. With the theory of explanation in hand, I can provide standards for good and bad reasons for action. I show how my view thereby accounts for how reasons support performing actions in both our deliberation and justification practices, and rivals the guise of the good model of practical reasoning without relying on evaluative content.

CONTENTS

INTRODUCTION	1
<u>CHAPTER ONE: THE STANDARD VIEW</u>	<u>15</u>
Section One: The Standard View	17
Section Two: Routes to the Standard View	20
Section Three: Virtues of the Standard View.....	31
<u>CHAPTER TWO: WHY REJECT THE STANDARD VIEW?.....</u>	<u>35</u>
Section One: Undermining the Intuitive Grounds for the Standard View	36
Section Two: Cases Against the Standard View	43
Section Three: What Could an Alternative Do Better.....	50
<u>CHAPTER THREE: TOWARDS A GENUINE ALTERNATIVE</u>	<u>59</u>
Section One: The Explanation-Based View.....	59
Section Two: How Explanation Relates to Justification	65
Section Three: Problems with Available Versions	71
Section Four: A Way Forward	76
<u>CHAPTER FOUR: EXPLANATION</u>	<u>79</u>
Section One: Standards for Accounts of Explanation	80
Section Two: Accounts of Explanation.....	84
Section Three: Van Frassen’s Account of Explanation	89
<u>CHAPTER FIVE: REASONS AS EXPLANATIONS.....</u>	<u>108</u>
Section One: Recall van Frassen	109
Section Two: Explaining Action	113
Section Three: Human Action.....	122
Section Four: Using Reasons to Decide What to Do	126
Section Five: Bad Reasons.....	135
<u>CHAPTER 6: A GENUINE ALTERNATIVE TO THE STANDARD VIEW</u>	<u>141</u>
Section One: Virtues of the Standard View	141
Section Two: Meeting the Desiderata	150
Section Three: Virtues of Reasons as Explanations.....	158
BIBLIOGRAPHY.....	163

Introduction

“I solemnly swear that I am up to no good”

-Incantation, J.K. Rowling,
Harry Potter and the Prisoner of Azkaban

“Why is he doing a handstand in the middle of a busy intersection?” “Hey, why did you hit me in the face?” “Should I help this guy or keep on my way?” We deal with questions like these pretty frequently. If we are puzzled by someone’s behavior, we ask her why she acted that way - inquiring into her reasons for so acting. If we are unsure of what to do, we consider the reasons for and against the action, weighing them in order to settle our minds and come to a decision. We are, moreover, typically ready to provide our reasons for our own actions. So, we can answer these questions: Why’d I hit you in the face? There was a bug. Why is that person doing something so unusual? He drank too much and thought it would be hilarious. Should I help or keep on my way? You should help, it would benefit her and no one else can. In each of these cases we are taking considerations to be reasons for performing an action.

When we take a consideration to be a reason for action we take said consideration to be supporting, or speaking in favor of, performing the action. When, for example, I take that coffee has caffeine to be a *reason* to drink the coffee, I am taking that it has caffeine as *speaking in favor* of my drinking the coffee. Or again when I take that the parking spot is close to my office to be a *reason* to park there, I am taking the consideration that it is

close to my office to be *supporting* my parking there. Because when an agent takes a consideration to be a reason, she take it to be *supporting* performing an action, it can guide her behavior. To this much all agree.

Trouble arises, though, when we try to unpack the “supporting” or “speaking in favor” relation. What does it mean to say that an agent takes a consideration to “support” or “speak in favor” of an action?

The standard answer to this question is the “guise of the good thesis.” This view, named by Aquinas and dominant since Socrates, unpacks the supporting relation in terms of what the agent takes to be *good* in her action.¹ More specifically, according to the guise of the good thesis, we take a consideration to support performing an action only insofar as we take it to point to the good in the action. For example, we take the consideration that the coffee has caffeine to “speak in favor of” drinking the coffee insofar as it points to the good in drinking the coffee. Or again, we take the consideration that the parking spot is close to support parking there insofar it points to the good in parking in said spot.

At base, the guise of the good thesis is a view about the relationship between our reasoning capacity and our *evaluative capacity* – our capacity to assess things as good. The thesis claims that how we evaluate the world constrains what reasons we take ourselves to have. That is, practical reasoning is grounded in evaluative assessments. In this dissertation, I argue that this is misguided; it leaves us with a distorted picture of human agency.

¹ See, for example, Augustine’s discussion of his motives in famously stealing pears from his neighbor in Augustine (2008). For most views, the commitment to the guise of the good thesis is implicit. For contemporary examples of views that explicitly adopt the guise of the good thesis, see Raz (manuscript), Tenenbaum (2007), Hawkins (2008).

The guise of the good thesis, while intuitive in some paradigm cases, fails to do justice to some cases of complex, and deeply human, behavior. It is important to account for a broad range of human action, including actions that are not represented by the traditional paradigm cases. Evil and self-destructive agents are agents too, and we should be wary of characterizing their flaws as flaws of reasoning rather than deviations from evaluative norms.

The basic idea involved in taking a consideration to be a reason for action is taking a consideration to support performing an action, and an agent can do that while being perverse, conflicted, wicked, immoral, evil, and evaluatively twisted in a number of interesting ways. I suggest that this is so because our rational and evaluative capacities, though of course intimately connected, are different things. The guise of the good thesis does not appreciate this fact.

My overall project, therefore, is to provide a viable alternative to this long-held view that accounts for the spectrum of reasoning agency. I will put forward and defend an *explanation-based* view of practical reason. In this introduction, I will orient the topic and clarify the way I will speak throughout this dissertation before outlining the upcoming chapters.

Orientation

We use considerations to come to decisions all the time. That the light was red led to my decision to stop. That there was cream cheese frosting on the cake led to my decision to splurge and buy the piece. That the coffee shop was open late led to my decision to

frequent them over a competitor. The reason that taking considerations as reasons leads to decisions is because in taking them to be reasons, we take them to count in favor, or support, a particular action. So, understanding what is involved in taking a consideration as a reason comes down to understanding this “support” relationship. Considering the sorts of decisions we take to be paradigmatic and which we take to be problematic can provide us with some good intuitive grounds for what is involved in taking a consideration to be a reason for action.

For example, we can make a variety of mistakes between taking a consideration to be a reason and making our decision to act. When I take the consideration that Eric is at a party to be a reason to go to the party, I could be mistaken in considering that Eric in fact is there, for he is in fact at home. This is part of what “take the *consideration*” language attempts to capture – I *consider* it to be the case that Eric is at the party, but I could be wrong about this.

There are other ways we could make mistakes in taking considerations to be reasons for action besides not capturing features of the world accurately. Whereas in examples like the above case, I mistook the empirical facts of the matter, I could make mistakes regarding how considerations support actions: I could make reasoning errors. These involve weighing reasons together to come to conclusions, so for instance I could take it that in order to avoid an itch on my pinky I should undergo open chest surgery. In such a case, I could take the consideration that it would allow me to avoid an itch to count in favor of undergoing the surgery, and compare it to the considerations regarding the risks and experiences involved in the surgery, resulting the in the decision to opt for the surgery. Something strange has

happened in this bit of reasoning: perhaps, say, a misapplication of the comparative strengths of the support relations of the considerations.

Further, it seems like there are cases that may appear to be of this second class of error (mistakes in reasoning) that are so extreme that they cast doubt on whether the mental transitions from taking something to be case in the world to coming to a decision qualify as taking a consideration to support at all. For instance, I could take the consideration that Sophocles is a Greek poet to support my buying a cup of coffee.² In typical cases, it is obviously clear that there isn't a rational connection between the Sophocles' status as a Greek poet and my purchasing coffee.³ Thus, it seem so unintelligible that an agent took the consideration that Sophocles was a Greek poet to support buying coffee, it seems the most likely explanation is that she is misspeaking, either confused about what reasons are or what her reason was. In short, it is unintelligible that an agent takes the consideration that Sophocles was a Greek poet to support buying coffee.

The possibility of reaching this degree of unintelligibility in taking a consideration as a reason for action marks the divide between the standard view of reasons for action and those that attempt to reject this view. According to the standard view, taking a consideration to be a reason to perform an action involves taking the action to be valuable (that is how it counts in favor). So, according to this view, to take a consideration to be a reason for action without taking the action to be valuable is unintelligible. Such a case

² This example comes from Raz (1999) p 8.

³ Though we could, perhaps, construct one: A local coffee shop could be having a trivia-dependent special where, for example, you could get 50% off a beverage of your choice if certain conditions are met. If, say, your nationality matches that of the trivia answer of the day. If I'm Greek and the coffee shop's trivia question today is, "Who wrote Oedipus Rex?", then the consideration that Sophocles is a Greek poet would speak in favor of my buying coffee because it gets me 50% off my coffee.

would be tantamount to taking the consideration that Sophocles was a Greek poet to count in favor of drinking coffee.

If we ask an agent why she performed an action and she tried to give a reason while denying that there was anything worthwhile in the action, according to the standard view of reasons this reaches a level of unintelligibility that suggests she wasn't taking there to be a reason at all. Indeed, for adherents of the standard view, it is conceptually impossible to take a consideration to be a reason for action while taking that consideration to have no bearing on the good of the action.

In rejecting the standard view of reasons, I claim that an agent *can* take a consideration to be a reason for performing an action without taking it to highlight the good in performing the action – this is an intelligible way of taking a consideration to count in favor of acting. When we ask an agent why she performed a behavior, she can respond with a consideration that she didn't take to bear on the value in performing the action without raising theoretical concern. The unintelligibility of taking the Greek-poet-ness of Sophocles to count in favor of drinking coffee is different, for my view, than taking an evaluatively neutral consideration to count in favor of performing an action.

On my view, if an agent's take on her reasons for action aren't tracking what she takes to be good to do, you could criticize her for not tracking what she thinks is in her best interests, but this is a criticism external to her reasoning ability. You are criticizing what aims she is adopting, or her higher order values, not her ability to take a consideration as a reason and have that result in a decision. The standard view of reasons takes it that the agent's view of goodness is a constitutive part of taking a consideration as a reason.

Clarifications

For the purposes of this dissertation, I use the phrase “take a consideration to be a reason”, hoping to indicate that the agent has a mental state of the sort that leads to a decision without necessarily committing my view to a particular sort of conscious cognition on the agent’s part. For instance, when I say that an agent takes a consideration to be a reason for action, I don’t mean to indicate that the agent is necessarily aware of doing so. I take it that we can take features of the world to count in favor of performing actions without realizing that we are taking the consideration to be the case. For example, when I drive to campus and stop at a red light, I took the red light to be a reason to stop most likely without conscious reflection or self-awareness of coming to the practical conclusion that I should stop because of the red light.

Note still that considering something to be the case can still be tacit or implicit for the purposes of discussion in this dissertation. In other words, that something is a consideration does not mean that you have considered it in a conscious or self-aware fashion. Again, the example of stopping at a red light can be instructive. That the light is red is the consideration I take to support stopping, but this consideration need not occur in conscious reflection about my experience or thoughts about what to do. Similarly, I could take the consideration that Will enjoys puns to be a reason to interject a doozy into our

conversation,⁴ but I need not consciously or explicitly consider that Will enjoys puns in order to reach this conclusion.

On the other hand, when we take a consideration to be a reason for action, this consideration can't be something that we are completely ignorant of. In other words, if I am completely insensitive to a realm of features and facts, I cannot take them up as considerations or reasons. We can intuitively see in cases like the following: Imagine I go to a coffee shop on a day that they happen to have a special that I was unaware of. This is fortuitous, but if I was unaware of the special it could not have been the feature that moved me to go to the coffee shop that day. That the coffee shop is having a special couldn't have been a consideration I took to be a reason to go to the coffee shop: it wasn't a consideration I took up at all.

Taking a consideration to be a reason is thus part of the mental transitions of deciding how to act. To sum up the above points: This process need not be conscious or self-aware and the consideration need not be a mental state we are conscious or self-aware of, but this process does involve the agent's engagement with the world and therefore must be in some way accessible to her.

Structure

I use the first chapter to discuss the appeals of the guise of the good thesis (the standard view of practical reasons). First, I suggest that besides the attraction of the view

⁴ "What's brown and sticky? – A STICK!" "I went to a zoo and there was only one animal, a dog. It was a shih tzu."

itself, there are philosophical positions that can lead us to find the view intuitive. Second, I make explicit the fact that the guise of the good thesis serves two theoretical roles that a viable replacement must also address. Besides illuminating the basic “support” relationship we take reasons to perform, the guise of the good thesis seamlessly accounts for the way reasons can both guide and justify our behavior.

To put this point briefly: A virtue of the guise of the good thesis is that it makes sense of both the guiding and justifying roles we take reasons to play. When we take a consideration to support an action, as in the cases above, it points us towards performing the action. When I take the consideration that the parking spot was close to support parking there, it has weight in my thinking about where to park; my action is guided by this consideration. This is often the most intuitive role practical reasons play: they guide our actions. Considerations that support or count in favor of behaviors are just the sort of thing that would move us.

We also use reasons to defend our action against criticism: we use reasons in the activity of justification. Often in life we need to defend our action against criticism. Sometimes this criticism is lodged internally, but more often we defend ourselves from the criticisms of other agents. We do this by offering reasons, i.e., considerations that support or speak in favor of our actions. If, unbeknownst to you, there was a dangerous spider on your face and I slapped you, you could criticize my behavior. “HEY,” you might say, “You don’t get to do that!” In defense of my action, I would offer a reason: I would say “there was a dangerous spider on your face”. I take the consideration that there was a dangerous

spider on your face to support my slapping you in the face and thereby contributing to my justification in slapping you.⁵

The standard view of practical reason is in a good position to account for these roles that reasons play. If taking a consideration to be a reason for performing an action involves taking there to be something good in the action, then this makes sense of why we would be guided by these considerations. If I wonder why you did something, and in providing what you took to be a reason for the behavior you point to what you took to be good or worthwhile in doing it, that makes really good sense of what happened.

A consideration construed as pointing to the good in an action can seem to be the sort of thing to guide one's behavior and the sort of thing to appeal to in offering a justification when defending one's behavior against criticism. The support relation in deliberation is linked to the supporting relation of justification, and the guise of the good thesis handily accounts for how we see reasons play these two roles.

In my second chapter, I show that the intuitive routes to the guise of the good thesis are insufficient grounds for adopting it. I address reasons to reject the standard view and suggest that there are sufficient reasons to look for an alternative. However, an alternative will have to fill the roles that the standard view so tidily handles.

In my third chapter, I outline the direction to develop a genuine alternative to the standard view. The most promising route will be understanding the "support" relation in terms of explanation. I suggest that the burden of this alternative explanation-based view of practical reason is to also fulfil the virtues of the guise of the good thesis and this will

⁵ You may disagree with me and take it that this consideration fails to support such a violent outburst.

structure the remainder of my project. First, an alternative to the guise of the good thesis must provide an account of how we take considerations to “support” performing actions, if not by highlighting the good in doing so. Second, in providing such an account, the alternative must address the way that this support is able not simply to guide our behavior, but also can be used in defense against criticism. In other words, a viable alternative to the guise of the good thesis will account for the way that taking a consideration to be a reason for performing an action relates to taking that reason to contribute to the justification for performing an action.

I propose an explanation-based account of the supporting relation in practical reason: we take considerations to “make sense of” performing the action. While this is the most prominent strategy for rejecting the standard view, it has thus far only been presented in the most rudimentary terms. I argue that in order for it to be a true alternative, we will need to develop what it means to take a consideration to count in favor of action by “rendering intelligible” and “making sense of” so acting, a task which hasn’t been taken on thus far.

I begin my positive project by addressing the most successful versions of the explanation-based view of reasons for action in the literature: those offered by Kieran Setiya and David Velleman. Though they leave the explanatory terms (such as “making sense”) undeveloped, thus failing to progress along the first desideratum, I argue that they provide two ways of addressing the second desideratum. While both of their accounts of the relationship between taking a consideration to support performing an action and taking

there to be justification for performing the action fail, the accounts fail in ways that are illuminating for a way forward, and this will inform how we fill in the first desideratum.

I show that both Setiya's and Velleman's accounts are too extreme in how they relate taking a consideration to be a reason and how reasons could play their justificatory role. Velleman's view reduces taking the consideration to contribute to the justification of performing the action to seeing a contribution "make sense" of our action. Setiya separates these two forms of supporting an action completely. I conclude this chapter by providing a way for the explanation-based view to use the flaws in these accounts to develop a successful account of the relationship between seeing a reason and seeing justification for an action.

By focusing on explanation, my view fills the gap left by rejecting the guise of the good thesis with "making sense": instead of seeing a consideration as highlighting the action as good to do, the agent sees the consideration as showing the action as making sense to do. This is the first step towards addressing the first desideratum of an alternative to the standard view, but we need to go further. We need a fuller understanding of what it means to see a reason "explain" an action. Currently, the proponents of the explanation-based account are making do with platitudinous phrases such as "making sense" and "rendering intelligible". Without more of an idea of what this means, the alternative is incomplete. I propose we mine the vast philosophy of science literature for a theory of explanation that can provide standards for what it means for a consideration to lend explanatory support for performing an action.

In chapter four, I review the plausible theories of explanation from the philosophy of science literature before laying out Bas van Fraassen's pragmatic account of explanation as the most promising view to make sense of the supporting relation in practical reason.

According to van Fraassen, an explanation is the correct answer to a "Why?"-question. So the explanation for a house fire is the answer to "Why did the house catch fire?" His account fills in the parameters for determining the standards for answers to "Why?"-questions, and I argue that the building blocks he lays out for event-explanations in the scientific literature are the material we need to understand what it means for reasons to "make sense of" performing an action.

With this theory in hand, in chapter five I apply van Fraassen's view to the practical sphere, articulating how a reason can support performing an action by being part of a good explanation for performing the action. Reasons can explain performing an action in this view without appeal to the good in the action because the parameters for a good answer to a "Why?"-question do not make necessary appeal to the good in the action. I lay out the relationship between the way that reasons explain and how they justify in this view.

In my final chapter, I address remaining concerns with how an explanation based view of practical reasons can meet the desiderata laid out in chapter three, and how it can stack up to the virtues of the standard view. I argue that in meeting the standards laid out in this dissertation, I have provided a theory of practical reasons that can engage in the debates and meet demands of the practical reasoning literature.

In drawing on the appeal of the guise of the good thesis, the limitations of proffered views, and the philosophy of science literature, I believe I have a plausible alternative to

the standard view of how we see reasons for action. This account has the virtue of avoiding putting our valuing selves at the center of our understanding of ourselves as reasoners. By putting our ability to make sense and explain at the center of our understanding of rational agency, I will be better able to understand pathologies of rational agency such as perversity and weakness of will, and look forward to engaging in further work discussing such pathologies, questions of autonomy, and how our assessments of reason relate to our commitments to what is good, both morally and non-morally.

Chapter 1: The Standard View

In this chapter, I will introduce the *standard view of practical reason*, often called the guise of the good thesis. The guise of the good thesis, roughly, is the claim that in order to take there to be a reason to perform an action, an agent must take the action to be good to perform in some way. This view unpacks the way that we take a reason to count in favor of performing an action in terms of what take to be good in performing the action. For example, when we take the low price of beef to be a reason to plan for burgers this week, we are taking the low price to be a good-making feature of buying and preparing beef for burgers.

This understanding of how we take considerations to support performing actions has been the default view for thousands of years – we can trace it back to Socrates. Nevertheless, at its core is a substantive commitment about the relationship between our capacity to engage with reasons for action and our capacity to take things to be valuable. According to the standard view, there is a *necessary connection* between our ability to take reasons to support performing an action and our evaluative take on the world; the standard view has it that we can only understand this first capacity in terms of the second. Our evaluative capacities are front and center.

The guise of the good thesis is a view in moral psychology: it concerns how agency works. There is a correlated commitment in meta-ethics regarding what it is to be a reason – namely, that in order for a consideration to be a reason for action, it must make reference to the good in the action. While it is possible to be committed to one of these views and not the other (by considering the good to be a necessary component of reasons but not a

necessary aspect that agents must engage with in order to take a consideration to be a reason), that is not common: If you think that value plays a constitutive role in making a consideration a reason, then it will be crucial to engaging with reasons as well. If you think value is a constitutive part of engaging with reasons, it is usually because you are committed to value being a constitutive part of what it means to be a reason.¹

The purpose of this dissertation is to provide a viable alternative to the standard view. Before turning to that project, it will be helpful to address why we may be committed to the standard view in the first place. To that end, this chapter will have three sections. In the first section, I will introduce the commitments involved in the guise of the good thesis along with some theoretical motivation.

In the second section, I will diagnose other philosophical commitments that have lead us to accept the standard view.

In the third section of this chapter I will outline the independent appeal of the standard view of practical reason. This section will explain the virtues of the standard view: what it does well. Any viable alternative will need to address this appeal, and thus this will shape a way forward.

¹ Accounts in meta-ethics that have value and reason constitutively linked include teleological views (roughly, reasons count in favor of good states of affairs) and buck-passing views (roughly, good states of affairs are those supported by reasons).

Section One: The Standard View

According to the guise of the good thesis there is a constitutive² connection between taking a consideration to support your action and taking there to be some value in the action. For instance, I take “that coffee has caffeine” to be a reason for buying coffee just in case I take the caffeine in coffee as a good-making feature of buying coffee. Similarly, I take “that the parking space is close to my office” to be a reason to park there just in case I take the closeness to be a good-making feature of parking there. When we take a consideration to support performing a particular action, we take the action to be good in some way. Thus, taking there to be a reason for action is tracking what we take to be good to do. According to the guise of the good, when we deliberate and consider what to do, we are in direct contact with our evaluative assessments of the world.³

The standard view of reasons has been the standard since Socrates, was named by Aquinas, and embraced by the scholastic thinkers.⁴ From these historical roots, many contemporary theorists have found the standard view of practical reasons appealing. Christine Korsgaard, a Kantian constructivist, asserts that in every choice, the agent is making a commitment regarding what is valuable that implicates a commitment to the value of humanity (human agency itself).⁵ Joseph Raz is committed to the guise of the good thesis, claiming that without a necessary connection to value, reasons will not have their

² By “constitutive”, I mean to mark a necessary connection that explains.

³ The issue at hand is what is necessarily involved in taking a consideration to be a reason, and thus we are discussing the agent’s perspective on her action. Nevertheless, I hope to remain as agnostic as possible regarding the degree of conscious cognitive judgments being made on the part of the agent in my discussion here. Wherever possible, I mean to be minimal in what I am ascribing to the agent.

⁴ See Plato (1997), Aquinas (1964), Dressel (2014).

⁵ See Korsgaard (2009), Korsgaard (1996), and Korsgaard (2008)

normative (justificatory) power.⁶ Samuel Scheffler is committed to the standard view of practical reasons through his consequentialism and views of deliberation: we have reason to pursue that which is valuable, and rational deliberation will guide us toward what is most valuable (Scheffler, 1988, 251-2).

These three commitments to the standard view of reason represent the variety of commitments that lead to adopting the standard view: Korsgaard has meta-ethical commitments regarding the grounds of morality, Raz is concerned with justification, and Scheffler has a teleological view in normative ethics that has reasons lead us to promote what is valuable. What they have in common is the commitment that in order for an agent to take a consideration to be a reason for action, she must be taking the action to be good to do in some way.

When an agent takes a consideration as a reason to act, she is taking the consideration to count in favor of acting. So, when she takes it that a meeting fits with everyone's schedules as a reason to schedule the meeting at that time because the consideration counts in favor of doing so. If you need a class to graduate, you can take it that it supports enrolling in the class – it is a reason to enroll. According to the standard view of reasons, the way that you take these considerations to be reasons – the way you take them to support performing actions – is by highlighting the good in behaving that way. So you are taking it that fitting in with everyone's schedules is a good-making feature of

⁶ See Raz (1996). He puts the point: “Does it account for the fact that defying reasons is irrational? That one may disregard a reason only to follow a more stringent one?” Raz (1999):28

scheduling the meeting at that time. A good aspect of enrolling in that class is that you need it to graduate; that is how needing it to graduate counts in favor of enrolling in the class.

Beyond this commitment that defines the standard view of reasons, accounts of practical reason can vary along a variety of dimensions. The way in which agents weigh the relative strengths of reasons is an open question, resulting in a range of views about the structure of rational principles and deliberation.⁷ The connection between the value involved in taking there to be a reason and other values, such as moral and aesthetic, is also open. This means that the good that is constitutive of reasons for action could be seen as continuous or discontinuous with the moral good – the standard view of reason remains agnostic about rationalism in this sense.⁸ Thus, views may vary on just what it means to take something to be good – on just what the evaluative material consists in.⁹

⁷ Debates over the way that principles of rationality bind agents take place between theorists who share a commitment to the standard view of reasons. For instance, John Broome and Niko Kolodny disagree on the “scope” of rational requirements, though both seem to qualify as committed in some way as committed to the standard view of reasons: Kolodny by being a sort of reasons-first buck-passer, reading value off of reasons, and Broome by being an existence externalist realist. See Broome (2013) for the connection between seeing what you ought to do and valuing what you take your prospects to be. See: Kolodny (2005), Broome (2007), Kolodny (2007). Also, both internalists and externalists about motivation and about the existence of reason can adopt the standard view of practical reason.

⁸ There are Kantian-inspired views of the relationship between reason and value that are committed to the standard view of reasons (See Korsgaard (2009), Scanlon (1998)), but there are also theorists who do not have a constitutive connection or conflation between moral value and reasons that nevertheless espouse the standard view of reason.

⁹ For various views on what it means to take something to be valuable, or to value it, especially as it relates to taking something to be a reason, see: Brady (2003), Bratman (2007) esp. Chapter 3, Copp (1995) esp. p. 177-8, Dancy (2004), Dworkin (1993) esp. ~230, Helm (2001). Heuer (2004) Jaworska (1999), Jaworska (2007), Korsgaard (2008), Korsgaard (2003), Korsgaard (1996), Knobe and Preston-Roedder (2009), Kolodny (2003). Nagel (1970) Raz (2001), Raz (2004) Scanlon (1998), Scanlon (2014), Scheffler (2011) Schroeder (2007a) Schroeder (2007b) Seidman (2009a), Seidman (2009b), Stocker (2004), Wallace (2012). Wallace (2010). Watson (2013) Watson (1975).

Commitments can also vary about the effectiveness of taking there to be a reason for action – positions on the possibility of weakness of will remain open.¹⁰

To hold the standard view of reasons, a theorist takes there to be a constitutive connection between taking a consideration to support performing an action and taking the consideration to point to the good-making feature in performing the action. As long as pointing to a good-making feature is substantive (that is, something beyond good-making by default insofar as it is something pursued),¹¹ this will be a standard view of reasons.

Section Two: Routes to the Standard View

In this section, I will diagnose the appeal of the standard view of practical reason. I begin with some intuitive appeal that can be traced back through its historical origins, and then point to areas in the contemporary agency literature that have a tight connection between reason and the good. Value plays such an important role in central questions of agency, I suggest, that this can make the guise of the good thesis seem like the natural account of taking considerations to be reasons for action.

¹⁰ This is true in principle, though a common way to challenge the standard view is to criticize how well it can account for weakness of will. See Baker (2015) Chislenko (2014) Draft

¹¹ To be substantive, the good in question simply must not be “whatever the agent pursues”. For instance, in the first sentence of Aristotle’s *Nicomachean Ethics* (Aristotle, 1999, 1094b), he proclaims that all action an art is aimed at some telos (often translated “good”). This, in itself, does not automatically invoke evaluative content. To qualify as the guise of the good thesis, and the standard view of practical reason, this must be an evaluative notion.

Diagnostic Ground 1: Making Sense of Action

The guise of the good's wide and prolonged acceptance is in part due to its intuitive appeal. After all, in a certain frame of mind it just seems obvious that to take a consideration to count in favor or support an action is to see it as pointing to the good, the value, or the merit in so acting. In the same frame of mind one might even be at a loss to come up with an alternative unpacking of the "counting in favor of" relation. How else could a consideration support an action?

It is no accident that that it has been the standard since Plato and Socrates. Socrates articulates a commitment to the standard on the grounds that it would be unintelligible for someone to intentionally pursue what they thought was bad.¹² This interpretation of the way that we work with reasons fits with some natural ways we engage with reasons. When we wonder about one another's behavior, for instance why someone performed some action, pointing out what they saw as good in it helps make sense of things. So often, we ask, "Why did Tammy do that?" – say, when she parked in a space we're pretty sure will lead to her receiving a ticket. In response, we will be satisfied by some information about what Tammy saw as a good thing about parking there – that it is close to her next appointment, or more spacious than the intimidating nearby parallel parking spaces. Often, good-making features help to make sense of why others pursued the actions that they did.

This idea has been preserved through to the contemporary discussion through cases like Warren Quinn's radioman: The standard view is that (at least part of) what it is to take a consideration to be a reason to ϕ is to see ϕ -ing as valuable. Warren Quinn's classic

¹² See Evans (2010) pp. 6 for a discussion of the rational commitments involved in Socrates' thesis.

example seems tailor-made to motivate this constitutive connection between seeing a reason for your action and seeing it as a valuable action to perform (Quinn, 1993, 233). In Quinn's example, a man turns on any radio in his proximity, without seeing anything good about turning on radios. The radioman isn't doing it because he has a desire to turning knobs, escape silence, or play with technology; he gets no pleasure out of it. The agent simply turns on radios (246-7). This example is understood to be missing a necessary feature of behavior that is done for a reason.

According to Quinn and the proponents of the guise of the good, the missing element is what the agent saw as good or worthwhile about performing the action. (Other agents we have difficulty seeing how they conceived of their actions as good, such as Rawls' grass counter (Rawls, 1971, 432), Anscombe's mud eater (Anscombe 2000, 70-71), and Raz's agent who drinks coffee on the basis that Sophocles is a Greek poet (Raz, 1999.)

For example, T. M. Scanlon in his discussion of the case says the missing element "is the element of evaluation: the judgment that there is something good – pleasant, advantageous, or otherwise worthwhile – about performing the action" (Scanlon, 1998, 43). In order for action to make sense as having been done for a reason, the thought is that we must identify what the agent saw a valuable in pursuing the action. Because we can't appeal to the agent's seeing good in making some noise, or avoiding silence, or playing with different radio knobs, we seem to be at a loss when it comes to making sense of this behavior. The behavior seems strongly resistant to being described in terms of the agent seeing reason to turn on radios.

In making sense of one another's behavior, appealing to what they took to be valuable in their action does a great job of explaining why they did what they did.

Diagnostic Ground 2: The Good in Discussions of Agency

The guise of the good thesis has been so prevalent in part because the good is so central to discussions across the literature on agency. The good plays important roles with other questions in agency besides making sense of someone's behavior. When we try to account for the distinctive qualities of actions that are fully agential, the agent's evaluative structure or ability to track the good tends to play a central role.

Consider: there is a sense that some behaviors are better representations of our perspectives than others. The extent that the behavior can be attributed to our authentic perspective, self, or standpoint is a central issue for how we can be held accountable, and the agent's evaluative perspective has been a popular touch point in these discussions.

While accidents may be a clear category of actions that do not stem from the agent's perspective, there are further classes of concern. Consider the following case:¹³ I miss lunch and become irritable, and a friend cancels some plans we had made for the weekend. I react to the cancellation by raising my voice and have a fight with my friend. Does this action represent my feelings, beliefs, and perspective on my friend and where the friendship stands in my life?

There are a variety of standards for determining which behaviors stem from what can be considered the "true agent", or which aspects of the mental life of an agent are

¹³ Adapted from Velleman (1992) 464.

“internal” to her. In other words, when would the anger in the case of fighting with my friend be considered “internal” to my perspective, and when would it be considered “external”? It seems to some that because it was produced by the hunger-irritation rather than my feelings produced by my friendship, this could indicate that the anger, and action produced by the anger, is external to where I stand, or my perspective, as an agent. It was the hunger rather than me acting, say.

Harry Frankfurt, who largely motivated this line of philosophical investigation regarding what counts as the agent’s perspective, suggested that in order for an action to be an agent’s own (to be internal to the agent’s perspective, to be performed freely), it must be the result of a desire *the agent wanted to be effective*.¹⁴ Many philosophers have found Frankfurt’s suggestion that a second order desire will delineate where the agent stands on her behavior to be insufficient. Evaluative models have fared better, following the line of criticism begun by Gary Watson, who instead emphasized the agent’s evaluative framework as the key for identifying her perspective.

Gary Watson, in *Free Agency*, suggested that we understand the agential perspective to be an individual’s valuational system – a system of evaluative commitments that issues in evaluative judgments about which actions to pursue.¹⁵ When an agent’s motivations are in line with these evaluative judgments, we can say that the agent stands behind the action.

¹⁴ See, for example the discussion in Frankfurt (1971), Frankfurt (1987) Frankfurt (1977)

¹⁵ See: Watson (1975). Though he quickly moves away from the emphasis on explicit evaluative judgments, see Watson (1987).

There is something plausible about tying the agent's perspective to what she sees as valuable, and Watson constructs the account to be able to withstand worries over whether the attitudes or system in question can truly be identified with the agent. One's values, he supposes, are the kind of thing that it is difficult to imagine an agent disavowing, or being alienated from. Even in cases where we can see an agent working against particular values, it must be from the perspective of other of her values, else from what stance will she be able to access a standard by which to judge the value in question to be unsuitable? So it seems in order to have a stance on her actions, an agent will have evaluative judgments, and something like the network of these judgments will constitute her practical standpoint.

Since Watson, elements of the agent's valuational system have been a mainstay in identifying her practical stance, or perspective on what to do. For many, valuing itself is conceived as constituting the agent's perspective. Agnieszka Jaworska proposes that a particular style of attitude she calls "caring", which is a sort of proto-valuing, is central to establishing an agential perspective (2007a). Jeffrey Seidman, in "Valuing and Caring", identifies valuing as the attitude that frames the deliberative perspective that constitutes the agent's practical identity (2009b). Bennett Helm argues that the central aspects of our moral psychology are value-laden (2001).

Consider, for instance, the account put forward by Michael Bratman. His account is a planning-focused view of agency, and on his account, to fully act, the agent will be acting from her values (2007). An agent, according to Bratman, is constituted by those parts of her mental economy that are continuous over time via interlocking and reflexive relations (2014). Plans and policies are such aspects of the agent's mental economy, and

thus acting on plans and policies will be acting in a full-bodied, or autonomous fashion – it will be action that speaks for the agent. To value is to have a policy about how one’s life should unfold, and this is constitutively connected to autonomous action.¹⁶

It is not just in articulating the agent’s perspective that the good plays a central role. The good also shows up when we consider the structure of practical reason. Practical reasoning is guided (or structured) by norms or principles. The grounds or aim of practical reason is another area that the good tends to be tightly entwined with agency. There are two main ways that this can happen.

First, some theorists take it that principles of practical reason are grounded in a parallel way to theoretical reasoning. Because in theoretical reasoning, truth plays a normative role, something like the good plays a normative role in practical reasoning. Beliefs are aimed at getting the world right in a way that other theoretical attitudes are not, and this is the sense that truth is a constitutive norm of the realm of belief.¹⁷ In the theoretical domain, truth plays this normative or governing role in the sense that to believe something, rather than hypothesize, guess, fantasize, etc., the attitude in question is truth-sensitive. It is not faulty for a fantasy to fail to track truth, but it is faulty for a belief to fail to do so. A success condition on a belief is that it be true, or match the world or evidence that the believer have. If I hold a belief about the world – say, that my car is parked outside the coffee shop – I should adjust the belief if information comes to my attention that

¹⁶ Ronald Dworkin’s account is similar to Bratman’s in that it appeals to the agent’s life as a whole in order to ground more temporally isolated psychological features. See: Dworkin (1993)

¹⁷ For a review of the discussion of truth as the aim of theoretical reasoning, see: David (2005) McHugh (2013), Owens (2003), Shah (2003), Shah and Velleman (2005), Velleman (2000a), Wedgwood, (2002), Zalabardo (2010)

suggests the truth of the belief has altered – say, my alarm starts going off and then gradually decreases in volume. If I were merely imagining that my car was parked outside (or hypothesizing, fantasizing, or holding some other attitude), then truth doesn't govern the belief, so finding out information that suggest it isn't true doesn't govern the belief in the same way.

Thus it can be said that in theoretical reasoning, there are norms governing belief. From this insight comes the notion that a constitutive norm of believing/theoretical reason is that it aim at the true. In other words, believing is governed by the true: in order to believe successfully, one must be sensitive to considerations that relate to truth.

The norms of the practical domain have been articulated in a parallel structure to this picture of the theoretical realm.¹⁸ Theoretical reasoning is guided by a norms concerned with getting the world right (the truth of propositions), and is aimed at producing or modifying attitudes of belief. Practical reasoning is guided by norms concerned with acting appropriately (the desirability or value of actions), and is aimed at producing or modifying attitudes of intention^{19,20} The same kind of normative structure grounding theoretical reasoning must be at work in the practical domain.

In our practical agency, the core attitude is one of motivation rather than assent, but theorists are committed to a picture where the attitudes in the practical domain function according to a standard in the same way as the attitudes in the theoretical domain.

¹⁸ And also have been criticized:, see: Millar (2009), Enoch (2006)

¹⁹ (or, controversially, behaviors – Aristotelian practical syllogisms – intentional action is the product of practical reason but that will only be intelligible as a reflection of mental states)

²⁰ See: Harman (1986), Bratman (1987)

Sergio Tenenbaum uses this strategy in his book, *Appearances of the Good*. Tenenbaum claims that “the good” constrains our action, with a central main focus on how agents *conceive of* this good in order for it to play the constraining/guiding role that the true plays in the theoretical domain. The aim of action, according to Tenenbaum, is what *appears* to the agent as “good”, which Tenenbaum defines as “worth-pursuing”. For something to appear to be worth-pursuing, no judgment is necessarily involved. In his account, the object of action *strikes the agent in some way* as worthy of being pursued in a similar way that, in the theoretical realm, the object of belief *appears* to be true.

Tenenbaum provides the following three examples of how appearances of states of affairs work in theoretical reasoning that will parallel appearances of good in the practical: “[1] from far above, the car appears very small;...[2] the raccoon appears to be dead;...[3] presented this way, the argument appears to be valid, but when we formalize it, we see that it is not”. When the world appears in particular ways, this is fodder for our theoretical reasoning, which is aimed at the truth.

Practical reasoning involves parallel “appearances”, namely, of the good. Actions can appear to be worth pursuing because we have a conative reaction to them, because of preliminary considerations that suggest they are pursuit-worthy, or because they are the means to a decided end. *Appearing* to be worth-pursuing is thus compatible with a variety of degrees of settled judgments, including those that judge an action to be bad.²¹

²¹ Appearances do not require judgments or endorsements of what appears to be the case. I can admit that the raccoon appears to be dead while not believing it to be the case. In a parallel way it can appear to me that jumping off a building is worth pursuing without any preliminary judgment or belief that this is actually the case.

Because for Tenenbaum desire is guided by such appearances of the good, the agent's conception of value is a constraint on what she desires – she can only desire that which at least appears good.²² This is meant to be parallel to the theoretical domain where agents can only believe what appears to be true.

A number of views have the good or valuable as the guiding norm of practical reason or deliberation on the grounds that there must be a guiding notion to such reasoning (it must have such a structure) without appealing to a parallel with the theoretical realm.²³ The thought here is that in order for there to be standards for good or bad reasons/reasoning, there must be something that this domain is aimed at, or something governing the domain. Goodness or value can be taken as a prime candidate for this role.

Many historical figures held views that lent tacit support to the standard view because of their intuitive connection or endorsement of an evaluative foundation of the structure of practical reason. Practical reason can be said to be guided by principles, such as the principle of instrumental reasoning. The structure of practical reason is an articulation, or the metaphysical reality, of the relation between such principles. For many theorists, in order for these principles to apply to an agent, or to have true normative force, they must be grounded in some form of the good. In other words, something valuable must

²² Tenenbaum makes it clear he thinks his view can accommodate a level of perverse action, stating that an intelligible aim “can, for example, be imprudent, immoral, or something that is not fit for a human being to want” (Tenenbaum, 2007). The focus here is on what is *intelligible*, whether the agent can see the *point* in what she is doing, not in whether the agent sees her action as *valuable* or *good* in the sense Velleman focuses on in (Velleman 2000). According to Tenenbaum, “actions, aims, projects, and so on can be foolish, frivolous, or cruel”, but by seeing the *point* of acting according to foolishness or frivolity, the agent sees frivolity or foolishness as worth-pursuing, and therefore acts under the guise of the good (Tenenbaum, 2007).

²³ Many of these views make up the constructivist or constitutivist class in meta-ethics. See, for instance: Bagnoli (2002), Bagnoli (2009a), Lebar (2008) Lenman (2010) Reath (2006) Reath (Forthcoming), Street (2008), Street (2010)

be a constitutive part of a foundational principle, or the norms must be guiding towards something of value.

Consider the Kantian framework of agency. Any right action must be in accord with the value of humanity, which grounds the value of any particular maxim an agent may act on. The worth of agency is the grounds for particular agents acting in particular ways; actions that disrespect that value are not in line with reason – they are unsupported by reason. A maxim (a principle underlying one’s reason for action) such as “in circumstances *c*, I will ϕ ” will follow practical principles if it is in line with valuing humanity.²⁴ For Aristotelians, the structure of practical reason is grounded in substantive truths of the good life. What we have reason to do will be aimed at what is good for us, understood in terms of distinctive human functioning.

This rough articulation of the role that evaluative content plays in these historical accounts, and the contemporary accounts inspired by them,²⁵ captures the *normative* motivation for connecting the good to practical reasons. In order for the principles that articulate practical reason to play their function (of binding agency, of making agents go wrong when they fail to follow them), the idea here seems to be that something besides reasons needs to be at work to ground them. There must be something of value that the reasons are capturing – otherwise, when we take there to be a reason to sit in a particular chair at a coffee shop, or take up a career path, or help a friend in need, will not be able to

²⁴See: Kant, Immanuel (2011)

²⁵ Bagnoli (2002), “Moral Constructivism: A Phenomenological Argument,” *Topoi*, 21: 125–138. Bagnoli (2009a) Bagnoli (2008) Lenman (2010)

be distinguished from one another: are we tracking the good life? (the Aristotelian concern)

Are we in line with the truly reasonable?

Section Three: Virtues of the Standard View

The standard view has two distinct virtues: it illuminates what it means for a consideration to be a reason, and it does so in a way that makes sense of the roles that reasons play in our practical lives. I will say more about these virtues here, for any viable alternative will have to address these virtues as well.

First, as I emphasized above, the view illuminates, or unpacks, the basic “support” relationship that reasons perform. When we take a consideration to be a reason for action, or when we act for a reason, there are a cluster of phrases and terms that we use in an attempt to capture the *mode* in which we are engaging this consideration. We can engage with considerations in a number of different ways besides taking them to be a reason – we can *imagine* that they are true, *hope* they are true, note that two considerations are *always true together*, that some consideration *isn't* the case, etc. When we engage with a consideration *as a reason*, we see it *count in favor*, or *support*, acting in a particular way, or believing something. Phrases like “counts in favor”, “favors”, and “supports” are taken to be getting at the way that the consideration plays its role as a reason – it points us towards something.

The standard view sheds some light on how this relationship may work between a consideration and performing an action. When we take a consideration to be a reason for performing a particular action it supports performing it by pointing to the good or value in

doing so. The standard view thus provides a further unpacking of the common phrases theorists use to understand what we see when we engage with considerations as reasons.

The second virtue of the standard view of practical reasons is that, in the way it unpacks the support relation, it also seamlessly accounts for the roles that we take reasons to play. Because reasons *support* performing actions, they play two roles: they *guide* our actions and they *justify* our actions. A virtue of the guise of the good thesis is that it unpacks these roles in a very natural way.

Reasons *guide* our actions because when we take a consideration to be a reason, we are engaged with it in the mode where we see it *support* performing the action, so we are pointed towards acting in that way. A central way we engage with reasons is in determining what to do. Reasons are relevant to this activity because of their directional nature – they favor and support. If we wonder what to do, and we take considerations to support performing some action or other, that will be relevant to our concern – what to do.

Reasons also *justify* our actions. Besides having bearing on the question of what to do, reasons are what we offer one another in defense of our behavior, this is our practice of justification. Because when we see a consideration as a reason, the consideration counts in favor of doing something, this support relation suggests that the action is *the thing to do*, allowing us to provide the consideration in defense of our performing the action. If I do something, and you don't consider it to be justified, you are taking there to be insufficient reason for doing it. If I can provide a reason in defense of it, that reason justifies the behavior.

A consideration supports, and thus can guide my behavior when I take it to support and defend my behavior if I offer it in support. The standard view of practical reason, in unpacking the support relation in terms of *value*, accounts for the way reasons both guide and justify. In taking there to be a good-making feature in an action I can be guided to perform it.

Imagine also when someone criticizes my action, suggesting that I am not justified. A natural way of understanding this is that they are suggesting that there is not value to what I have done, or too little value. The standard view of reasons can make sense of this justificatory role of reasons, for I will be justified when considerations were tracking value and fail to be justified when I fail to track value. A natural way of responding to justificatory criticism is by pointing to the value in the action.

The supporting relation in deliberation is linked to the supporting relation of justification, and the guise of the good handily accounts for how we see reasons play these two roles.

The ability to seamlessly account for these roles of reasons is crucial for a view of practical reason. The support relation of reason is what allows reasons to play the role they do in our deliberative practices of guidance and justification, and any alternative to the standard view will not only have to illuminate the support relation, but do so in such a way that preserves these roles.

Conclusion

In this chapter I discussed reasons for adopting the standard view. I contend that there are commitments in meta-ethics of the structure of practical reasoning, and action theory relating to agential authority, that can lead to finding the standard view intuitive.

I articulated virtues of the standard view of reasons that any alternative will have to account for in attempting to replace it. In the remaining chapters I will turn to developing a viable alternative to the standard view, and these virtues will be desiderata for my replacement. In the following chapter, I will use the most fruitful attempts at rejecting the standard view thus far to further set up my successful alternative.

In the following chapter I will address weakness of the standard view and reasons to turn to an alternative.

Chapter 2: Why Reject the Standard View?

In the previous chapter, I outlined the standard view and reasons that it may seem like the natural view to adopt. In this chapter, I will attend to reasons to reject the standard view.

I will first show that despite the way that the good shows up in many places in the literature on agency and practical reasoning, this should not mean that we are too quick in adopting the standard view, which commits us to a necessary connection between taking a consideration as a reason for action and taking there to be something good in performing the action.

Next I will outline the major cases against the standard view from the literature rejecting it. These are counterexamples that represent instances of action that the standard view fails to account for. There are classes of cases that the standard view has a difficult time accommodating, which I suggest indicates a deeper concern about the way that the standard view attempts to account for our ability to track reasons. Because the standard view puts our ability to track evaluative content front and center, the possibility that there are classes of cases where evaluation does not play any key role is troubling, and suggests the standard view misidentifies goodness as the guiding factor of our reasoning.

I conclude with a theoretical advantage of moving away from the standard view of practical reasoning and towards an alternative. Accounts that are committed to the standard view of practical reason have special problem accommodating what Josef Raz terms the “basic belief”, and I suggest that an alternative will do better.

Thus, this chapter points us towards adopting an alternative to the standard view of practical reasons. The commitment to a constitutive connection between taking a consideration as a reason for action and taking the action to be valuable is misguided. Here are some reasons why.

Section One: Undermining the Intuitive Grounds for Standard View

Though the good shows up in a variety of places in the discussions of agency, this does not lead to a necessary connection between reason and the good. In the previous chapter, I outlined some intuitive connections between agency and the good that could explain the appeal of the standard view. The prevalence of the good in this literature may make it unsurprising that theorists have found the commitment to a connection between reason and the good to be natural. However, unsurprisingly, I will now outline how this connection need not be a necessary one.

When we attempt to articulate the standards for full-fledged action, or what will count as the agent's authentic perspective, the agent's evaluative perspective naturally arises. However, this doesn't mean that in order to act at all the agent's evaluative perspective is invoked. Views of agential authority are not usually explicitly in the business of providing accounts of practical reasons as such. Intuitively, these are different questions – we sometimes act for reasons but nevertheless behave in a way that doesn't truly “speak for” us. However, this doesn't leave the accounts of agential authority and the standard view of reasons in a clear relationship. Because philosophers like Frankfurt are providing accounts of “free action” in distinguishing agential authority, and others like Velleman

oppose full-fledged action and action of non-human animals, this suggests that the account of action they provide is a minimal threshold of action, say acting for a reason. However, action that is “full-fledged” or “speaks for” the agent meets a higher bar of action – say acting with agential authority. It is not always clear in this literature which class of action is being analyzed, as John Martin Fischer discusses in Fischer (2012).

Setting aside the discussion of lack of clarity regarding the domain of behavior the discussion is meant to address, we can consider the connection between agency involving taking a consideration as a reason and that involved in taking something to good. We can see that there is not a *necessary* connection here. Consider two views with the tightest connections between value and the agent’s perspective: Watson and Bratman – even on such views we can see that evaluative commitments need not be involved in seeing a reason for action (Bratman 2007, Watson 1975). Though full-bodied action may involve her values and evaluative commitments, taking there to be a reason for acting need not.

For Watson, full-blooded action will stem from the agent’s valuational system. However, as we can see in his latest discussion of psychopathy, he sees the possibility of a distinction from action that flows from a full agent’s perspective and action that is done from taking there to be a reason to perform it (Watson, 2013). Psychopaths may populate this class of agents, who are distinctive in his view for lacking a diachronic concern for themselves.

Watson characterizes valuing as investing in something that provides a standpoint for self-correction (Watson, 2013, 277). The ability he finds to be crucially missing in psychopathic agents is the ability to transcend present aims as a way to be able to relate

and value others. This is tied importantly to the ability to transcend present aims as a way to be able to relate to your future self,¹ which may be conceived of as a stepping point to relating to and valuing others.

The psychopaths are not self-critical. They don't see their behavior as having any impact on the value or worth of themselves, as having normative impact on their future behavior – being grounds for adjusting their behavior in the future. Watson characterizes this deficiency as not seeing *themselves* as having any worth; indeed, perhaps an inability to take up anything as having practical significance beyond actual specific aims of the moment.² So, according to Watson, a psychopath is capable of “practical rationality...the acceptance of an objective that to some extent structures and guides his thought about what to do next (or soon)...he will typically have practical knowledge of what he's up to” (277). However, “they provide no basis for self-criticism or evaluation of the objectives themselves” (277). Thus, the psychopaths show a type of reflexive complexity that falls short of caring. They are often into what they're doing, just as we all are often freely engaged in activities that do not matter much to us. That capacity does not constitute a critical standpoint.

Thus, according to Watson, it is possible to act for a reason without invoking the evaluative and reflective commitments that constitute full-blooded action. The agency of

¹ Note: Bratman also emphasizes the importance of diachronic investment to establish agency and a will. Without such investment, you are not really exerting your will at a time.

² Another way of conceptualizing the psychopath is as a “rational wanton” in Frankfurt's taxonomy of reflective willing. Such a person has reflective capacities, and therefore can act on reasons, but is incapable of caring about this capacity or how she uses it. See: Watson, 2013, fn 14. To care, or have reflective evaluation of one's will, you have to do more than freely accept the movements of your will. It requires a diachronic dimension – an investment in the movements of your will over time.

an agent with the diachronic commitments referenced here has a perspective from which she can judge which behaviors to take up. Acting for a reason is a more minimal process that can have this diachronic perspective, but need not.

The distinction between acting with full agential authority and taking there to be a reason to act comes through even in Bratman's view, which connects valuing and autonomous action constitutively as well. Recall that for Bratman to act autonomously is to behave on the basis of a self-governing policy, and that to have a value is to have a policy about how to live one's life. This is just what it is to act with a full-blown agential perspective, so what it means to act from a perspective that is fully yours may be to act from your evaluative take on the world. However, we can see explicitly that it will be possible to act for a reason without performing an action with all of this machinery.

In "Valuing and the Will", Bratman constructs a creature that is capable of valuing via Gricean creature construction in order to articulate the capacities involved in full-blooded agency (Bratman, 2007, 49). He is motivated by the distinction between valuing and judging good, as well as the distinction between valuing and "mere" wanting (Bratman, 2007, 48). For Bratman, valuing will be tied to his planning theory of intention and connected to the person's will. He begins with a creature who has the barest agential capacities and who clearly lacks what we would require for valuing, and proceeds through Gricean creature construction to explore what capacities are required for valuing and acting from self-governing policies. The agents/creatures he constructs have aspects of rational agency and what may be called valuing along the way to full-blooded valuing or autonomy at full development, but what is clear from his construction is that at the most basic level,

an agent can take a reason to support performing an action without the higher capacities that allow for true agential authority. Bratman, with his creature-construction method of argument, has created a creature that can act for a reason before he has defined a creature that acts from valuing.

This much should be intuitive – we act from an appreciation of reasons even when such performances don't fully speak for us. While the intuitions in the beginning of this section surely are valid ones – acting for reasons can be a good litmus test for when we are acting intentionally, and will feed into determinations of when the action is our own (certainly in the negative cases – when we don't act for a reason it seems likely that the action doesn't speak for us) these intuitions are limited. It takes more than acting for a reason to be sufficient for an action to be “fully” the agent's, or to indicate where the agent stands.

Recall the example where I fight with my friend when hungry. I can do so in a state where I see reasons, despite this not being my considered opinion of what is best to do. Thus, we could say that the behavior doesn't have full agential authority, despite being performed by reasons, and thus is a case of my taking there to be reasons for action but not speaking for me.

Though a concern with agential authority can make the standard view, with its tight connection between reasons and value, seem attractive, it is important to keep a distinction between acting with agential authority and acting because you took there to be reason to do so, which is a potentially weaker (and broader) class of actions. Espousing a view of agential authority that invokes evaluative content, or evaluative commitments on the part

of the agent thus shouldn't automatically lead to a commitment to the standard view of reasons for action.

Also, as I discussed in the previous chapter, when we inquire into the necessary structure of practical reason, many views have the good front and center. Value underlies the force of the principles of reason in both realist and anti-realist ways – either grounding reason claims by appeal to what is good in the world (as a realist would), or grounding reason claims based on the agent's assessment of the good (as a constitutivist would).

Theorists who have sympathy with the view that theoretical and practical reasoning have a parallel structure need not accept the standard view of practical reasons, however. An account of practical reason can be committed to the claim that practical reason must have a normative aim or constraint just as belief, or theoretical reasoning, is governed by truth without having something evaluative play such a role. David Velleman suggests that an appropriate parallel or analogy to the truth would be the “attainable”, as what we desire, or intend, is distinct from the rest of our conative attitudes (like wish, hope, etc.) in that we must see it as possible to achieve (Velleman, 2000).³ Gilbert Harman argues that practical reasoning is aimed at increasing explanatory coherence (Harman, 1997). Thus, a commitment to a parallel or analogous structure between the theoretical and practical need not lead to a commitment to the standard view of practical reasons. There will need to be

³ Though the attainable doesn't seem to play the same role he attributes to the defining feature of belief in the theoretical realm. It is unclear whether he is aiming for a sufficient feature to delineate among attitudes in a direction of fit class or a necessary feature. The “attainable” does seem to be a necessary feature to distinguish the attitude in question, but may not be sufficient, in which case some evaluative content may nevertheless be required and the guise of the good may sneak in.

further reason to think that something evaluative is needed to play the structuring, governing, or guiding role that truth plays in the theoretical domain.

Certainly, accounts make such cases or have such commitments, and some do so without appeal to analogy between the domains of reason. In other words, parallels between the theoretical and practical domain are not the only way that the good has been envisioned as the guiding norm of practical reasoning.

These accounts are clearly amendable to the standard view of practical reasons. If something evaluative – like a substantive notion of the good life, or the good will/value of humanity/etc. – is grounding the structure of practical reason, it is intuitive that reasons themselves include an evaluative element. It can be difficult to articulate the commitment of these views in a way that doesn't simply amount to a commitment to the standard view.

Is it possible to think that the structure of practical reason requires evaluative grounding but yet taking a consideration to be a reason need not involve evaluative content? Perhaps. Imagine a theorist who is committed to a view in meta-ethics about the nature of reason, but nevertheless thinks that agents can (mistakenly) take a consideration to support an action because it is bad. In other words, depending on this theorist's commitments about the possible mistakes an agent can make while still qualifying as taking a consideration to be a reason, this position is entirely coherent. Though a consideration, in order to be a reason, must make some appeal to the good, an agent may not need to take the consideration to make appeal to the good.

It seems conceivable that there is value grounding a structure of principles that delineate how reasoning should take place, and though reasons are the material of these

processes they themselves don't include the ultimate grounding for the system. This seems a clunky way of making sense of our reasoning, however, and engaging with these views regarding the role of evaluative content in the structure of practical reasoning may amount to engaging with the debate over the standard view of practical reasoning itself: should we adopt the guise of the good thesis? I'll spend the rest of this chapter suggesting that we should not.

The good is taken by many theorists to play crucial roles in agency and practical reason. This can lead to finding the guise of the good thesis as a natural or attractive view in moral psychology. However, there are a number of concerns with the thesis, which I will turn to now.

Section Two: Cases against the Standard View

In the last several decades, the standard view has been criticized by theorists who have attempted to incorporate a richer pantheon of human agency into the fold than the guise of the good seems to allow. Because the guise of the good thesis is a claim about a necessary perspective the agent must have in virtue of taking a consideration to be a reason for action, these attempts put pressure on the standard view of practical reason's ability to account for all instances of agency. The dialectic has largely been focused on purported

counterexamples to the view.⁴ These would be cases where an agent acts intentionally, or takes there to be a reason for her behavior, yet nevertheless does not take there to be good in so acting.

Some theorists are concerned about preserving the possibility of weakness of will in the face of the constitutive commitment of the standard view. In such cases, the weight of reason or motivation is out of whack with the degree of the evaluation. If the standard view of reason claims that the agent must take her action to be good to do in order to act for a reason, what about when agents have a view on what would be good, but act some other way? Sergio Tenenbaum (Tenenbaum, 2013) and Derek Baker (Baker 2015)⁵ discuss this concern, which may arise out of principles of intentional action Donald Davidson discusses in 1980, which include: “If an agent judges that it would be better to do x than to do y, then he wants to do x more than he wants to do y” (Davidson, 1980, 23). A result of this principle is that our motivations will be determined by our evaluations.⁶ For example, you may purchase new shoes despite taking donating to Oxfam to be the more valuable option.⁷

⁴ Some notable exceptions to discussing the guise of the good thesis in terms of counterexamples have been Kieran Setiya’s discussion in Setiya (2007) and Sergio Tenenbaum’s account of the guise of the good in Tenenbaum (2007)

⁵ Baker takes it that in order to accommodate the phenomenon of weakness of will, the standard view will have to take on a form that lacks substance to the point of being uninteresting.

⁶ (Davidson does not clearly accept this principle, but draws out tensions in characterizations of principles we would readily find acceptable)

⁷ I take it that there is some conflation going on in examples like these. “Better” is not evaluative content in the sense that “good” is, quite, so these cases may not fit into our framework as neatly as the others. Nevertheless, they persist as purported difficulties for the guise of the good thesis. I take it that they are potential problems for how theories of practical reason can account for deliberation. It is a way of articulating worries over weakness of will itself, more than a problem with how our take on evaluative content and reasoning relate.

Another class of counterexamples are those that may be categorized as evaluatively neutral. Donald Davidson can be interpreted as providing such a case in “Actions, Reasons, and Causes”, where he describes an agent with a yen to drink a can of paint (Davidson, 2001). Discussing the class of attitudes that can be taken to rationalize (be a reason-explanation) for an agent’s behavior, he says,

In general, pro-attitudes must not be taken for convictions, however temporary, that every action of a certain kind ought to be performed, is worth performing, or is, all things considered, desirable. On the contrary, a man may all his life have a yen, say, to drink a can of paint, without ever, even at the moment he yields, believing it would be worth doing. (Davidson, 2001, 4)

The action of drinking a can of paint isn’t performed out of tracking some good, but is done intentionally.⁸ Gary Watson considers such cases: a woman who considers drowning her screaming child and a tennis player considering smashing his racket, neither of whom have positive evaluations of their actions (Watson, 1982, 101). If agents can act for reasons but nevertheless don’t view what they are doing as good, then the standard view of reasons isn’t capturing the way they are being guided by reasons.⁹

The class of counterexamples that have generally been taken to be the most compelling share a form with Michael Stocker’s discussion of Satan from his classic article “Desiring the Bad: An Essay in Moral Psychology”. In it, Stocker considers the case of bringing about a bad state of affairs on the basis or grounds that it is bad (Stocker, 1979,

⁸ This is, of course, controversial. See Mele (2003).

⁹ Kieran Setiya proposes a series of examples where an agent acts for a reason but nevertheless does not see her action to be good in Setiya (2010)

739-40).¹⁰ Consider the case from Milton's *Paradise Lost*. Satan famously proclaims, "Evil, be thou my good!" thereby adopting the bad as his aim in acting (Milton, 2003, Book IV Line 10). "Good" here is playing the functional role of the aim in one's action and Satan, in his declaration, is having his conception of evil (substantive badness) play this role in guiding his actions.¹¹ If humans can be guided in their actions in this way – namely, by their conceptions of what is bad, worthless, or evil – then the standard view of practical reason is missing something.

Structurally similar to this counterexample is the self-destructive agent proposed by David Velleman in "The Guise of the Good" (Velleman, 1992). He asks us to consider an agent who has perhaps suffered a disappointment and has found herself in a depressed and self-destructive mood (Velleman, 2000, 121). The only actions this agent will consider are those she sees as bad or undesirable. Any action that she realizes would improve her situation or mood or that does something positive is rejected as a course of action on those very grounds. The agent's perspective is described thus:

I am moved to stay at home, refuse all invitations, keep the shades drawn, and privately curse the day I was born... What's more, I engage in these actions not only out of despair but also in light of and on the grounds of despair. That is, despair is part of my reason as well as part of my motive for acting... My reason for acting thus includes not only my mood but also an unconditionally negative evaluation of the action. (Velleman, 2000, 121)

¹⁰ Stanley Benn also takes for granted that humans can act in this fashion: Benn (1985)

¹¹ David Velleman interprets the case in a weaker sense, though he rejects the standard view of practical reasons. Velleman (2000)

The agent in question here considers what to do and chooses an action because it expresses her despair and she considers it to be bad. This agent, in other words, takes a consideration to count in favor or supporting her action insofar as that consideration points to the bad in the action. This example suggests that the counting in favor relation need not be unpacked in terms of the good but can also be unpacked in terms of the bad. It suggests that the guise of the good thesis is overly restrictive and thus false.

An adherent of the guise of the good thesis has two ways of responding to this example without rejecting it outright. She can either characterize the behavior as in fact under the guise of the good (because the agent sees something good in the action after all) or as not involving the agent seeing there to be a reason for her behavior.

If the proponent of the guise of the good thesis takes the latter route, then she is in effect assimilating Velleman's self-destructive agent to Warren Quinn's famous radioman. Quinn's radioman turns on any radio in his proximity, without taking anything to count in favor of doing so (Quinn, 1993, 233). The radioman isn't doing it because he has a desire to turn knobs, escape silence, or play with technology; he gets no pleasure out of it and sees no point to his behavior. In other words, he is not acting for a reason. He is simply turning on radios.

It seems, though, highly problematic to assimilate the two cases; they are, after all, importantly disanalogous. While the radioman does not see any consideration as supporting his behavior, the self-destructive agent does: she sees considerations as counting in favor of her action insofar as they highlight the bad in her actions. While the radio man is not guided by any consideration, the self-destructive agent is.

As Velleman describes the case, then, it looks like the better strategy for the guise of the good adherent would be to try to make a case for the agent's seeing reason for her behavior, and doing so under the guise of the good. But to go this route, the guise of the good proponent must say that the self-destructive agent is viewing her own destructive behavior as good. However, in this case, if she judged that the action was good, she would not see that as a consideration that *supports* performing the action because she only wants to do what is bad. The interpretation of the depressed agent here as acting under the guise of the good misses how her behavior and attitudes would change if she came to think of a potential action as good or worthwhile. In saying that the agent sees her action as good, the guise of the good thesis has thus classified the agent's reasoning in a way that she would not accept.¹² The thesis cannot give an account of the agent who takes the very fact that the possible action is bad to be reason to perform it.

Thus, counterexamples take a few forms, all purporting to illustrate cases of agents who take there to be reasons to perform an action while nevertheless not taking that reason to highlight some good in performing the action. This can be the case because the agent is conceiving of her action as bad to do, or because she is acting weakly, or because the agent is conceiving of her action in neutral terms. While examples on their own will likely not decide the debate over the relationship between taking there to be a reason for ϕ -ing and

¹²The proponent of the guise of the good theorist may suggest that the agent is making an *implicit* assessment of her action as good to do, and that is what is guiding here despite her commitment to the badness of the action. We do, after all, make implicit or inchoate judgments and assessments frequently in our practical lives. However, this is problematic or at least not simple in cases like this one, where the agent is guided by her assessment that is directly opposed by the judgment the proponent is suggesting to ascribe. Introducing an implicit judgment of the action as good introduces an internal conflict to the agent. There isn't a way of retaining the structure of this example and ascribing an assessment of value on the agent's part.

seeing ϕ -ing to be valuable, cases like this at least challenge the standard view. What the examples show is that it does seem intuitive that in some cases the agent doesn't *conceive of herself* as pursuing what is valuable. In a debate concerning the agent's perspective it is *ad hoc* or a distortion to claim that she is acting under the guise of the good. We should avoid drawing conclusions about the agent's perspective by insisting that she is really committed to the value of her action when in fact she wouldn't agree to such a claim herself.

If these counterexamples are compelling, it is because they can account for the way that agents take there to be reasons for action without making necessary appeal to the agent's taking there to be good in performing the action. Those theorists that are committed to the standard view consider these cases to be missing a crucial element of the agent's perspective – namely, the evaluative take on the action. In other words, in cases such as Milton's Satan and Velleman's self-destructive agent, there is a way in which the agents take their action to be good, or else they are incoherent. Similar claims will be made against the other cases.

In the counterexamples above, we saw that there may be reason to think of the reasoning of human agents in ways that don't commit us to the standard view. A further motivation for breaking with the standard view is that we can imagine (and encounter) reasoners that don't have evaluative capacities. These would succeed in qualifying as counterexamples to the standard view.

For instance, consider human agents with deficits such as patients with advanced dementia, or children. Agnieszka Jaworska considers these agents to be at the "margins" and one of the prime agential distinguishers between adults with full capacities and such

agents is their diminished evaluative capacities. However, she claims that they still retain key features of reasoning agency that ought to be nurtured or respected.¹³ Further examples may be non-human animals, which can be conceived as agents that do act for reasons, but lack the cognitive development to take things to be good (evaluative perspectives being more cognitively demanding than the skills and abilities required for tracking features in the world and acting on them).

Where ever we land on the question of these possible counter examples, they represent mounting pressure on the standard view of reasons. Yet it remains the standard. I turn now to a reason to move away from the standard – a feature of our agency that an alternative will be able to capture better.

Section Three: What could an Alternative Do Better?

The standard view of reasons has it that reasons depend on value in a constitutive way. This is the underlying commitment to the view, and that should be the focus of the discussion. I suggest that this is what is inherently problematic.

This commitment not only leads to worries over the above counterexamples, it leads to problems understanding the shape of our rational agency. This can be brought out regarding approaches to what Josef Raz calls “the basic belief” (Raz, 1999a and Raz 1999b). Views that have reason necessarily connected to value in the way that the guise of

¹³See: Jaworska (2007b), Jaworska (1999).

the good thesis requires have difficulty accounting for this commitment, and an alternative will be able to better address it, for reasons will be able to “outstrip” value.

It is easy enough to note that there being reason to do something is intimately connected to what is valuable. Most accounts of reasons and value will have this be the case - we take it to be data that reasons and value are tied in some way. Teleologists go beyond claiming this connection to be intimate and understand one in terms of the other. To put it quickly at first: For “teleologists”, an agent’s reasons are *functions* of what is valuable.¹⁴

In other words, a teleologist claims that an agent’s reasons at a time can be understood once we understand the values in play.¹⁵ Values determine the agent’s reasons. One starts with an account of what is valuable, and so what states of affairs would be good to bring about, and that will set what an agent has reasons to do, for (roughly speaking) we have reason to bring about the good states of affairs. If we lay out everything that is valuable in a situation, we will have laid out all of the relevant aspects of the reasons as well. Reasons are given from what is valuable.

This specific picture of deliberation is an inessential aspect of teleologism for my purposes (it includes specific principles of reasoning), but it highlights a motivation for adopting the teleological framework. The crucial aspect of teleologism is that reasons be understood in terms of what is valuable. Part of the appeal of such a view (with such a strong connection between reasons and value) is the picture of deliberation that invokes

¹⁴ See Scanlon (1998) for a characterization of “teleologist” similar to this one. Raz (2001) has a slightly different characterization but non-problematically so.

¹⁵ Teleologists, according to this definition, include Samuel Scheffler, Philip Petit, and Joseph Raz.

these principles. If the purpose of deliberation is to select *the best state of affairs* to be brought about, and reasons are the material of deliberation, then this would motivate the teleologist. But as far as I am claiming, to be a teleologist is just to be someone who thinks what one has reason to do is fully determined by what one sees as valuable. Thus those who find a consequentialist/maximizing view with regard to deliberation attractive will likely find a teleological view of the relationship between reasons and values attractive.¹⁶

What an agent has reason to do is exhausted by the valuable for the teleologist. This view leads to adopting the standard view of practical reason because the way a reason will support performing an action will necessarily involve evaluative content.

The teleological relationship between reasons and values is problematic, because agents seem to have different reasons in circumstances where the values in question are the same. Say, for instance, a career in the military is valuable. This value translates to different reasons for me, as a graduate student in California, to perform actions than for the Commander in Chief in D.C. I do not have a reason to sign a bill giving a selection of Navy personnel raises, whereas the Commander in Chief doesn't have the reasons I do to enlist at my local Navy recruitment center. Given the values in these circumstances, different reasons seem to result.

The teleologist can claim that the reasons that track value are conditional, and thus capture the different circumstances that the Commander in Chief and I find ourselves in.

¹⁶ I won't be discussing deliberative forces and pressure explicitly here. A more consequentialist or Moorean picture fits the teleological view I lay out here. Incidentally, the normative shape of deliberation will become pertinent later in this paper as I address Raz's view of reasons for action. Raz is not a deliberative maximizer, whereas Scheffler and Pettit are. Velleman (2000) argues against maximizing as a deliberative strategy. Broome addresses this in Broome (2013). Dancy engages the question in Dancy (2004).

The value of the military means that there is reason to-enlist-at-your-local-Navy-center-given X,Y,Z-circumstance and there is reason to-support-legislation-that-supports-the-troops-if-is-in-your-power. Both the Commander in Chief and I have both of these reasons, but we are not in the circumstances to act on both. (Similarly, we could say that I have reason to ease the suffering of my fellow man where I can, but I am not always in the circumstances that trigger this reason.) Thus, every value leads to all the reasons for the actions that would be the appropriate responses to that value.

This pushes the question to the agent's perspective – how can the teleologist capture when an agent “has” a reason. Arguably, this is the most important aspect of reasons, for they are what point agents toward action, and thus work via an agent appreciating that they are the case, through an agent taking them to apply to her. That agents take reasons to point to performing an action – that they suggest that they, *the agents*, perform an action, is a key part of what it means to be a reason. The way that the teleologist makes sense of the difference in reasons between two agents in different relations to value elides the difference in the position these agents are in regarding reasons for action. We can capture this difference by attending to the practical perspective, or the perspective of deliberation.

The teleologist needs to posit some subjective condition to mediate the value of an object and how an agent should interact with it.¹⁷ So reasons will still depend on the value of the object, but different agents have different kinds of interaction that are warranted depending on their circumstances. Teleologists can attempt to make what an agent *values*

¹⁷ For instance, Scanlon discusses the role of subjective conditions in determining reasons for his buck-passing view in Scanlon (1998). In many ways, the buck-passer is the other side of the coin to the teleologist. While the teleologist considers reasons to be able to be “read off” of what is valuable, buck-passers consider value to be able to be “read off” of what there is reason to do. See Heuer (2004).

play this role. This would preserve something evaluative being the thing determining the reasons in play, so has seemed to some teleologist to be a good save, for instance to Joseph Raz.

Consider an example. Because I value Russian literature, and this relationship is appropriate, the *attachment* is valuable (Raz 2001).¹⁸ Now all agents have reason to respect this valuable attachment. It would be wrong for people to disrespect my valuing Russian literature, perhaps by ruining my books or distracting me while I read. These are reasons shared by all agents in virtue of the value of the attachment I have to Russian literature. My valuing Russian literature also gives me special reasons that only I have, however. I have reasons to read Dostoyevsky and Bulgakov, for example, or to keep up on scholarship pertaining to the academic field. Raz's attempt to accommodate agential partiality by pointing to the value of the valuing relationship does not answer the question of partiality. It points to another impartial value that is created by legitimate partial relationships of valuing. My partiality to Russian literature gives everyone reasons to respect this relationship. This does not explain my special reasons, or how they themselves may relate to value.

Some of the reasons involved in my valuing Russian literature are not shared by those who, though they may recognize the value of Russian literature (and even the value of my relationship to it), do not value Russian literature themselves. There is some

¹⁸ (when they are appropriate in the sense that the object of attachment is in fact valuable and that the role the agent has put it to in her life is appropriate)

subjective condition that explains my reasons with regards to the literature that is not a function of the value of it that a teleologist cannot account for.

So how can a teleologist explain that I have reasons regarding Russian literature when another agent doesn't? As long as no one disrespects the value in question, no one goes wrong, but there are still reasons on the table.

One response would be to point to the distinction between wrong-making reasons and reasons that render intelligible.¹⁹ For it is only the reasons of respecting value that an agent can go wrong when acting against. The reasons of engagement – the reasons that I have to value Russian literature given my circumstances – are of a sort that I would be foolish to not follow but not wrong. Such reasons may not be the sort that teleologists mean to address when they say that reasons are a function of value.

Something other than these kinds of reasons are at work when it comes to agents valuing things, and accounting for when such valuing are legitimate or appropriate. These are normative situations that aren't explained by the teleologist's story, so at the very least they outline the limit of the relationship the teleologist attempts to account for the relationship between reasons and value with. More things need to be explained in terms of reasons relationships than value can accommodate. When Raz acknowledges that some valuing are appropriate and some not, that some agents are foolish to not value while others are in the clear, he is acknowledging that there is normativity going on beyond what the teleological framework can accommodate.

¹⁹ See, for instance, Dancy (2004) Chang (2005). MacNamara (Manuscript).

Another way of putting this concern is in terms of an idea that Raz discusses throughout his work. A core commitment he has is to what he labels the “basic belief”, a commitment that we naturally have about our agency: that usually when we consider what we have reason to do we are faced with multiple options, and that we may act in more than one way without going wrong. In other words, what the teleologist calls “reasons” often under-determine what an agent must do – agents can act according to some reason rather than another without going wrong. This raises the question that, if reasons are not determining the thing to do, what will? If, when faced with the relevant reasons for a particular case, the agent has genuine options regarding what to do, how will she proceed to *choose* to act?

According to the teleologist, the reasons on the table are a function of the values in the case. They often leave the agent with options of how to behave, and then subjective conditions and circumstances of the particular agent will make some options more intelligible, or more or less foolish or silly, than others, but not more or less wrong insofar as they truly are options. The teleologist cannot call the grounds on which the agent chooses from the relevant options ‘reasons’, for they will not be connected to the value of the options in the right way. These are merely the subjective conditions that explain what the agent did.

Now, however, the teleologist is left piecing together an important part of our agency – the process by which we choose to act and the grounds by which we do so. There are two issues here. First, these subjective conditions are more than merely explanatory, they have normative weight in that they make it foolish, silly, or in some ways criticizeable

for some agents to behave in ways that are fine for others. Because it falls short of the kind of “wrong” that is at stake in some circumstances doesn’t make it non-normative, and doesn’t make these considerations any less reasons that can ground the action of these agents. Second, by having the reasons that are derived from value set up the options of deliberation, and then contextual relevance go to work afterwards, the teleologist has delineated a break in the deliberative framework that is counterintuitive. The view has one thing explain why an action is an option and another thing explain why it would be intelligible to choose one option rather than another.

The subjective conditions that explain why it makes sense for me to choose one option while it makes sense for you to choose another is separated completely from what made those actions options in the first place on this teleologist view. Why? To preserve the primacy of value in reasons-talk. For the teleologist, the reasons by which we could go wrong are determined by what is valuable, and these reasons frequently under-determine the case. So we have a set of options that what was valuable set, and the teleologist is out of explanation for how we choose what to do from there.

However, the better picture of what is going on in deliberation takes into consideration the relevance to the agent, rather than starting with what is valuable. What makes these actions options in the first place for agents is that they are relevant to the context of action. This is what the buck-passer will point to – the action’s pertinence to the context and our practical engagement with the world is all that their value will amount to. So the line that separates showing up as an option (value-reasons supporting) and rendering

some options more and less intelligible (subjective conditions, non-wrong-making supporting) naturally becomes blurred.

Conclusion

Thus, there is reason to be hesitant to adopt the standard view, if not actively adopt an alternative. The reasons that the standard view appears natural do not necessarily commit one to the standard view, and there are cases that prove problematic for the view. Further, the standard view has difficulty accounting for some basic intuitions we have regarding deliberation. For the remainder of my dissertation, I will focus on providing an alternative.

Chapter Three: Towards a Genuine Alternative

In this chapter, I begin to lay out an alternative approach to analyzing what is involved in taking a consideration to be a reason for action. I proceed as follows. In section one I introduce the explanation-based view. An explanation-based model of taking a consideration to be a reason for action holds that when an agent takes a feature to count in favor of performing an action, she takes that feature to make sense of performing the action. This contrasts with the standard view, which holds that agents take reasons to count in favor by taking the action to be *good* to do in some way. Thus this alternative replaces the good in the standard view with making sense, or rendering intelligible, which are the work of explanation. I argue that this explanation-based alternative underlies both Setiya's and Velleman's rejections of the guise of the good.

In section two I lay out Setiya and Velleman's underdeveloped versions of the explanation-based view accounting for the relationship between explanation and justification. In section three, I suggest that, though both versions don't get this relationship quite right, the way in which they fail illuminates a way forward for the explanation view. I conclude in section four with my view of this way forward, articulating desiderata for my view to meet in later chapters.

Section One: The Explanation-Based View

I have urged in the previous chapter that the guise of the good thesis is too quick in unpacking the counting in favor relations in terms of goodness: the intuitive appeal of the standard view does not lead directly to adopting it and we should be open to alternatives.

But if we assume this is correct, then we would still need a viable alternative unpacking of this relation. How can we make sense of taking a consideration to count in favor of acting if *not* in virtue of tracking value?

Setiya's and Velleman's work is explicitly aimed at rejecting the guise of the good thesis. Importantly, though, there is implicit in their accounts the beginning of an alternative account of what is involved in taking a consideration to speak in favor of performing an action. I call this the explanation view.

The explanation view is that agents engage with reasons for actions in an explanatory mode: reasons *explain* performing actions. Explanations of events or actions *make sense* of why they occurred, and so seeing reasons as explanatory is seeing them *make sense* of performing an action. The key to the strategy is that taking a consideration to make sense of performing an action does not necessarily involve taking the action to be valuable to perform. Thus the main tenets of the view are:

1. to take a consideration to be a reason for action is just to take it to make sense of performing the action, and
2. taking a consideration to make sense of performing the action does not commit an agent to any evaluative view on that action.

What is novel about the explanation view is that it claims that this explanatory work is the *only necessary* supporting relation from the agent's perspective – so when an agent takes a consideration to be a reason for ϕ -ing, she is taking it to explain ϕ -ing.

I suggest that both Setiya's and Velleman's views can be seen as explanation-based views.

Setiya, in order to reject the guise of the good thesis, focuses on what is necessary to the agent's perspective when acting intentionally. He uses Anscombe's "Why?" question to articulate the agent's reasons for acting. In *Intention*, Anscombe uses the agent's ability to answer this question about why she so acted to point to the direct knowledge we have regarding the reasons as causes of our behavior (Anscombe, 2000, 17). Setiya says: "The answer to Anscombe's question is an explanation of what one is doing and why one is doing it" (Setiya, 2010, 89). Because the explanation that is an answer to Anscombe's "Why?" question needn't involve an evaluative assessment, Setiya claims the guise of the good thesis is false.

On Velleman's view, agents engage with reasons as a result of a desire to make sense of ourselves as agents, and so actions will make sense to perform in virtue of how they fit together with our understanding of the world and ourselves as agents (Velleman, 2009, 13-16).¹ Velleman thus characterizes the agent taking there to be a reasons for action as taking there to be a feature that highlights the way in which behaving in such a way is intelligible, or makes sense to the agent.

Velleman and Setiya are the best examples of those who reject the guise of the good thesis by focusing on the explanatory role of practical reason. However, I suggest that those who share the concerns raised in the previous chapter regarding the standard view's potential to account for the breadth of human action should find the explanation-based view to be a compelling alternative.

¹ See also Velleman (2005) and Velleman (2000c) for Velleman's repeated references to the agent making sense of herself through action.

The debate over unpacking the supporting relation deals with the *necessary* elements involved in taking a consideration to count in favor of performing an action, so it makes sense to be cautious of invoking any unnecessary commitments on the part of the agent.

The standard view claims that certain commitments, namely positive evaluations, are necessarily involved in seeing a consideration supporting performing an action. In the previous chapter, we saw that it seems that we take reasons for a broader range of actions than those we have positive evaluation towards. The guise of the good thesis, while intuitive in some paradigm cases, fails to do justice in cases like those of the self-destructive agent. It is important to account for a broad range of human action, including actions that are not represented by the traditional paradigm cases. Evil and self-destructive agents are agents too, and we should be wary of characterizing their flaws as flaws of reasoning rather than deviations from evaluative norms.

The basic idea involved in taking a consideration to be a reason for action is taking a consideration to support performing an action, and an agent can do that while being perverse, conflicted, wicked, immoral, evil, and evaluatively twisted in a number of interesting ways. I suggest that this is so because our rational and evaluative capacities, though of course intimately connected, are different things. The guise of the good thesis does not appreciate this fact.

Reasons are the material of explanation (both within and outside of the practical domain) and our evaluative commitments are the material we use to determine what is good. Our capacities to deal with these things can come apart. We see things as good

besides potential actions, and have evaluative commitments that may never lead us to see reasons for action. Also, reasons are primarily the material of explanations. We should look to what reasons are doing: putting behaviors into a context where performing the action would make sense.² When we provide an explanation of why an event or action occurred, a reason is what renders it intelligible.³

The reason we take to favor an action is what makes sense of performing the action. So what is seeing a consideration supporting action on this proposed alternative to the guise of the good thesis? Explaining why we would perform the action.

I find this view compelling but as it stands, it is woefully underdeveloped.

First, we can note that by focusing on explanation, this view stands to replace our understanding of the “supporting” relation from rejecting the guise of the good thesis with “making sense”: instead of taking a consideration to highlighting the action as good to do, the agent takes the consideration to show the action as making sense to do. This is the first step towards producing an alternative to the standard view, but we need to go further. We need a fuller understanding of what it means to take a reason to “make sense of” an action. Currently, the proponents of the explanation-based view are making do with platitudinous phrases such as “making sense” and “rendering intelligible”. Without more of an idea of what this means, the alternative is incomplete. I propose we mine the vast philosophy of

² We can see this as naturally linked to the justificatory intuition.

³ This could be how we understand reasons in the theoretical domain as well, though this is more controversial a commitment. When we provide a reason for a conclusion, we are providing an explanation for the conclusion, either in the sense of why we reached it, why one would, or why one should. These are different, surely, but we could understand each as a kind of explanation.

science literature for a theory of explanation that can provide standards for what it means for a consideration to lend explanatory support for performing an action.⁴

Second, as I explained above, one of the key virtues of the guise of the good thesis is that it seamlessly explains how seeing the support of a reason in its guiding role relates to the justificatory role. The explanation account needs to be able to do the same. In other words, we need to see how the explanation view accounts for the way that agents see reasons as the material we offer in defense of our behaviors. There is an intuitively tight connection between seeing a reason for action and seeing the action as justified. How does seeing that a consideration lends explanatory support relate to seeing the consideration lend justificatory support?

While Setiya and Velleman do not offer progress on the first point,⁵ in each of their views they implicate a relation between explanation and justification. Because of this, it will be fruitful to take a chapter to consider the possible ways that these explanation-based views have dealt with this burden. In this chapter, I address this aspect of their accounts in order to discuss the prospects of the explanation-based view for meeting this desideratum laid out above. A genuine alternative to the standard view will have to make sense of how reasons justify, and an explanation-based view of practical reasons must illuminate how the fundamentally explanatory role of reasons can accommodate these roles.

⁴ I discuss this further in the upcoming chapters.

⁵ Though neither elucidate what it means to explain, Velleman does expand on information he takes to contribute to explanations in Velleman (2003). There, he intimates that what he takes explanatory force to do, basically, is “convey understanding of some kind or another” (Velleman, 2003, 5). I take it that we can do better than this. That there is a literature in the philosophy of science devoted to explanation also is reason to be optimistic.

Section Two: How Explanation Relates to Justification

According to the explanation strategy, taking a consideration to be a reason for action is just to see why the action makes sense to perform. To be successful it will need to make sense of the common intuition that taking there to be a reason to perform an action involves taking there to be justificatory support for the action – this involves considerations that say why you *should* perform the action, and why it would be proper to offer in defense of the action. Yet it seems there are cases when we are able to use a reason to explain why we performed an action without taking that reason to add to the justification for performing the action. For example, imagine I have a moth phobia, and this leads me to perform some action. If I take the presence of a moth to be a reason to run from the room, then according to the explanation strategy, the moth's presence supports running from the room: it makes sense of the behavior. We can see the explanatory support most clearly when we note that the moth's presence successfully explained my action. However, in taking the moth's presence as a reason to run from the room, did I take it to justify running from the room? That seems to be a stretch, for it strikes most that running from the room in the face of a moth is unjustified behavior.

Must proponents of the explanation-based view, therefore, reject the common intuition that taking a reason to support an action involves taking there to be some justificatory support for performing the action? Most think that when we take there to be a reason for action, this involves some justificatory support: we take that reason to support performing the action in a way that we can offer it in defense against criticism. In short, when we take a consideration to be a reason, we take it to be a justificatory reason. If the

explanation-based view were to reject the notion that taking considerations to be reasons for actions always involves taking there to be justification, then taking there to be justificatory support would be something beyond simply taking there to be reasons in support of an action. On the other hand, the explanation-based view could have a similar structure to the guise of the good thesis, and claim that the way that agents take there to be reasons in support of actions (they make sense of performing them) is how agents take them to contribute to the justification in performing them.

To hold the explanation-based view, the crucial claim is that the fundamental support relation involved in taking a consideration to be a reason is explanatory and is not under the guise of the good. There are two versions that are consistent with this commitment, however. First, one could take this fundamental support relation to be distinct from the support we take to be involved in justification, which may be under the guise of the good. Second, we could take the support involved in justification to be the same that is implicated in taking a consideration to be a reason, no part of which is under the guise of the good. In this version, value is simply one among other considerations that may support performing the action, while explanation is the fundamental supporting relation. The second version of the strategy aligns the explanatory and justificatory aspects of reason, while the second makes a sharp distinction between the two.

The two forms of the explanatory strategy achieve the same end – separating what is essentially involved in an agent taking there to be a reason for an action from taking that action to be valuable – by using explanation to account for the supporting relation involved in taking a consideration to be a reason. In order to make making sense the fundamental

feature in seeing a consideration as a reason, these versions commit themselves to different views on the relationship between explanation and justification.

I take it that Setiya and Velleman adopt the different versions of the explanation strategy.

Setiya rejects the intuition that when we take there to be a reason for action, we take it to justify our action (to some degree). According to his view, we can take a consideration to support performing an action, making sense of performing it, without taking it to contribute to the justification in doing it.

For Velleman, to take a consideration as a reason is to take it to make sense of performing the action. His view is, like the standard view, in line with the common intuition about the tight connection between our view of reasons and justification: when we take a reason to support performing an action, we take it that the reason lends justification to performing the action. For Velleman, this is by virtue of making sense of the action: we take considerations to count in favor of and contribute to the justification in performing an action when we take them to make sense of performing it.

I'll outline each in turn here.

1. Version One: Two Support Relations

An agent can take there to be a reason for action without taking it contribute to the justification for the action, according to Setiya. Setiya uses the Anscombian commitment that agents have non-observational knowledge of their reasons for action to try to separate taking there to be a reason and taking there to be justificatory support. At the root of

Setiya's position is the idea that an agent can believe that she is ϕ -ing because P without believing that P justifies ϕ -ing. So, P being the reason for ϕ -ing has a different supporting relation than P justifying ϕ -ing.

Setiya uses the "Why?" question to isolate the necessary features of acting for a reason from those that the guise of the good involves. Because believing you have explanatory reason to ϕ is sufficient for answering the "Why?" question, you can believe you have explanatory reason to ϕ while not believing you have justificatory reason to ϕ . Setiya concludes that an agent can believe that she is acting for a reason while not believing she is acting for a justificatory reason.

To see how this can work, consider an example from "Sympathy for the Devil": Setiya provides a case of a house fire, where a family man runs out of the house, abandoning his wife and children (Setiya, 2010, 90-91).⁶ The man takes the danger of the fire to do nothing to justify his behavior, but yet to have explanatory power and be the reason for which he acted. Thus the danger of the fire is the man's reason for running out of the house and the man does not take it to justify running out of the house at all. Because it is consistent that the danger of the fire is the man's reason for running out of the house, and that it doesn't justify running out of the house, Setiya takes the case to establish the possibility that an agent can take the danger to support his running out of the house without seeing it as justifying running out of the house.

⁶ In the text, Setiya relies on the controversial "silencing" doctrine in order to get the intuition that the man's behavior has no justification whatsoever. Intuitions about the actual intelligibility and justification involved in this case will, of course, vary but my hope in including it is to illuminate the structure that Setiya has in mind for his argument here against the guise of the good thesis.

Setiya's view preserves an aspect of the guise of the good thesis: when considerations contribute to the *justification* in performing an action, they highlight the good in performing the action. Justification is a matter of value on this account, but is not an essential part of seeing a consideration supporting performing an action.

2. Version Two: One Support Relation

Velleman is in line with the standard view in a different way: when an agent takes a consideration to count in favor of an action, he thinks the agent takes the action to be justified, thereby preserving the intuitive connection between taking there to be a reason and taking there to be justification. However, Velleman doesn't link justification to value, or tie taking an action as justified to taking it to be good. He follows the common conception of practical reasoning in claiming that taking yourself to have a reason is taking there to be justification for so acting.⁷ He says:

A reason for action is something that warrants or justifies behavior. In order to serve as the basis for a subject's behavior, it must justify that behavior to the subject – that is, in his eyes – and it must thereby engage some rational disposition of his to do what's justified, to behave in accordance with justifications (Velleman, 2000, 9).

⁷ This is mainstream in the sense that in most articles about reasons it is backgrounded that insofar as an agent is deliberating about what to do, she is either constitutively guided by the standard of what she *should* do. You can see this is the norm by observing the discussion of reasons is simply normally in terms of what the agent "should" do, "ought" to do, or would need stronger reason in order to get out of being "required" to do. See: Broome (2013), Korsgaard (2009), Scanlon (1998) for some telling examples.

Taking a consideration to be a reason for action will be necessarily tied to defensibility and justification for Velleman. However, this doesn't lead him straightforwardly to the guise of the good, as it would with views that link justification with being worthy of pursuit and valuable.

Velleman treats the relation between justification and explanation as especially tight, with making sense and justification being paraphrases of one another. In this case, justification isn't a special variety of making sense, but rather another way of articulating what is going on when something is making sense to some extent – it is to some extent justified. This interpretation is supported by the places where Velleman paraphrases making sense as justification and vice versa (“I cannot act for reasons if I don't care about doing what's justified or (as I would prefer to put it) what makes sense” (Velleman, 2000, 121)), as well as by the fact that he claims that the relevant support behind seeing reason or actions is making sense *as well as* being justified:

Rational agents have a desire to do what makes sense, or what's intelligible to them, in the sense that they could explain it. (Velleman 2000, 141)

A reason for action...engage[s] some rational disposition of his to do what's justified, to behave in accordance with justifications. (Velleman, 2000, 9)

When it comes to seeing reasons for action, Velleman's view seems to be that justification *is* the sense of intelligibility in question. While taking a consideration to be a reason **does** lead an agent to taking the supported action to be justified, it does **not** lead her to taking it to be valuable. Thus, Velleman preserves the common intuition we have about the

relation of taking there to be reasons for action and justification, but this may be at the cost of being too revisionary regarding what it means to see an action as justified. For Velleman, seeing an action as justified is reduced to seeing the action as intelligible to do, and doesn't appeal to value.

This may seem to be a revisionist view of justification - the most common understanding of justification is that it is tied to value. However, it fits with a certain understanding of what it means to justify. When we justify behavior, we defend an action against criticism, saying why it is something that should have been done. One way of understanding this is that in defending behavior we are saying why it makes sense to have performed it. There is substantive disagreement here in the views of justification, but perhaps also some intuitive support for both.

The different issues faced by the two forms of the explanation-based view's answer to the relationship between explanatory support and justificatory support will point to more explicit standards for a successful version of the view. I turn now to drawing these out.

Section Three: Problems with the Available Versions

The ways that Setiya and Velleman relate seeing a reason to seeing that consideration contribute to justification each lead to problems.

1. Version One: Two Support Relations

By allowing for the possibility of taking a consideration as a reason for action without considering the action to be justified, Setiya isolates the basic supporting relation from seeing actions as valuable.

Setiya presents his position as grounded in examples of agency, but I suggest that these very examples may point to a problem with this view.

Consider the following set of cases:

Pete the Puritan – Pete thinks sexual desire is the work of the devil and considers any pursuit of such behavior to be unjustified. However, he does experience the average spectrum of human urges and sees some considerations (the attractiveness or availability of some potential partner, say, or how good he has been in depriving himself up to this point) as reasons to indulge in these desires. However, though he sees these considerations as reason to indulge, but he does not believe them to justify indulging (Case from Watson, 1975, 210).

Academic Al – Al takes great pleasure in the life of a philosopher, constructing arguments about abstract positions and engaging in debates with other sharp minds of his day. It gives him a sense of power he very much enjoys, though he sees the life of a philosopher and the pursuits that make it up to be empty of meaning and not worthwhile in the least. He sees considerations as reasons to engage in

philosophy while at the same time, he doesn't believe these considerations do anything to justify participating in the activities of philosophers. (Setiya, 2010, 76)⁸

Setiya the Smoker – Setiya, though he has smoked for years up to this point, considers smoking to be completely unjustified. He is going to quit at midnight tonight, and while he sees this as a reason to smoke a pack of cigarettes before then, he does not take this consideration (i.e., that he is quitting at midnight) to do anything to justify smoking the pack of cigarettes. (Setiya, 2007, 37)

In each of these cases, the agent is identifying some consideration to be a reason for an action in the sense that it makes sense of performing the action. However, according to Setiya, in each case the agent considers the reason to fail to add any justification to so acting.

When we take a consideration to be a reason for action, we are taking the action to make sense to perform; we take the consideration to render it intelligible in some way. For Setiya's picture to get off the ground, a consideration must be able to render an action intelligible without the agent thinking that she is in any better position to defend her behavior on the basis of that consideration. From the agent's perspective, the intelligibility of ϕ -ing will do nothing for how justified she is in ϕ -ing. This is crucial – that it do nothing at all to change her position with regards to defensibility or warrant. If the agent takes the intelligibility to contribute to the justification of ϕ -ing, then she thereby takes there to be

⁸ Adapting Burnyeat (1980)

something worthwhile in her action according to Setiya's view, and the behavior will be under the guise of the good.

According to Setiya's view, though the agent can take the reason to rendering her behavior intelligible, she can take the consideration to be justificatorily inert. The justification for the action does not change when the consideration is added to the scenario.

How plausible is it that in the above examples the agents don't take themselves to have *anything at all* to say in their defense with regards to the suggested behaviors?⁹ Admittedly, for each it is clear that they consider their behavior to be *overall* without worth, and *all things considered* the wrong thing to do, and therefore without a justification that would withstand any sustained reflection. But does this mean that there is no added justification for the action from the agent's perspective once he sees there to be an explanatory reason to perform the action?

If the agent takes there to be a consideration that renders an action intelligible, she seems also to be able to offer that consideration in defense of that action. For Pete, say he took the attractiveness/availability of a potential partner as a consideration in favor of ϕ -ing. If he considered whether he had *any* defense for ϕ -ing, something that he grants is unjustified, would this consideration not alter the weight of justification at least somewhat? It seems unlikely. Similarly it is doubtful that if Pete takes the attractiveness/availability of the partner to make sense of his behavior it would not have any effect on the behavior's warrant. Consider also the case of Al, pressed for justification for his philosophical

⁹Keep in mind pragmatic considerations that may come into play regarding social thresholds for when we would cite relevant reasons. See: Schroeder (2007) and Schroeder (2008) for pragmatic considerations on citing reasons for our reason-citing behavior in ordinary speech.

behavior. It seems intuitive that he would have some response to questions along these lines, even though he believes his behavior to not have worth overall. Finally, in Setiya's smoking case, wouldn't we say that the mere intelligibility itself of the behavior may do something to alter the status of the action with regards to how defensible or justified the behavior was?

Setiya's strategy in rejecting the guise of the good is to isolate the explanatory support from the support of reason which is connected to seeing your action as valuable. I have suggested a potential worry with this strategy – that it seems unlikely an agent will see an explanatory reason to act in a particular way without also seeing that reason as altering the action's defensibility.

2. Version Two: One Support Relation

Velleman's view may seem to fare better, but I suggest it runs into different problems. Because it isn't trying to allow for intelligibility without justification within the agent's perspective Velleman's version of the explanatory strategy has a different structure than Setiya's. It therefore doesn't have the potential problems with regards to the examples of Pete the Puritan and company from above. However, by making justification *reduce* to intelligibility in his example and in the quoted text, Velleman suggests a picture committed to a strong claim in another direction: that these forces aren't independent at all from the practical perspective. This may have its own troubles with the cases.

If justification is just making sense, and this is the single fundamental way reasons support performing actions as Velleman says, then Setiya's three examples, the

considerations that count in favor of the relevant actions must lend just as much justification as intelligibility. If justification just is intelligibility, then an agent should take her behavior to be justified to the extent that it is intelligible. It is hard to see that this is the case, especially in the circumstances described above.

For example, the fact that philosophical activities are pleasurable to Al and give him a sense of power make it very intelligible that he partakes in them, but it is difficult to see how he would be justified *to the degree* this makes sense. Similarly with the case of Setiya the akratic/rationalizing smoker. His smoking a pack of cigarettes tonight is intelligible given the fact that he is quitting at midnight, but the explanatory support here seems out of sync with the justification of the behavior – it makes more sense than the amount he takes himself to be justified. This is problematic for a view where justification just is intelligibility, for it is unclear what could account for the difference in degree here.

For Setiya, the problem was that along with all the intelligibility that the pertinent consideration lends the relevant action, there may be a smidgeon of justification as well. For Velleman, it may be a problem that it is just a smidgeon, in comparison to all that explanatory support. These issues turn on the relationship between justification and explanation. I take it the issues are not insurmountable.

Section Four: A Way Forward

I began by suggesting that implicit in the literature is suggestion of an alternative. Implicit in the work of Setiya and Velleman, central objectors to the guise of the good, is the suggestion of what I am calling the explanation-based view of practical reason. This view,

though to my mind promising, has thus far been too underdeveloped to live up to the criteria for an alternative to the standard view. I proposed that there are two broad ways this has been the case.

First, in order for the explanation-based view to be a genuine alternative to the guise of the good thesis, we need it to develop past a thin or rudimentary understanding of “making sense”. To meet this challenge, we can bring in theories of explanation and standards of explanation from the philosophy of science literature. For over a century, theories of explanation in the philosophy of science have explored what features make better and worse explanations. In the scientific explanation literature, theorists have benefited from decades of engagement with examples from science and scientific theory. This has resulted in precisified standards for what makes a good explanation, and using the work done in the scientific domain will help us develop a more robust view of explanation in the agential domain. Mining this literature can help illuminate what it is that agents see when they appreciate that a consideration helps explain why one would perform an action.

To meet the second challenge facing the explanation strategy, we need to be clear about how taking there to be explanatory support for performing an action relates to taking there to be justificatory support. From the discussion of the two views on the table, I suggest two parameters for going forward. The first version of the explanation view separated the support essential to taking a consideration as a reason to perform an action and the support involved in taking there to be justification in performing that action. The problems with this view led us to see that we cannot understand these two forms of support to be completely independent: take a consideration as a reason is linked to taking there to be

justification for the action. The second version took this commitment to the extreme, claiming that when we take a consideration as a reason, it is just to see that reason contributing justificatory support for performing the action. On this version, there is no way we can take justificatory support to differ from explanatory support: justificatory support reduces to explanatory support. The problems with this view highlighted the difference between taking considerations to make sense of actions and taking those consideration to justify. Thus, going forward, we need a view of seeing support for performing an action that links making sense of performing the action with justifying performing the action while preserving the possibility that the amount of justification can vary from the amount of intelligibility.

The two broad desiderata are related, for it will be through developing an understanding of explanatory support for actions that we can further illuminate the relationship between taking a consideration as a reason for action and taking there to be justificatory support for the action. We saw here that taking there to be justification is not reducible to, nor entirely separable from, taking there to be reasons in support of the action. In getting into a theory of explanation from the philosophy of science literature and developing what it means to make sense of performing an action, we will be able to see what relationship would be appropriate here. Answering these challenges are not entirely separable projects. Responding to the first will inform the second.

Chapter 4: Explanation

Science offers impressive explanations of a vast array of empirical phenomena. Examples of the phenomena to be explained include the sky's being blue, the tendency of unimpeded objects to fall to the ground, and the fact that there is a correlation between animals with hearts and animals with kidneys. A theory of explanation should provide an account of what makes a collection of information a good explanation of some phenomenon under investigation while other collections of information are either bad explanations or even not explanations at all.

We have an intuitive grasp of what makes something a good explanation. When we explain what makes the sky blue, we intuitively accept that information about how light disperses, features of blue light, and characteristics of our atmosphere would strengthen an explanation whereas information about the height of the Eiffel tower, the number of knuckles on the foot of the panther chameleon, or the date the last stegosaurus died do not explain the sky's blueness.

In this chapter, I will outline standards for an adequate theory of explanation and review standard accounts before outlining van Fraassen's pragmatic account. Theories of explanation provide standards for good and bad explanations, which are crucial for my project. With standards of good and bad explanation in hand, I suggest we can provide a viable alternative to the standard view of practical reason, and I will turn to this issue in the following chapter.

Bas van Fraassen argues that an explanation is the right answer to a “Why?”-question,¹ and I suggest that besides accommodating the desiderata involved in scientific explanation, this theory is well-suited to account for the way reasons count in favor for reasons for action.²

Section One: Standards for Account of Explanation

In order to succeed as a theory of explanation, there are three features of explanations that an account must accommodate. First, in providing standards for a good explanation, an account should respect the asymmetry of the explanandum and the explanans. Second, an account of explanation should provide some standards for selecting the salient information to include in an explanation. Third, it must account for cases in which an event doesn’t call for an explanation, or when requests for explanations should be rejected. I will fill out these standards in turn.

The first standard (capturing the asymmetry of the explanandum and explanans) is important for framing the task of explanation. In the motivating example above, we seek

¹ Sylvian Bromberger originally suggests that explanations are responses to “Why?”-questions in Bromberger (1966), where an explanation responds to “Why P?”, where P is a statement and Why? is something that turns statements into questions, into requests for explanations for P. From the practical reasons literature, this relates to Pamela Hieronymi’s view (See in Hieronymi (2005) that she treats a reason as a consideration that bears on a question – though through applying the logic of “Why?”-questions in particular, van Fraassen (and I in the practical reasons domain) am going farther than Hieronymi.

² Readers familiar with the intentional action or practical reasoning literature will recognize the use of “Why?”-questions for analyzing action from Anscombe’s *Intention* (Anscombe (1957)). She uses this structure to discuss what it means to say that an action is intentional – it is behavior where the “Why?”-question has particular application; behavior where the agent is aware of what she is doing and the cause of what she is doing non-observationally. Here I am more concerned with how to evaluate answers to the “Why?”-question than the nature of the behavior where the question finds purchase. I hope to highlight a continuity between the intentional cases where the “Why?”-question can apply and the non-intentional application of the question.

an explanation for why the sky is blue. An explanation will provide information that illuminates why this is the case, and the work of these two pieces are importantly different: the explanandum (the blueness of the sky) is the phenomenon that we take ourselves to be missing information about, and the explanans (the information regarding light dispersal, etc.) is the information we are missing. We know that the sky is blue. We do not know what it is that makes this so. An explanation bridges this gap.

A satisfactory theory of explanation will illuminate this connection between explanandum and explanans while preserving the directionality of the way the explanans illuminates the explanandum. Consider some examples. The relationship between the change in reading on a barometer and an oncoming storm is asymmetric in the sense that the oncoming storm explains the reading on the barometer, but the reading on the barometer doesn't explain the storm: if we sought an explanation for the storm, we would typically be seeking information about meteorological events, the time of year and how it relates to storm trends, etc. The reading on the barometer does not explain the oncoming storm, while these meteorological events would.³

A similar structure exists for other classic cases of explanations. Consider how astronomers note the shift in light we're receiving from galaxies is shifting towards the red end of the light spectrum. The receding of galaxies away from us explains the red shift, but the red shift does not explain the receding of galaxies. Thus, if we were seeking an explanation for the red shift, the movement of the galaxies illuminates, or provides

³ "Typically" is doing real work here. The context in which we are seeking an explanation determine what needs to be illuminated, and that is largely what sets up this asymmetry. Relatedly, though the reading on the barometer can be a good reason to think there is an oncoming storm, it does not explain why there is a storm oncoming in the meteorological context. This distinction will come up later.

information that makes sense of, the phenomenon of the red shift. It renders the red shift intelligible. However, if we were seeking an explanation for the receding of galaxies, the red shift doesn't play the same role – it doesn't render this motion intelligible, make sense of it, or illuminate it. The explanandum and explanans are asymmetric in this sense, and this asymmetry generalizes to typical contexts: explanation is directional, the explanans is meant to illuminate the explanandum.

The second standard for an adequate account of explanation is selecting the salient features that will qualify as explanatory from amongst the array of broadly relevant features. For instance, in seeking an explanation for the blueness of the sky, we could cite everything about the physics of light, everything science tells us about our atmosphere, etc. – basically everything such that, if it were different, the sky wouldn't be blue.⁴ A successful account of explanation will be able to differentiate among these factors in order to isolate those that illuminate the sky's blueness, render the blueness of the sky intelligible, etc.

Van Fraassen provides an example that illuminates this standard (van Fraassen, 1980). Why did the Irish elk go extinct? Any number of factors were part of necessary history that led to the extinction: a difference in the elk's speed, changes in the contact area of the hoof, alterations in height, distribution of weight in the body, distribution of its food supply, migration habits, changes in surrounding fauna and flora. If some of these had been different, the elk would have survived.⁵ However, the explanation for the elk's extinction

⁴ This example brings out the difficulty in using accounts of causation that have the structure of *condition sine qua non*: the feature without which the phenomenon would not be. There are an infinite number of features that must be the case in order for the phenomenon to be the case. The question on the table is: Which of these features explain the phenomenon?

⁵ We can articulate this standard in a fashion that isn't specifically related to causality. The features that led to the elk's extinction, or are good grounds for believing in the elk's becoming extinct, or all the laws of

is the selection of males with large antlers, which were cumbersome in their environment. (The rest of the factors weren't screened off, it is pure salience that makes this the explanation, so philosophers have attempted to provide an account of this selection.)

The final standard for an account of explanation to meet will be to have some way of determining when a request for an explanation should be declined. Not all phenomena have explanations, such as fundamental axioms or assumptions in math or scientific theories. Consider Newton's laws of motion: what explanation can be provided for why an object in motion stays in motion? Which questions fall into the category of phenomena or features of the world that are inapt targets for explanation can't be determined a priori, so an account of explanation should be able to accommodate this feature of our theories and understanding. Basic axioms or foundations of theories are not the only examples of such cases either. We should reject a request for an explanation for why this chunk of uranium depleted now rather than 5 minutes from now: our theories are not built in such a way that this phenomenon has an explanation. Some features of our world don't get explained and instead the request for an explanation should be rejected.

Thus, there are three accommodations that an account of explanation must make. First, a successful account of explanation will accommodate the asymmetric relationship between the explanandum and the explanans. Second, a successful account will accommodate the selectivity of explanation – in order to be a successful explanation,

nature implicated in the extinction, each over determine the explanation for the elk's extinction. Because each over determine the explanation, some selection of the explanatorily salient amongst them is required.

salient factors must be selected. Finally, an account of explanation must distinguish between cases where explanations are apt and inapt.

In the following section, I will lay out the broad approaches to articulate standards for explanation before turning to the approach I take to be most successful and well suited to application to the practical domain in Section three.

Section Two: Accounts of Explanation

In providing an account of explanation, theorists are articulating standards for when information succeeds at explaining a phenomenon or event. In this section I briefly outline major approaches to identifying a standard for when a factor explains. Each approach attempts to characterize the principle by which information or features render phenomena intelligible, with a focus on the scientific domain. This means that in providing a standard for explanation, theorists have been articulating what makes one scientific explanation (theory) do a better job of explaining a phenomenon than another.

Consider a central approach to explanation, the “deductive-nomological” approach. Carl Hempel laid out the standards for the deductive nomological view in the mid-20th century, focusing on the following structure: “the explanandum must be a logical consequence of the explanans” (the “deductive” part of this approach) and the explanans must contain at least one “law of nature” (hence the “nomological”) (Hempel, 1965, 248). These standards allow us to see why some generalizations can be used to *explain* (because they are generalizations that capture laws) while others do not explain.

For instance, Hempel uses the example of two stipulatively true generalizations to bring out how these connections can be explanatory. There is, according to this approach, a crucial difference in the explanatory power of the following generalizations: “All members of the Greensbury School Board in 1964 are bald” and “All gases expand when heated under constant pressure”. The generalization about gases captures a law of nature, whereas the generalization about the bald school board members does not. Therefore, according to this model of explanation, the gas generalization can be used with information that a particular gas has been heated under constant pressure to conclude (deductively) that it will expand, resulting in a successful explanation. On the other hand, the generalization about the school board is not a successful explanation for the person’s baldness, though we can use it with information about a particular person who is a member of the school board to deductively conclude the person is bald. Thus, according to this approach, both the deductive relationship between the explanans and explanandum and the reliance on a law of nature are standards for a successful explanation.

As should be intuitive, a major burden for this approach is to provide standards for what generalizations qualify as laws of nature. What precisely is the difference between the generalization about the behavior of gases that makes it count as a natural law while the generalization about the school board members fails to qualify? There has yet to be consensus on what constitutes law-hood which weakens the above approach.

Further, there are questions regarding whether laws of nature are necessary for successful explanation (Woodward, 2000). Consider the generalizations that are the result of study in biology, psychology, economics, etc. that likely do not, or may not, qualify as

laws of nature: leopards breed in spring, say, or that humans' behavior is impacted by authority figures. The shape of these generalizations differs from the archetypical exception-less natural laws we come across in physics. However, it would be strange to say that they fail to explain. There are a variety of tensions and difficult-to-capture examples for the deductive-nomological approach to explanation along these lines.⁶ This leads to accounts of explanation that diverge from the above model in the strength of the connection between the explanans and explanandum (not deductive) and the type of connection that will replace the laws of nature (no longer nomological).

According to one alternative approach, an explanation is a grouping of statistically relevant factors: the “statistical relevance” model adopted by Wesley Salmon (Salmon, 1971). Instead of relying on generalizations that are exception-less or law-like, this approach allows for statistically predictive generalizations. (This change moves us away from a strictly deductive connection as well.) According to this model, when we look for an explanation for a phenomenon, the factors that are explanatory are those that affect the probability of the phenomenon in a statistically relevant fashion. A fact, A, is statistically relevant to a phenomenon, E, exactly if the probability of E given A is different from the probability of E simpliciter. So, for instance, the probability of a cold sufferer getting well

⁶ In the scientific explanation literature, most tend to adopt pictures in the Humean tradition, where (this is Lewis' articulation) “all there is in the world is a vast mosaic of local matters of particular fact, just one little thing and then another” Lewis (1986). Laws characterize generalizations that survive over time without exceptions, but do not have reality beyond this. Some approaches to what it means to be a law: David Lewis's systems approach Lewis (1983) and David Armstrong's universals approach Armstrong (1983). Other treatments include antirealist views (van Fraassen (1989)) and antireductionist views (Carroll (2008)). The question of what qualifies as a law is exacerbated by the number of philosophical questions and issues that the answer will touch on, such as: (i) whether laws supervene on matters of fact, (ii) the role laws play in the problem of induction, (iii) whether laws involve metaphysical necessity, and (iv) the role of laws in physics and how that contrasts with the role of laws in the special sciences. See Carroll (2016).

given that they took vitamin C is equal to the probability of the cold sufferer getting well given the factor that they didn't take vitamin C: they'd get better regardless and thus taking vitamin C is statistically irrelevant. On the other hand, the probability of a syphilitic patient getting better given that they took penicillin does not equal the probability of a syphilitic patient getting better given that they didn't take penicillin: taking penicillin is a statistically relevant factor.⁷

Recall, however, the first desiderata for accounts of explanation: that accounts accommodate the asymmetry between the explanans and the explanandum. This is difficult for views that focus on statistical relevance. Consider the case of the rising barometric pressure and the oncoming storm. The oncoming storm *explains* the rising barometric pressure but not the reverse. However, approaches focusing on statistical relevance can have difficulty accounting for the asymmetry in cases like this where the explanation involves an indicator such as the barometric pressure. In order to be a good indicator, the indicating feature and the event that they indicate can have symmetric varying predictability. The rise in barometric pressure and the oncoming storm co-vary: when one happens, so does the other. Thus, with a view where the guiding principle is statistical relevance, there isn't a clear reason for the oncoming storm to be the explanans for the rising barometric pressure rather than the reverse, and this could be troubling.

A final approach to accounting for explanation appeals to causation. According to this approach, in order to explain a phenomenon we must look to what factors caused the

⁷ Hempel required that the probability be increased over $\frac{1}{2}$ in order for the statistically relevant factors to succeed at explaining, Salmon didn't.

phenomenon to occur. Salmon outlined such an approach in more recent work than I referenced above.⁸ Central to this view is that explanations trace causally relevant factors. Salmon defined *causal processes* and *causal interactions* as the crucial features of an explanation in order to hone in on how an explanation will show how the event fits into a “causal nexus”.⁹ A causal process is a physical process where a property or feature (called a “mark”) continues through space and time.¹⁰ Commonly cited examples are a baseball moving through the air with the mark of a scuff or a car moving down the road with the mark of a dent in the fender. Causal interactions are points in space and time where two causal processes intersect in a way that alters both. Cars travelling down a road constitute causal processes; when they collide and each take on properties like scratches and dents, this is a causal interaction: for now both have been altered. An explanation for an event, E, will trace the causal properties and interactions that led to E. This locates the event in the causal nexus.

Recall the second desideratum in the above section: a successful account of explanation will be able to pick out the relevant features that are explanatory. The difficulties in picking out the salient feature that explains an event is most stark with approaches that focus on causation. When we seek an explanation for the extinction of the Irish elk, we see that if we altered any number of features, the extinction would not have happened. This suggests that these features caused the extinction, because most accounts

⁸ See, for example, Salmon (1984) and Salmon (1997)

⁹ Salmon (1984) page 9

¹⁰ The purpose of the “mark” is to distinguish genuine causal processes from the changes in the world that track these processes but aren’t themselves the true causal “roots”, as it were. Consider the difference between the car moving through space and the shadow of the car. They track one another, but interfering with the shadow at T1 doesn’t alter the shadow at T2, whereas interfering with the car at T1 does alter the car at T2.

of causation rest on some sort of counterfactual intuition. Any event has any number of environmental sensitivities, not all of which are explanatorily relevant.

These varied approaches to providing standards of explanation have been developed and modified in response to concerns and counterexamples. They represent different motivations and each capture intuitions that drive our attempts to illuminate phenomenon. When we seek explanations, we are trying to make sense of some event. In what follows, I outline a different approach – a pragmatic approach developed by Bas van Fraassen. I consider this approach to be better suited at meeting the desiderata laid out in the previous section and to be well-structured to adapt to the practical domain.

Section Three: Van Fraassen's Account

Van Fraassen has a pragmatic account of explanation, where a correct explanation will be determined by contextual features. Thus, in order to understand his account, we need to understand what he means by context and how this sets standards for a correct explanation.

For van Fraassen, to provide an explanation is to provide a correct answer the question “Why P?”, where P is a statement of an event. So, roughly, an explanation for a broken window will be the correct answer to the question: “Why did the window break?”, and the explanation for the extinction of the dodo bird will be the correct answer to the question: “Why did the dodo bird go extinct?”

Of course, that is van Fraassen's view *most broadly*. To get clear on the theory we need to see what goes in to properly forming and answering a “Why?”-question. To start,

consider the now-famous example of requesting an explanation for Adam eating the apple in the Garden of Eden. We could form the “Why?” question as:

1. “Why did Adam eat the apple?”

As van Fraassen emphasizes, however, this question is under-developed: it doesn’t pick out what *aspect of the event* is in need of illumination.

If I were wondering about why Adam ate the apple, this could be (at least) three various requests:

- 1a. Why was it *Adam* who ate the apple?
- 1b. Why was it the *apple* that Adam ate?
- 1c. Why did Adam *eat* the apple?

This points to the idea that “Why?”-questions are not questions about events full-stop. Events have many aspects and, in asking for explanations of them, we are asking for an explanation of one of their many aspects. For example, in the Adam case, we could be asking about different aspects of Adam’s apple-eating: that it was Adam, that the event involved eating, or that it involved the apple.

But things are actually slightly more complicated. When we are asking for an explanation of an event, we are not just asking for an explanation of an *aspect* of the event, but even more specifically, an aspect of the event *under a description*.¹¹ When we ask why *Adam* ate the apple, we might be asking why Adam *qua person near the apple*, or Adam *qua man*, or Adam *qua Middle Eastern person*, or Adam *qua Bible character* ate the apple.

¹¹As van Fraassen states, “The range of hypotheses about the event which the explanation must ‘weed out’ or ‘cut down’ is not determined solely by the interests of the discussants (legal, mechanical, medical) but also by a range of contrasting alternatives to the event. This contrast class is also determined by context.” (van Fraassen (1988):129)

In each of these cases, Adam is the aspect of the event under scrutiny, but Adam is being considered under different descriptions.

Or again, we could ask “Why?” with the *apple* as the aspect of the event under scrutiny. When we ask for an explanation of why Adam ate the *apple*, we could be asking why Adam ate the apple *qua Eden vegetation*, apple *qua forbidden fruit*, or apple *qua food offering from Eve*. In each case, the inquiry is focused on the apple, but under different descriptions. When we ask for an explanation, we do not ask for an explanation of the entire event. Instead, we ask for an explanation of an aspect of the event under a description.¹²

To see this point even more clearly, imagine two different people requesting an explanation for a house burning down during a storm: a neighborhood representative (who runs neighborhood watches and organizes tree trimmings, etc.) and a housing materials specialist (whose career is focused on determining safe housing plans). The event to be explained is that the house burned down, so both people are seeking a response to the rough “Why?”-question: “Why did this house burn down?” Further, both are seeking an explanation for why it was *this house* that burned down – if we attempted to break down the question as we did in the case of Adam eating the apple, we’d hone in on the same aspect: “Why did *this house* burn down?” (as opposed to “Why did this house *burn down*?”) However, the two parties are seeking answers to different questions, and in articulating the aspect of the event (“this house”) under the correct description we can see this.

¹² It is likely the case that most events have a prime feature that inquirers focus on under different descriptions, and thus are less like the Adam eating the apple example, that had discrete features as well that we can readily understand someone requesting an explanation for. However, this isn’t an important point. Key to this view is that any time we request an explanation, it will be of an event under a description.

For the neighborhood representative, she is concerned with the houses in the neighborhood and how they ended up at the end of the storm. She is thus seeking an explanation of the event *qua house in the neighborhood*. The description of the event the materials specialist keys into is that it was struck by lightning. In seeking an explanation for why this house burned down, she is considering houses that are struck by lightning, thus considering the event *qua house that was struck by lightning*. Both are interested in the same feature of the event – the house – but under different descriptions.

By recognizing that requests for explanations are requests for explanations of *aspects of events under particular descriptions*, we begin to tap into van Fraassen’s main insight. According to his theory, an explanation is the correct answer to a “Why?”-question. However, as we’ve been noting, a request for an explanation doesn’t just ask “Why P?”, it hones in on some aspect of P and considers it under a description. The generic form of a “Why?”-question, such as “Why did Adam eat the apple?”, or “Why did this house burn down?” doesn’t capture the relevant aspect that the inquirer seeks to be illuminated. These generic questions are not yet fully articulated, according to van Fraassen. Specifically, they are not fully articulated because they don’t include what he calls a *contrast class*.

Asking why something happened under a relevant description allows the event to be *contrasted* against a set of relevantly similar cases. According to Van Fraassen a fully articulated “Why?”-question has the following form:

2. Why (is it the case that) P *in contrast to* (some other member of) X?¹³

¹³ The notion that reasons are essentially contrastive is not unique to views like this one (that take reasons to be answers to contrastive “Why?”-questions). Justin Snedgar, for example, has had success in defending such a view – see Snedgar (2017). (A contrastivist view of reasons is also a popular response to the problem of Buridan’s Ass - See also: Chandler (2013)). Some competing accounts: Fitelson (2012). Contrastive

Recall the differentiation involved in our first example of Adam in the Garden. If we are asking why *Adam qua person near the apple* ate the apple, then the contrast class will include all the other persons near the apple. Essentially we would be asking why Adam ate the apple instead of Eve. Or, if we are asking why Adam ate the *apple qua available piece of fruit*, then the contrast would be all the other available pieces of fruit. Essentially we would be asking why Adam ate the apple instead of the nearby pear, peach, or grape. Or again, if we are asking the third question above, we can intuit from the italics that the question asks why Adam decided to eat the apple instead of *doing something else* with it instead of returning it, or juggling it. These questions hone in on different aspects of the event and are thus asking different questions.

Thus, even though the generic question is what we may ask pretty frequently – often, we get by just fine asking “Why did Adam eat the apple?” – the fully articulated question involves the implicit *contrast class* of alternatives that is part of the inquirer’s request for an explanation: “Why is it the case that it was Adam who ate the apple instead of some other member of the class of potential people who could eat the apple?” The contrast class is the set of events that are brought into the question by the “rather than” or the “instead of” clause. It is the class of things that the inquirer is comparing the event to – in this case, other people that didn’t eat the apple.

Note that in the example of the house fire, we can see that there is a tight connection between the description under which we are considering the event (*qua house in the*

Bayesianism. Note: you can think reasons are the answers to “Why?”-questions without thinking that reasons are contrastive, and you can think reasons are contrastive without thinking that they are the answer to “Why?”-questions. I am drawing on this theory of explanation, which is committed to a contrastive view of “Why?”-questions, and I claim this is how reasons work.

neighborhood or *qua house hit by lightning*) and the contrast class that will complete the fully articulated “Why?”-question. The neighborhood representative seeks an explanation for why this house burned down when she has a number of houses in her neighborhood that didn’t burn down. She asks why the house burned down *qua house in the neighborhood*, and the other houses in the neighborhood is the class of concern for her explanation. The house that did burn down is in the neighborhood and is being contrasted from the houses in the neighborhood that are still upright.

Thus, once we grasp the description under which she is seeking an explanation, we see the class she is contrasting this event from: the houses in the neighborhood that didn’t burn down. A fully articulated “Why?”-question includes the contrast class, and in this case, therefore, the fully articulated “Why?”-question for the neighborhood representative’s request for explanation is “Why did this house burn down rather than the rest of the houses in the neighborhood?” Capturing the description under which she is considering the aspect of the event (this house *qua house in the neighborhood*) is tightly connected to the contrast class.

We can see that the contrast class falls out in a similar fashion for the materials specialist. The housing materials specialist is interested in the event *qua house hit by lightning*. Her project regards lightning-struck houses, and this is the class under consideration. Once we know that this is the relevant aspect of the event (that it was hit by lightning), we know that the thing that made the fact that this house burned down interesting is that there are these other houses that didn’t burn down that the materials specialist is comparing it to – the other houses that were struck by lightning. Thus, the

contrast class that will fill in the fully articulated “Why?”-question is houses that were struck by lightning and didn’t burn down.

According to van Fraassen, the aspect of the event that is relevant, and therefore the contrast class, is determined by the *context*, i.e. the interests and background information of the question asker. Thus, it is the *context* that accounts for the difference in the two “Why?”-questions in this case, and the three different “Why?”-questions in the Garden of Eden case. Recall the distinctions between the cases of the neighborhood representative and the materials specialist. For each case, the individual cared about something different regarding the house – that it was part of the neighborhood, or that it was struck by lightning. Also, the inquirers had different background understandings – the neighborhood representative knew there was a storm but hadn’t heard a house had been hit by lightning, whereas the materials specialist wouldn’t have shown up without the background knowledge that the house had been hit.

These are the elements of the context for van Fraassen: interest and background understanding. The inquirers’ *interest* in receiving an explanation and *background understanding* of the event determined the contrast class, and therefore the fully articulated “Why?”-question for each case.

We can see how our competence in answering “Why?”-questions is a matter of grasping the inquirer’s interest and background understanding by considering some cases. Norwood Russell Hanson provides a lovely example that highlights the contextual nature of the salient features of causes and the contextual nature of the explanation of events:

There are as many causes of x as there are explanations of x. Consider how the cause of death might have been set out by a physician as ‘multiple haemorrhage’, by the barrister as ‘negligence on the part of the driver’, by a carriagebuilder as ‘a defect in the brakeblock construction’, by a civic planner as ‘the presence of the tall shrubbery at that turning’. (Hanson, 1958, 54)

The request for an explanation of the death of this individual could be framed in a similar way to “Why did Adam eat the apple?” above. In this case the generic question would be something like, “Why did he die?”, and will have different appropriate answers depending on the context parameters van Fraassen articulated, namely the background understanding paired with interest of the inquirer, which will determine the contrast class.

The physician is interested in a medical explanation, which is still pretty generic – it picks out the relevant part of the event to be explained as one of a *patient’s* death, but still the physician could be interested in the patient’s death in a number of ways. His nurse will likely be able to intuit the interest he has when he is seeking the explanation for “Why did he die?”: Is he interested in why this patient died *qua carriage-accident-victim*? Then the request for an explanation will be the contrast between this patient and the other carriage accident victims they’ve treated. Or perhaps the physician could be interested in why this patient died *qua patients-he-treated-that-day*? Then the relevantly similar cases will be his patient list – and instead of “multiple hemorrhage”, the nurse may say “carriage accident”. Note that if the nurse mistakes the physician’s interest, and says “carriage

accident” when the physician’s interest was the former rather than the later, it could be a humorous or morbid exchange.¹⁴

The way we can so frequently use short-hand generic “Why?”-questions speaks to our ability to understand the context – and more – from the scenarios in which we seek explanations. An important part of this context is the implicit contrast class that our interest in the case determines. For the physician, the interest of patients who have had carriage accidents creates a different contrast for this case than the interest of the patients of the day.

These different interests determine contrast classes, which more fully articulate the “Why?”-question for the physician. If the physician is interested in the patient’s death *qua-carriage-accident-victim*, this would be “Why did this patient die rather than these other victims of carriage accidents?” The correct answer will be the feature of this event (this person’s case) that distinguished it from the other cases of carriage accidents. That this patient experienced multiple hemorrhage (a frequently fatal complication) would be a good explanation for why he died (stipulating that those carriage accident victims who lived didn’t suffer this complication).

Note that the nurse, in providing the answer of “multiple hemorrhage”, is intuiting not just the interest of the physician but also some background understanding on his part. The physician’s understanding of medicine, his patients, and the effects of different

¹⁴There is a great example of failed explanation from the realm of sport: During a post-game interview, college athlete Taurean Prince is asked “Why was Baylor out-rebounded by Yale?” The student replied, “They scored after getting possession of the ball after a failed shot more times than we did.” In this case, the student purposefully fails to take the inquirer’s interests and background understanding seriously, deliberately mistaking the context of the “Why?”-question.

conditions are all implicated in this case, and the nurse's ability to recognize the physician's background beliefs and understanding allow her to identify the implicit contrast class. If the nurse didn't think the physician had some of the relevant facts, she would fill in the generic question differently, or answer it differently.

For example, if a layperson asked "Why did he die?", with an interest in the difference between this patient and other carriage accident victims, different answers may be appropriate depending on the layperson's understanding – does the individual have a working understanding of internal bleeding? If so, mentioning that could do the job. Otherwise, the destruction or dramatic impact of the crash on the body described in brute causal rather than medical terms could be the way to go ("the dash hit him really hard in the chest").

Explanation thus has this basic structure: It is a response to a request for information about why something happened under a relevant description. The correct explanation will be something about the event that distinguishes it from the alternatives in the contrast class, which was determined by the interests of the discussants and the background understanding of the scenario. Having the interest in agents who could have eaten the apple set the aspect of the event to be explained in 1a with our interests in why Adam was the one who ate the apple, and the contrast is other possible people who ate the apple. A good answer here will be what about Adam distinguished him from these other people that made him the one who ate the apple.

The context of the scientific "Why P?" question includes a contrast class of things that appeared to be the same until P happened (all the houses weren't burned down, then a

house burned down). The contrast class is a set of relevantly similar cases, against which the event (under a description) must be distinguished. *Something* is the relevant difference between the houses that didn't burn down and the one that did. So, the responder's job is to identify what relevant difference explains P happening to the now homeless questioner instead of her neighbors, or instead of the rest of lightning victims, in light of the presupposed understanding of the questioner.

Now, we can start to see how we can evaluate rival explanations. What makes one explanation do a better job of capturing the relevant feature than another?

Recall that an explanation is an answer – “Because A”. This answer has the implicit form of “A is the relevant feature of P that made P occur rather than the rest of contrast class X” or “P in contrast to the rest of X because A”.¹⁵ With this structure in hand, we can evaluate explanations on three main parameters: how likely it is that A is true (occurs), how much A leads to the likelihood of P's being true (occurring), and how much A leads to the likelihood of the rest of the contrast class being false (not occurring).¹⁶

With such evaluations in hand, we can compare “Because A”s as answers to the “Why?”-question to the other possible answers.

If A: (i) is likely, (ii) leads to P occurring, and also (iii) leads to the rest of the contrast class not occurring, then A receives the highest marks in this context. These

¹⁵ Or as van Fraassen prefers: A bears R relation to [Pk, X]. ‘k’ here is the relevant background theory and knowledge assumed by the discussants.

¹⁶ Van Fraassen breaks down the parameters of evaluation slightly differently than I do here. He claims there are (at least) three ways of evaluating explanations, including how relevant the answer is to the topic at hand. He also breaks down what I have here as the second criterion into three parts. However, for space considerations and clarity I am focusing on the central evaluations that will be helpful in grasping the strength of an explanation in order to apply the pragmatic account to the practical domain.

evaluations will be relative to the background understanding: knowledge of the discussants and any relevant theory. If A doesn't straightforwardly meet these conditions (i)-(iii), we compare answers based on how well they meet the conditions. A could still be the strongest answer if it favors P against the alternatives (it lowers the probability of the rest of the contrast class instead of implying the falsity of them).

Thus, good explanations are evaluated based on how they redistributed the likelihood of the events in question (the event P, and the contrast class). We compare answers to one another based on how well they perform at the above conditions, but they also could interact with one another. This brings in the notion of "screening off", where for answer A to screen off B: the probability of event, E, given A is the same as the probability of E given A and B: $P(E/A) = P(E/A \& B)$. This blocks B as a good explanation to the event – it didn't add to the probability of E given A. In such a case, A is a stronger answer, even if B did a good job at the (i)-(iii) conditions above.

Further, we could assess *how many of the cases* in the contrast class A lowers the probability of. Recall that there are two measures besides A's empirical likelihood: A's effect on P and A's effect on the rest of the contrast class. If it doesn't do both straightforwardly, it will be a better answer the more it does both of these things. Thus the better an answer A is, the more it will increase the likelihood of P occurring and the more it will decrease the likelihood of the rest of the contrast class.

Consider the first example again, of Adam eating the apple. Interested in why Adam decided on the action that he did, I ask a "Why?"-question articulated with similar cases determining the contrast class, where similar cases are Adam's other behaviors with apples.

(Basically, my interest in this event can paraphrase the inquiry as: I'm asking why Adam did this with the apple instead of the other things he does with apples, given my background understanding.) I ask, "Why did Adam eat the apple rather than (juggle the apple), (throw the apple to the ground), (add it to a collection of apples in a nest he'd be keeping), (simply hold the apple for a while), (hide it safely away from those who may eat it), etc.?" Different reasons fare better and worse at differentiating amongst this contrast class. Possible reasons may include "it will impress Eve", or "it will allow him to gain the knowledge of good and evil".

This first option, "it will impress Eve", does distinguish eating the apple from many of the options in the contrast class – simply holding the apple for a while is unlikely to impress Eve, and neither is hiding it safely away from those who may eat it. However, Eve may find juggling the apple to be impressive, and she may also be impressed by nicely arranged fruit (though likely not as impressed by juggling). Thus, this reason does speak in favor of eating the apple, but doesn't do a good job of distinguishing the eating of the apple from the contrast class.

On the other hand, "to gain the knowledge of good and evil" does distinguish eating the apple from the contrast class. No options in the contrast class have a connection to this reason, whereas P (the event, eating the apple), is connected. If he wanted to gain the knowledge of good and evil, none of the options in the contrast class are more likely, whereas eating the forbidden fruit is more likely. Thus, this reason is a better explanation for the event than "it will impress Eve", even though the latter does go some way in making P more likely and distinguishing amongst the contrast class. The correct explanation will

be the feature that does the best job of promoting the truth/likelihood of P and lowering the truth/likelihood of the rest of the contrast class.

Another case will bring out how this works. In medicine, we know that it is very rare for individuals to present with the condition of paresis.¹⁷ One of the most common ways that someone will develop paresis is if they have untreated syphilis. However, among patients with syphilis, developing paresis is still very rare.¹⁸ So imagine a doctor presented with a patient who has paresis. A doctor can seek an explanation with the interest of a doctor seeking the underlying condition, asking “Why does this patient have paresis when the rest of my patients of similar age and health do not have paresis?” In this case, the contrast class is healthy patients without paresis, constructed just as in other cases, by the interest of the doctor and her background understanding of cases like these. Possible reasons will be the variety of underlying or acute medical conditions that bring about paresis. Syphilis ends up being the correct answer, as after medical investigation it is the diagnosis that holds.

However, note that syphilis does not raise the likelihood of paresis by a great deal; indeed, most syphilitic patients do not develop paresis. Thus, a correct explanation need not be “strong” in the sense of raising the likelihood of the event by a great deal. It need only be true and play the difference-making role between the event and the contrast class.

¹⁷ Paresis is a condition of muscular weakness caused by nerve damage or disease; in other words, paresis is partial paralysis.

¹⁸ Van Fraassen uses this example to bring out how some requests for explanation should be rejected. “Why does this patient have paresis?” is a valid question, whereas “Why does this syphilitic patient rather than that syphilitic patient have paresis?” is not.

We can use another medical example to bring out this point: Consider a patient who is losing his hair. A doctor can seek an explanation for this, asking “Why is this patient losing his hair rather than (my other patients who are not losing their hair)?” This contrast class is set up by the doctor’s background understanding and interest in understanding the medical causes involved here. Rival explanations will be medical causes for hair loss, such as alopecia, chemotherapy, etc. Say investigation results in chemotherapy succeeding at differentiating amongst the contrast class and this patient, making receiving chemotherapy the correct answer.

Receiving chemotherapy is the correct explanation for this second patient’s hair loss, and having syphilis is the correct explanation for the first patient’s paresis. However, note that these explanations have differing strength, in a way that van Fraassen’s account can capture. Syphilis, while succeeding at distinguishing amongst the contrast class and supporting the event in question (developing paresis), does not do this as well as receiving chemotherapy does in our second case: Receiving chemotherapy raising the likelihood of losing your hair more than having syphilis raises the likelihood of developing paresis. We have, in this sense, a better explanation for the hair loss than for the paresis, though we have an explanation for both.

According to van Fraassen, standards for explanation are based in our interests and background understanding. These set up a contrast class, and rival reasons are judged based on how well they make the contrast class less likely to be the case, and how well they make the event in question more likely to be the case. The correct explanation (the reason for the event) will be the feature that does the best job of this, given the interests and background

understanding that set up the “Why?”-question. As we saw in the Adam case and in our medical cases, different explanations can do better or worse jobs at explaining.

The standards for this account of explanation are thus pragmatic – the context of providing the explanation set the standards for a correct explanation. With these standards, van Fraassen’s account meets the three desiderata for a theory of explanation laid out above.

The first desideratum for a successful account of explanation was that it accommodate the asymmetrical nature of explanation. The classic example here is that the height of the flagpole is meant to explain the length of the shadow, but not the reverse. The explanandum and explanans should be asymmetrical in this manner. This desideratum is related to the second, for it is addressing how we pick out the salience of the events in question and the features that explain them. Recall from above that theories of explanation that don’t seem to be able to account for the asymmetry of explanation, such as the statistical significance of a feature for an event occurring, or whether a feature is grounds for believing the event occurred, are the views that also had difficulty in selecting the salient features to do the explanatory work.

Van Fraassen’s view fares better here, accommodating the roles of explanandum and explanans in particular context. In order to explain, we must ask a “Why?”-question of a particular explanandum, and we will get an explanans relative to an interest-dependent explanandum. The structure of his account means that asking different “Why?”-questions gives you different explananda, so in switching explanandum and explanans, you’ve switched interests as well. In other words, asking “Why is the shadow this long?” and “Why

is the flagpole that tall?” are two different questions, with different interests. This is possible on van Fraassen’s view, but is not pernicious in the way that it was in the previous views, where being able to explain shadows and heights in terms of one another elided salience. In van Fraassen’s account, it brings out the role salience is playing in his view.

For the extant views, the shadow and the height of the pole do roughly the same quality job at explaining one another, which is strange. On van Fraassen’s view, it would be strange for the flagpole’s height to be a well-answered “Why?”-question, but that doesn’t mean we couldn’t find ourselves in such circumstances. For the flagpole’s height to be an explanation for the length of the shadow, the circumstances would have to be pretty out of the ordinary, and van Fraassen provides such a tale, which he calls “The Tower and the Shadow”. He provides a context where the length of the shadow does explain the height of the flagpole.

Explanation has an asymmetric structure, but one that nevertheless is flexible. Changing the context can shift the asymmetries. Van Fraassen’s view can account for this.

Second, a major motivation for the view is the standard that the alternative theories have the most difficulty with. While many of the above views can have difficulty selecting among features that distinguish the explanans for explanatory relevance, van Fraassen contends that the selection of the salient causal factor is a matter of context. In *The Scientific Image*, he articulates the failed attempts of alternative approaches to explanation and causation (van Fraassen, 1980, Chapter 5). Contextual factors such as a person’s interests and “various other peculiarities in the way he approaches or comes to know the problem” affect what counts as salient. He formalizes these peculiarities in terms of

interests and background knowledge, and claims we cannot arrive at a “real” or “more objective” or more “true explanation by simply combining context dependent answers” (van Fraassen, 1988, 126). The selection is thus a matter of context, and the correct answer is set by this selection as well. The pragmatism of van Fraassen’s view is a response to this desideratum.

Finally, a successful account of explanation must allow for some requests for explanation to be misguided, illicit, or denied. The standards for the “Why?”-question accommodate this. In asking a “Why?”-question, you are presupposing:

1. P is the case (the event happened, as described)
2. All other members of the contrast class did not occur (so, in the case of the house burning, the rest of the houses in the interested party’s contrast class didn’t burn down)¹⁹
3. There is a feature, A, that distinguishes between P and the rest of the contrast class, given our background understanding

A request for an explanation will be rejected (or is illicit) under the circumstances that the event didn’t occur how the request presumes, or if the contrast class has been mis-constructed. A contrast class has been mis-constructed if some member of it did in fact occur.²⁰ There thus should not be an explanation for why events that did not occur occurred (unless we switch to a hypothetical mode).

¹⁹ There are other conditions I am eliding here for clarity and brevity, that don’t affect the rest of the discussion. Van Fraassen’s formalization has it that P (the event) is a member of the contrast class, which is a class determined by a similarity relation. X’ is the contrast class, without P. Thus, a presupposition of a “Why?”-question is that the members of X’ did not occur.

²⁰ Or, technically, if it doesn’t bear a similarity relation to the rest of the class. I will not be addressing this condition (it isn’t very clear, and I don’t believe this will pose difficulties down the road). The answer,

A further way an answer can be ruled out is if something in the background understanding rules out an answer, thus conflicting with (3) above. For instance, if I ask “Why did the uranium disintegrate today, rather than the other days I checked on it?”, my background understanding of the behavior of uranium rules out the presuppositions implicit in the question. We understand why uranium, rather than helium disintegrated (its radioactive nature), but we can not predict when uranium will in fact disintegrate.²¹ There is thus no explanation for why the uranium disintegrated *today*, and van Fraassen’s view can say why.

Van Fraassen’s theory of explanation thus meets these desiderata from the philosophy of science literature. I suggest it also is thereby well positioned to be applied to the practical domain, where individuating actions is interest-sensitive, and where an alternative for the standard view will need standards for good and bad reasons. In what follows, I will apply van Fraassen’s view to the practical domain, and show how the standards will satisfy the desiderata for an alternative view from the previous chapter.

The framework I will bring over from the philosophy of science is the structure of the “Why?”-question and the method of evaluation. I turn to that now.

according to van Fraassen must also bear a relevance relation (R) to P and X, given the background understanding (this relevance relation is not reducible to causal relevance). Thus, a formal articulation of a “Why?”-question, where Q is the question, k is the background understanding, so Pk is the member of the contrast class being questioned, would be the following : $\langle Q, Pk, X, R \rangle$. It is a request for the feature that bears R to Pk, in X. The form of a direct answer would be “Pk in contrast to X’ because A”, which paraphrases “A bears R to Pk”. In my discussion I leave out the paraphrase and deal instead with the direct answer.

²¹ Further, we don’t consider it to be a gap in our understanding of physics that we fail to predict the decay of uranium in this way.

Chapter 5: Reasons as Explanation

In this chapter, I apply van Fraassen's account of explanation to the practical domain. I articulate how a reason can support performing an action by being part of a good explanation for performing the action. After summarizing van Fraassen's view in Section One, I outline how the practical domain complicates the picture in Section Two. When we explain behavior, rather than non-intentional phenomenon, we implicate the perspective of an agent. Therefore, our understanding of their agency and the way they engage with the world will be an important factor in the context: our background understanding will include our grasp of the agent in question.

In Section Three I turn to human behavior. I show that the same framework that has provided standards for explanation throughout the discussion thus far still applies: the interests and background understanding of the individual seeking the explanation determine the contrast class. This allows us to see the fully articulated "Why?"-question, and evaluate rival explanations. In all domains, the correct explanation is the feature that distinguishes the event from a contextually determined set of alternatives. In Section Three I show how we can understand this in the practical domain.

Section Four is dedicated to applying this structure of explanation to the first person perspective. When an agent considers what to do, she is taking considerations to be reasons in support of her behavior. I suggest that taking the consideration to support her behavior can be modeled as an internal version of the same explanatory framework I've discussed thus far. The way a reason supports is by distinguishing an event from a given set of alternatives, and when an agent takes a consideration to have this sort of support, she has

a perspective made up of interests and background information and is taking the consideration to do the same distinguishing work amongst behaviors that features were doing in the other sections.

I conclude this chapter by addressing a major burden for an explanation-based view of reasons, and a major reason to bring an account of explanation into the practical domain. In Section Five, I outline how the standards for successful explanation provide an adequate framework for evaluating good and bad reasons for action.

Thus, overall in this chapter I lay out an explanation-based view of reasons for action. In the following chapter I will go into main objections to this view.

Section One: Recall van Fraassen

On Van Fraassen's account, what counts as a good explanation is determined by pragmatic standards. In other words, what counts as the correct explanation will depend on context: the background understanding and the interests of the discussants. This context sets up a contrast against which we explain the events we seek to understand.

For example, imagine we seek an explanation for a car crash. This can be captured by a generic "Why?"-question: "Why did the car crash"? However, according to van Fraassen's view, a fully articulated "Why?"-question includes a contrast class determined by the context. Context is made up of the *background understanding* and *interest* of the inquirer. When we attempt to explain a car crash, we are seeking an explanation with a particular interest, and we understand the event with a background in circumstances like

these (for instance, we understand something about cars, driving, roads, etc. and likely more specialized things depending on our interests).

Say you are a police officer seeking an explanation for the car crash. You have an understanding of how different speeds affect driving conditions, and this road, the behavior of drivers, etc. As a police officer, you could be interested in the *legal behavior* that led to the crash – say you are interested in the speed of the cars on this road. This interest in finding an explanation allows us to hone in on the aspect of the event you are seeking an explanation for. In this case, because you are interested in the speed of the cars on the road, you would be asking for an explanation for why *this* car crashed when other *cars of this kind* didn't crash, where other cars of this kind is understood in terms of speed. You thereby are interested in the event *qua citizen driver on the road*. The policeman's context determines a contrast class of cars driving on that road that didn't crash. So, the fully articulated "Why?"-question would be "Why did this car crash rather than these other cars on the road that didn't crash?"

Other contexts determine other fully articulated "Why?"-questions. For instance, you could also be interested in an explanation for the car crash as a municipal designer. In such a case, you would also be asking for an explanation for why this car crashed when other *cars of this kind* didn't crash, but what counts as other cars of this kind for the municipal designer will be in terms of the context here. Your background understanding would go beyond the general understanding shared with most people about cars, driving behavior, etc., for a municipal designer has specialized knowledge about road design and driving conditions related to visibility, turning radius, etc. Also, your interest as a municipal

designer would differ from the police officer. You could be interested in potential obscured vision of drivers, for instance, making cars of this kind those in similar driving conditions – similar plant growth on the sides of the roads, similar turns and intersections, etc. The municipal designer is interested in the event *qua crash in these driving conditions*. In other words, the municipal designer is looking for the feature that made this crash happen when in other similar conditions crashes did not happen. So, the request for explanation would be “Why did the car in X, Y, Z conditions crash when the cars in (ABC) conditions, (DEF) conditions, (GHI) conditions, etc., did not crash?”

Requesting an explanation as a tire vendor would result in still a further fully articulated “Why?”-question. As a tire vendor, you could be interested in the condition of the tires of the car that crashed and the difference between those tires and the tires of cars in similar conditions that did not crash. Thus, the vendor will hold the rest of the conditions of the crash constant (such as those visibility conditions that interested the municipal designer), and consider different tire conditions in the cases that didn’t lead to a crash. Thus, as a tire vendor, the discussant would ask, “Why did these tires lead to a crash, rather than X tires in similar conditions (which didn’t lead to a crash), and Y tires in similar conditions (which didn’t lead to a crash), etc.?”

In this way, different interests pick out a different aspect of the event that is to be explained and pick out a different set of contrast cases that the explanation will distinguish the event from. The police officer is looking for what happened with this car that made it different than the other cars on the road at that time. The municipal designer is looking for features of the environment that made the driving circumstances dangerous and is

understanding the event in terms of external conditions. The tire vendor is looking for how equipment can be more effective and is understanding the event in those terms. These different contexts hone in on different aspects of the event and ask for an explanation of the event relative to their interest and background understanding.

The event has some feature that the contrast classes don't. The request for an explanation is a request for this distinguishing feature. The correct explanation for the police officer could be that at the speed the driver was going, skidding is more likely. The correct explanation for the municipal designer could be that there was a bush obscuring the speed limit sign. The correct explanation for the tire vendor could be that the tread on the tires was worn down. These features of the event are what distinguish the event from the contrast class of these respective contexts.

In order to reach a correct explanation for the car crash, often some empirical investigation is required. For example, the tire vendor needed to have information about the tires of the car that crashed and the relevant cars from his contrast class. Rival explanations will be possible features that differentiate between the crashed car's tires and the contrast class, but the quality of explanation will increase the better a job the reason does at making a difference between the contrast class and the event (crash). Then the explanation is the feature that did the best job of distinguishing between the contrast class (the tires on cars in similar conditions that didn't crash), and the car that crashed.

Thus, for van Fraassen, the correct explanation will be the feature of the event that does the best job distinguishing the event from the contrast class. This will be relative to the context. I bring this structure of evaluation to the practical domain: the standards for

good reasons will be pragmatic in the way outlined here in van Fraassen's view, articulated by the "Why?"-question and the context.

Section Two: Explaining Action

In the practical domain, reasons explain actions. Van Fraassen focuses on scientific explanations of natural phenomena, not intentional actions. Van Fraassen and the rest of the scientific explanation literature aimed to capture the way that scientific theories explain events like why a sample turned green rather than another color and why the water/copper substrate reached equilibrium temperature at 22.5 degrees Celsius rather than some other temperature (van Fraassen, 1980, 127).¹ In this chapter I outline my proposal for the way that explaining actions is continuous with such explanations.

For scientific explanations, the focus is on natural phenomena in the world. In contrast, when we explain an agent's behavior, we are locating the agent as a relevant part of the story of how a phenomenon came to be. For instance, rocks and lizards are both part of the natural world. However, while it is implausible to explain the rock's presence in that location in terms of its perspective on the world, it is not similarly implausible to explain at least some animal behavior in terms of their perceptions of the world and their inclinations. When we see a lizard on a rock, there is an intuitive difference between asking "Why is that rock there?" and "Why is that lizard there?" To capture the rock's presence, we appeal to forces acting upon it. On the other hand, when it comes to the lizard, another

¹ I will refer to this context as a scientific context (rather than causal context), because I take it that on any reasonable view the way an agent works in the world will be able to be understood causally.

approach can be appropriate: namely, appealing to the lizard's perceptions and inclinations. Roughly, what is distinctive about agential explanations is that the reason that explains the event is a feature that moved an agent to bring the event about. What the "Why?"-question is seeking is thus something that the agent took to be the case that resulted in her action.²

A typical case of action explanation shares the key features I noted in the above scenario from a non-practical domain. In other words, the same factors determine the fully articulated "Why?"-question when we seek a reason behind the rabbit's behavior as when we sought an explanation in the car crash case: interests and background understanding. In the car crash case, a fully articulated "Why?"-question included a contrast class which identified the cases against which the crash was being distinguished. For the police officer, the contrast class was made up of cars on the road going different speeds and this was determined by the description under which the officer was considering the event: her background understanding and interest in the event. The structure will be the same here.

However, while the constituents of context are the same (background understanding and interest of inquirer), the background understanding in cases of intentional explanation implicates our understanding of an animal's perspective on the world. When we explain the location of a rock at the bottom of a hill by appealing to gravity and incline, we identify the forces that acted on the object. When we explain the behavior

² There is, of course, a vast literature on what it means to act intentionally or act on a reason. I hope to remain agnostic regarding what it precisely takes for an action to be for a reason as opposed to "mere" behavior beyond what I say here. The relationship between acting for a reason and intentional action has been much discussed, see: Alvarez (2009), Chan (1995) Mantel (forthcoming), Schlosser (2012) Shute (1992), Tuomela (2006). For a sampling of the issues involved in acting for a reason, see: Velleman (1992) and a response by Zaragoza (2006), as well as: Korsgaard (2005) and McAninch (2015) for a discussion of the way we discuss reason in motivation.

of an animal, we attempt to track what it took to be the case in the environment that led to its behavior. For instance, we could explain why a lizard, rather than the rock, is at the bottom of a hill. To do this, we get into the way the animal sees the world and responds to it – something we take the rock not to do. The lizard could see the sun shining on the rock and so move in that direction. This sort of explanation for the lizard’s behavior appeals to background understanding of how lizards work: They’re cold-blooded animals that warm themselves largely in the sun, and are able to recognize differences in light, and propel themselves forward on the terrain in question. My grasp of the access lizards have to information about the world and what incentives they have for action is an important part of my background information.

This is another important parallel with scientific explanation: the ability to identify relevant features of the environment for investigation. My background understanding suggests that the proximity to Mercury retrograde, while a potential distinction between today’s lizard behavior and the contrast class, doesn’t distinguish amongst lizard behaviors – it doesn’t make it more likely for a lizard to go towards a rock than do a lizard dance or stay still. My understanding of lizard-sensitive things in the environment will guide my investigation just as the municipal designer is looking for features of the environment that affect visibility, not just any old thing in the environment. The municipal designer understands that weeds on the side of the road don’t affect driving conditions in the way that potholes and trees do. In short, my understanding of lizard behavior plays a parallel role to the background understanding of the inquirers in the scientific explanation examples above.

To get clear on the picture of intentional action generally, we can start with the simpler cases of non-human animal action, for such animals are non-reflective creatures but nevertheless we often capture their behavior best with agential explanations. In cases of such animal behavior, we commonly take there to be something in the environment that the creature is responding to, and our explanations pick this feature out under some relevant description.

Imagine a non-human animal behaving: say, a rabbit running around in a field. We can explain this behavior in the frame provided by the scientific explanation literature by citing a feature that explains the event in scientific (non-intentional) terms: the cause (large muscle contractions), or law of nature (a complex list of nomological interactions), or statistically relevant feature (relations of likely behavior in the wild). For van Fraassen's account, this would mean requesting an explanation in a context that seeks to understand the behavior and event in natural (non-intentional) terms: locating a force that acts on or through the rabbit. Because we seek explanations of aspects of events under descriptions, the interest and background understanding would thus be in the form of an explanation request for the rabbit running away in a non-agential context.

Alternatively, we could provide an explanation that appeals to the rabbit's perspective and seek what the rabbit saw about the world that moved it to act.³ In other

³ Approaching action in this way is similar to how Aristotle discusses the lantern example in Aristotle (1984) II, 11. In this example, we see how different contexts yield different explanations of the same event. When someone leaves a light on outside the house, asking why it is on can receive four different answers depending on the concern. In other words, different concerns make different answers to "Why?"-questions appropriate. Considering the mechanics of a lit light, the social meaning of the lit light, etc. will result in different appropriate explanations. Similarly, in the Posterior Analytics, Aristotle develops the four causes, which can be understood as different answers to "Why?"-questions depending on one's interest in the event (Aristotle, 1984, I, 13). See van Fraassen (1980)

words, when an animal acts, we can seek an explanation for its behavior in an agential mode. Our interest in understanding the behavior and our background understanding of the circumstances will determine the aspect of the event/behavior we are seeking an explanation for, just as in the scientific (non-agential) cases.

Imagine observing that rabbit running in a field and being puzzled by its behavior. There could be a hawk overhead that you don't see, or perhaps you do see it, but you would suspect a different pattern of franticness given the presence of such a predator. If you seek an explanation for the rabbit's behavior, your background understanding of rabbit behavior and this particular situation will determine the "Why?"-question. If you have a great deal of rabbit understanding (you know when they dart around versus stay still, what sorts of environments they prefer to graze in, which predators are most dangerous and what behaviors they take up in response to them, etc.) but don't see the hawk, you would ask a different "Why?"-question than someone with very little rabbit understanding but who sees more of the empirical field. All else being equal, in the latter scenario a good explanation would provide you with rabbit information. In the former, you need to be informed about the hawk. Further, if you own the hawk and are training it, you could seek an explanation for the rabbit's behavior with a different interest than if you are a zoologist studying small mammals: you'd hone in on different aspects of the event to be explained, potentially, just as in the above car crash example.

In seeking an explanation for the behavior of the rabbit, besides having an interest that will hone in on the behavior under a relevant description, the background understanding will include information about rabbits. Because, in this case, we are

explaining an agent's behavior, part of our background understanding will be information about the agency in question – in this case, rabbit agency. So, we have an understanding of the rabbit in a similar way as we have an understanding of how cars work and what driving is. We may have an understanding of what rabbit goals are, their abilities, tendencies, limitations, etc.

We also have our own read on the rest of the empirical situation – what the field is like and its features, just as we did in the crash scenario – its size and location, perhaps its contents, etc. This is again a parallel to the non-practical scenario as well, for we had a read on the rest of the empirical situation the car was placed in as well.

If we look to explain the rabbit's behavior, we will ask “Why is it doing X?” The interest in the rabbit's behavior and the special background understanding of the interest are what create the contrast class that would determine the fully articulated “Why?”-question. In the case of the rabbit, it may head towards an obstacle in the field rather than to a hole in the fence surrounding the field. You may seek to explain this behavior with a variety of interests and this would lead to different reasons/explanations.

With the interest of the hawk's trainer, you could be contrasting this animal's behavior with other fleeing animals. Thus you are interested in the event *qua fleeing small mammal's behavior*. Say your background understanding concerns fleeing small mammals, and you want to know what struck this fleeing small mammal that made it avoid the hole and go for the rock, perhaps to better train your hawk. The class under consideration is small fleeing mammals, and you are interested in what made this small mammal's behavior different.

Thus, you would ask, “Why did the rabbit hide behind the rock rather than behave like that rat that went through the hole, the fox that went through the hole, the vole that went through the hole, etc.?” A good explanation will be what about the rabbit’s take on the world made it hide behind the rock rather than flee through the hole. What distinguished the rabbit behavior from other small mammals? (A feature about rabbits in general would do a good job here, perhaps, or maybe this individual rabbit is weird. I leave it to the rabbit specialists.)

Note also the role that background understanding can play in determining the fully articulated “Why?”-question. If you, as the hawk trainer, had a different understanding of rabbit behavior or the empirical situation, this would determine a different class of alternatives. If, for instance, you were under the impression that rabbits were possessed of great strength and bravery and stood up to all predators, you could be quite confused by any fleeing behavior. You could think that rabbits didn’t have fur but rather had feathers and wings and thus would class them with birds that your hawk would predate on, creating a contrast class of behaviors including flying away. On the other hand, you could understand the scenario other than in terms of features of rabbits. You could think that there is a warren right next to the rabbit, and thus include diving for it in the contrast class, for instance. In this manner, it is not just the interest of the hawk-trainer but also the background understanding that determines the contrast class, and thus the fully-articulated request for an explanation.

Other interests could lead you to seek an explanation for the rabbit’s behavior. With the interest of the rabbit’s owner, you could be interested in what made the rabbit behave

in this way in particular today. Say you have observed the rabbit going for the hole when pursued every other time it was spooked. In this case, you are seeking an explanation for the event *qua rabbit today*. Thus, your interests and background understanding set up the contrast class of the rabbit's behavior on every other observed occasion, and your request for an explanation is, "Why did the rabbit go for the rock rather than (behavior on day X, which was going for the hole), (behavior on day Y, which was going for the hole), (behavior on day Z, which was going for the hole), etc.?" In this case, a good explanation will cite the feature that made it more likely that the rabbit went for the rock rather than the behaviors in the contrast class. The feature is what distinguished today from these other days in making it most likely that the rabbit did this rather than the behaviors it performed on the other days.

The explanation for an event is the feature to which the agent is responding that made the difference between the action and the contrast class. This is the same thing that was happening in both the rabbit case and the car crash case. When someone asks for an explanation of the rabbit's behavior, they are looking for a feature that made the rabbit do what it did rather than the set of alternatives. Similarly with the car crash: the "Why?"-question sets us up to identify the feature that plays the "difference-maker" role – what made it the case that the car crashed (under the relevant description) instead of the salient alternatives? Explanations pick out the features that plays this "difference-maker" role. Given a scenario, the feature in question makes the event that took place the one to expect rather than the contrast class of alternatives (the other things that are implicated by the questioners interests/understanding).

So the interests that we have in the case are as paramount here as in the cases in the scientific explanation literature. You can't explain the rabbit's behavior, full stop. You can explain the rabbit's behavior relative to the context: your understanding and interest in making sense of it. Here, I have modeled how to explain the rabbit's behavior *qua fleeing small mammal*, or the rabbit's behavior *qua my pet*, for example.

A further parallel with the scientific cases is that frequently we need to investigate in order to identify the correct answer to the "Why?"-question. Recall that in the car crash and other scientific examples some empirical investigation may need to be added to our background understanding in order to judge explanations to reach the best one. If I am looking for the reason my rabbit went to the rock today rather the hole (as he has done all the other times I have observed in similar situations), merely identifying the contrast class isn't enough to bring about the correct explanation. I need to investigate the features that distinguish my rabbit today versus the other instances I have observed him, or perhaps go into the field to try to see things from the rabbit's perspective.

Thus the best explanation is relative to context. The interests and background understanding constitute a contrast class against which the behavior is being understood. This contrast class sets the standards for a good explanation, because the feature that is the explanation for the behavior is what distinguishes the behavior from the contrast class, given background understanding. Key to cases of agential explanation is that we seek to "get inside" the perspective of the agent in question to capture what it saw in the environment that moved it to behave.

Section Three: Human Action

Human action creates more complicated cases than rabbit action because humans participate in complex practices and behaviors and have the ability to reflect on reasons and act on the basis of this reflection. Humans can therefore see competing reasons and consider which to act on, creating complexity in the relationships between the agent and the reason for acting. However, when we seek an explanation for human action, it is still our interests in explaining the action and our understanding of the agent and situation in question that help determine the correct explanation. The model of explanation applies to cases of human action, incorporating our complexities in to the background understanding.

The background understanding here works just as in the rabbit and car crash case. In identifying someone's reason for acting, we take into consideration relevant information about her, the kind of agent she is, and the situation she was in. For example, let's consider the case of my office mate coming in to the office on a Wednesday morning. This explanation scenario could be captured roughly by the generic "Why?"-question: "Why did Zac come to work Wednesday morning?" If we are interested in the case as an instance of agential behavior, our interest will involve getting at Zac's perspective – what he took to be the case that moved him to come to the office. There are a number of ways of characterizing what moved Zac to action, and the pragmatic account of explanation captures how context (our interests and background understanding) determines the appropriate explanation.

For instance, I could be interested as an officemate who wants the office to myself. Along with my interest, my background understanding will determine the context that will

allow for a fully articulated “Why?”-question. My background understanding includes how offices get assigned in my department, and general habits of graduate students, etc. and, in particular, what I’ve noticed about my office mate’s, Zac’s, habits. Specifically, I have noticed that his habits are to work from home on Wednesdays until the afternoon, and have constructed my own work schedule accordingly. As an officemate wondering about Zac’s choices today that are preventing my solitude, I am interested in the event *qua Zac working at the office today*. The contrast class this picks out is Zac’s behavior on other Wednesdays. Thus, my seeking an explanation could be modeled by the “Why?”-question: “Why did Zac come to the office at 8am today rather than (his arrival time last Wednesday), (the Wednesday before that), (the Wednesday before that) etc.?” The answer to this question will be the feature that differentiates the event today from these other events.

Another interest that could motivate explaining Zac’s behavior could be that of an administrator tracking office use to see how crowded offices get on different days. Suppose there is a big increase in office use on Wednesdays that is causing trouble in assigning offices, which otherwise get used in a scattered fashion that would allow more assignments per room. Such an administrator would have a background understanding of how universities have offices for graduate students and perhaps some specialized knowledge of trends and how they are assigned. Her interest could be in explaining trends in office occupancy throughout the week. The administrator is interested in the event *qua student using an office on Wednesday*. Thus, her explanation could be modelled: “Why did Zac arrive on Wednesday rather than (students who arrived on Monday), (Students who arrived on Tuesday), (Students who arrived on Thursday), (Students who arrived on Friday)?” The

administrator is looking for a reason for the surge in office attendance in the middle of the week, and is wondering why Zac came to school today. The good explanation here will distinguish between Zac's Wednesday arrival and the rest of the week's arrivals.

Alternatively, we could consider the interests of a parking and transit worker looking into commuter times. There could be plans under consideration that would charge different rates for parking, depending on commuters who come to campus at off-hours, and the parking and transit worker could be attempting to track how effective the policy has been in altering the time commuters have been arriving. The commute studier is seeking an explanation for the event *qua morning commuter*. Thus, I could be asking, "Why did he arrive at 8am rather than (student that arrived at 9am) (arriving at this other time), (arriving at that other time), etc.?" If the reason is because of parking rates, this will be informative for future policy initiatives. The commute studier wants to know something about Zac's motives, but under a different description than my interests as an officemate, and under a different description than the office-assigning administrator.

In the first context, where I am interested in Zac arriving earlier on Wednesday than my background understanding would suppose, and my contrast is the other Wednesdays where he arrives in the afternoon, I may have some guesses at the correct explanation. My background understanding of Zac and the dynamics of our office and department life can suggest some: he could have a meeting with a student, he could be trying something new in order to be more productive, he could have caught a convenient ride to work, there could be a department meeting I don't know about, an event on campus, something could be happening at home that would have been distracting, etc...

As with the car crash and rabbit cases, rival explanations will be compared based on how well they make a difference between the contrast class and the event that occurred, given our understanding of the event and the background. So, for instance, if it's the case that there have been distractions at home every Wednesday morning, that reason is discarded – it doesn't distinguish today from the rest of the contrast class of other Wednesdays when Zac arrived in the afternoon.

So, my understanding of Zac, his work habits, and campus life will help to evaluate things here. However, just as in the cases above, there likely would have to be some investigation when it comes to making some final determination between rival explanations – some cluster of possible reasons may just be really likely and we may have to ask Zac some questions about his life, or check email in order to see if there are campus events, or just ask Zac what's up and why he's at the office, to see which are actually the case and which not.

It could turn out that the difference-maker this Wednesday was that he needed to print something out for work and came in to use the department printer. This distinguishes today from the other Wednesdays when he came in later in the day (on those days he did not need to use the department printer). My interest and background understanding constituted the context that set up the contrast class of the other Wednesdays when Zac came into the office in the afternoon, and thus placed the demand on an adequate explanation that it distinguish the event today (Zac's arriving early) from the events on other days (Zac's arriving in the afternoon on each of the other Wednesdays).

The reason was that he needed to use the printer. Other reasons could fail for the same reasons they would fail in the scientific context: because they fail to distinguish between all of the other Wednesdays and today, or because they weren't in fact true today, or because they are ruled out by my background understanding of Zac, offices, and the working of the university. This is how the mechanisms from the scientific context apply to the practical domain. Our interests and background understanding set up standards for a correct explanation, which is the reason for the agent's behavior.

Section Four: Using Reasons to Decide what to Do

Van Fraassen's model of explanation characterizes the way that features explain events: they make a difference between an event and relevant alternatives given a set of interests and background understanding. In other words, a feature *explains* by *making sense of*, by fitting well into a contextual understanding of the world given a set of interests. For scientific explanations, a feature does this well by raising the likelihood of the event while lowering the likelihood of the contrast class. In this chapter thus far, I have laid out how this is also how we can understand explanation of events that involve agency: When we seek to explain an agent's behavior, we are interested in a feature that moved the agent, and thus capture something that made it more likely that the behavior occurred rather than a set of alternatives. In other words, we are seeking to identify a difference-maker in the same way as in the scientific explanation literature: our background understanding and interests identify the description under which the event was relevant and a contrast class of alternatives against which the event will be distinguished. The best explanation will

distinguish the behavior from the contrast class by raising the likelihood of the event while lowering the likelihood of the contrast class.

From the agent's perspective while deliberating, the frame of the pragmatic model remains the same. In other words, the interests and background understanding of the agent determine what she takes to be a reason for action. A reason for action will be a feature that distinguishes a behavior from a class of alternatives determined by interests and background understanding. When we decide what to do, we take a feature of the action (or the world) to count in favor of performing an action by taking it to make the difference between the action and a set of relevant alternatives given our interests.⁴ Van Fraassen's account of explanation lays out the way that a consideration explains in the scientific context. I suggest that this model can be extended to cases of agency both in and out of the deliberative perspective by applying the structure van Fraassen develops to understand explanation, and thereby reasons, as difference-makers in a context. The model was directly applied to cases of behavior that have occurred, but there will be some adjustments for the deliberative perspective.

Let's take officemate Zac as an illustrative example of how this view can be transferred to the deliberative perspective. I claim that when an agent takes a feature to be a reason to act, she is implicitly taking a feature to explain. So, when Zac takes there to be a connection between needing to print something at the office and going in early on

⁴ Velleman has a similar way of articulating the deliberative perspective – in terms of making sense of one's self in the world. See Velleman (2005) and Velleman (2009)

Wednesday, I suggest he is tacitly taking there to be an *explanatory connection* between performing an action and his understanding of the world.

Zac had interests (say, graduate-student-interests in getting research done cheaply) and a background understanding of his resources and how the world worked. In scenarios where we are explaining an agent's behavior, we have been focusing on using context to identify the event under the relevant description, such as the behavior *qua Zac's coming to the office today*, or *qua student using an office on Wednesday*, or *qua morning commuter*. Recall that identifying the event under the relevant description determines the contrast class - the relevant alternatives. Zac's perspective includes interests and background understanding, and this leads to him approaching potential actions under descriptions as well. Zac could be imagining his potential actions with his graduate-student hat on, interested in being productive on the cheap, for instance, thereby imagining potential actions *qua cheap research options*. My suggestion is that with this perspective on potential actions, when Zac takes a consideration to be a reason in support of an action, he is tacitly taking that consideration to *differentiate among potential actions given this context* – and thus the pragmatic model of explanations/reasons captures the framework here in the same way it captures how features differentiated between an event and the contrast class when we attempted to explain his behavior.

When an agent takes a consideration to be a reason for action, she is taking it to distinguish between a behavior under a description and a contrast class given her interests and background understanding (given a context), thereby taking the reason to be a “difference-maker” of sorts. The pragmatic model captures the way that an agent takes a

feature to “count in favor” of performing an action in this way: the interests and background understanding frame the perspective. However, there are key distinctions between the perspective of deliberation and the scientific and agential explanation circumstances.

First, note that when we deliberate, we are not seeking an action we are most likely to take up compared to alternatives. In other words, whereas the method of evaluating rival explanations, or reasons, in the scientific and explanatory cases was in terms of how well the feature increased the likelihood of the event and lowered the likelihood of the alternatives in the contrast class, this aspect of the model does not transfer literally to the deliberative perspective. Instead, the factors involved in the pragmatic model are still the relevant determining factors from the deliberative perspective – the agent’s interest and background understanding shapes a contextual perspective. It is this context that determines potential alternatives behaviors. Features will count in favor of performing the behavior under deliberative consideration to the extent that they distinguish the behavior from the alternatives in the contrast class, which is determined in the same way as in our other settings, but this is no longer a predictive relationship.

When we consider a potential action – say, making a birthday call to a sibling – we are considering actions *qua good sibling*. Our interest in being a nice sibling together with background understanding of communication practices, birthdays, and relationships amongst family members determine a contrast class of alternative potential behaviors (send a card, purchase a gift on amazon and have it sent directly, go in on a gift together with family, go visit during break, etc.) Reasons to make the birthday call to the sibling will be features of that behavior that distinguish it from the contrast class given the background

understanding. For deliberative cases, for instance, it may be more relevant that the behavior has the feature while the alternatives lack it, rather than that the feature raises likelihood. A good reason will be a feature of the behavior in question while not a feature of the alternatives, and be relevant to the interest – in this case, being a good sibling.

A second key distinction between the deliberative perspective and the settings I have outlined thus far is that the agent is considering *potential* events. This perspective is forward-looking and therefore the explanations are hypothetical. In the cases in the scientific explanation examples, and cases where we are seeking explanations for actions that have taken place, we are looking for a feature that differentiates between the action in question and the scenarios that the context picks out as similar, that this action would have fit in with. Why did this action occur rather than the class of actions that is relevantly (given our interests and understanding) similar? Roughly speaking, the explanation is the feature that does the best job of raising the likelihood of the event while lowering the likelihood of the events in the contrast class.

From the perspective of deciding what to do based on reasons, the actions in question have not yet taken place. Thus, the agent isn't considering a contrast class of circumstances where an agent *behaved* one way, and seeking the feature that makes a difference between that class and how she behaved in some relevant circumstances, contextually described. Instead, she is articulating a *potential* action, given her interests and understanding, and evaluating potential explanations from an imaginative, hypothetical perspective. If she behaved in such a way, *would there be* a good explanation?

Is the feature she is considering a good explanation: namely, does it distinguish the action from the given set of alternatives?

Recall how the interest determines the contrast classes. When the municipal designer was considering the car crash, his interest in visibility set up the contrast class in terms of strips of road with similar design where crashes didn't take place. In order to set up a contrast class, the interest determines what is held fixed (day of the week for me as an officemate, visibility for the municipal designer). This will work in the same way with the forward-looking deliberative perspective, but concerning potential events.

Given a context, features of the world differentiated amongst his options, highlighting the action he ended up taking. This need not involve an explicit or conscious weighing of different considerations, but we can ask how this process might work, if taken up explicitly.

Consider Suzanne, a bright young millennial considering her job options. She's been accepted to the Cordon Bleu in France and is trying to decide whether to attend. The relevant context is made up in the same fashion as in the previous cases. The background understanding and interest can be modeled by a "Why?"-question that will allow considerations to be evaluated as explanations (or reasons) for pursuing the relevant action. The distinction is that the question doesn't ask why an event that has occurred happened instead of the events in a contrast class. Rather, because Suzanne is concerned with acting for a sufficient reason, she is concerned with having a good answer once she acts to the question: "Why did she ϕ rather than X?" where the action she is considering is ϕ and the contrast class is X.

Suzanne could have a number of interests in pursuing a career: a child of demanding parents (with an interest in making them proud), a college graduate (with an interest in putting her degree to use), and as an adventurer (with an interest in traveling and trying new things).⁵

Suzanne is concerned with whether she has sufficient reason to pursue culinary school in France: thus tacitly the concern is how strong an explanation she would have to be a culinary student. She has a background understanding of herself and how she would fit careers in her life (how these careers work, how she relates to others, and how things make sense in the world generally). She can evaluate potential explanations given her different interests. Thus for Suzanne, if she is considering the culinary career path, the alternatives that make up the contrast class under consideration at a given point will be fixed by her interests.

With the interest of making her parents proud, she can ask “Why pursue a culinary school in France rather than (go to law school), (become a teacher like my mother), (volunteer at a non-profit), (pursue other careers that make parents proud)?” The contrast class here is set up in the same way as in the above cases – the interest sets the similarity condition for the instances, like “cases where there was similar visibility but no crash” for the municipal designer. In this case, the contrast class is made up of careers that Suzanne imagines would make her parents proud, framed by her interest and background understanding. The “Why?”-question, framed thus, is seeking a feature that makes the most

⁵ These interests may be loosely similar to how Christine Korsgaard discusses practical identities in Korsgaard (1996). They are commitments that the agent has for how to act that shape how she will evaluate the strength of her reasons for action.

difference between pursuing culinary school and these options. Will culinary school distinguish itself from these other things that would make her parents proud?

With an articulated “Why?”-question, we can evaluate potential reasons. Why would Suzanne pursue this career? She could see a number of considerations that may differentiate between pursuing being a chef and these alternatives. Perhaps she likes funny hats, or the hours fit her preferred schedule (working late and sleeping in), or the culinary world allows her to work with eccentric people, etc.

Suzanne can consider these features and how well they differentiate between the contrast class and pursuing a culinary career. The contrast class set up by her interest of making her parents proud is constituted by a set of career choices, and she can consider whether her affection for funny hats would do a good job of distinguishing between pursuing them rather than culinary school. Similarly for the nature of the hours and working with eccentric people. The best explanation for pursuing culinary school rather than the contrast class will be her consideration that makes the biggest difference between the contrast and pursuing becoming a chef, say, the flexibility of the work hours.

This consideration is the strongest reason she has to pursue the culinary career, given this interest. She can consider whether she has sufficient reason to pursue this career given her other interests as well, of course.

With an interest in pursuing a life of adventure, she can ask “Why go to culinary school in France instead of (teach English abroad), (take up a job with lots of vacation hours), etc.?” where the contrast class is made up of action options that are articulated based on similarity to Suzanne’s conception of her commitment to adventure. The features

that seem to differentiate between going to culinary school and these alternatives for Suzanne will construct the best explanation – the best reasoning – for pursuing this path from the perspective of her interest as an adventurer. The possibility of meeting eccentric people, the fact that it's in France, and similar considerations are strong candidates for this role.

There can be meta-questions posed from the deliberative perspective regarding which interest should have priority. For instance, Suzanne can ask, “Why should I take up the interest of making my parents proud, rather than (interest of adventurer) (interest of college graduate), etc.?” It is important to note that this “Why?”-question, like all inquiries for explanation, comes from a context that is guided by an interest.

We can see a similar structure if we reconstruct the Garden of Eden example. If Adam considered what to do with the apple, his background understanding and interest would set up the “Why?”-question.

With an interest in being impressive, the contrast class is set up with actions that would show off. “Why eat the apple rather than (juggle it), (add it to his fruit display), (throw it on the ground dramatically), etc.?” That it would gain him the knowledge of good and evil will differentiate amongst these options strongly if it would be more impressive than these alternatives. It is thus the correct answer here. However, how strong Adam considers the strength of this reason to be will depend on how impressive gaining the knowledge of good and evil is compared to the alternatives.

Note again the further question that can be raised here. Adam can begin to wonder about what interests are appropriate. In the above example, his guiding interest is

impressing Eve. He can start to think that this may not be the best interest to have, however. Indeed, this is a key capacity of human agency. This higher order consideration of what to pursue would be a further application of explanation and making sense of ourselves.

Section Five: Bad Reasons

Above I outlined the way we identify the best explanation in cases of agential behavior. Doing so identifies the agent's motivating reason under different descriptions, depending on the context. In this section I will address the ways that we can criticize an agent's motivating reason. We can say that an agent has a bad motivating reason when it doesn't do a good job of explaining. As I laid out in the previous chapter, the ways a reason doesn't explain well are the ways it fails to distinguish between the behavior and the contrast class. So we would say a motivating reason is a bad reason when we evaluate the reason according to an interest and finding it lacking along the lines of a bad explanation: it doesn't add to the likelihood of the event or it fails to diminish the likelihood of the alternatives. Let me explain.

In some cases, we may seek to explain a case of human action where the best explanation given our interests and background understanding still doesn't do a particularly good job of distinguishing the event from the contrast class – ie., of justifying. For example, say I go to the latest super hero movie this weekend. My friend and my adviser could each seek an explanation for my behavior, with different interests behind their inquiries.

The friend may be interested in my going to the movies qua weekend behavior. Thus, the contrast class will be other things I do on the weekend, which could include going

to the pool, watching Netflix at home, bringing friends over, working on a project, grading, or doing other research projects – quite a variety of behavioral options that her background understanding of me helps to construct.

The adviser could be interested in my going to the movies *qua* behavior of a student with a deadline. In this case, the contrast class will be behaviors that students (me in particular, most likely) take up when a deadline is nearing. Student-related behaviors include reading, writing, attempting to get feedback, workshoping the project, etc.

From the perspective of the friend, the feature of my going to the movies that differentiates it from the rest of the weekend behavior will be the explanation. My excitement over the opening weekend of a movie telling the story of a character I'm a fan of would be a distinguishing feature here, something that going to the movie has that the other activities don't.

The interest in my behavior *qua student* doesn't capture things in quite this way. When I've worked on projects with deadlines in the past, I have experienced excitement for fandoms and nevertheless pushed through. Further, it is common knowledge that others have worked through fandom excitement and thus the frequency or ease of such behaviors can easily be imagined as part of the background understanding as a sort of normative expectation for how people (like me) tend to act. This understanding informs the contrast class. This excitement thus may not fully distinguish between what I did today and my behavior on these other occasions, when I performed student behaviors.

My reason for going to the movies does a fine job of differentiating my action from those in the friend's contrast class, but doesn't make a difference among the adviser's. This

parallels the way that the Adam eating the apple cases distinguishes good answers for different interests. For the adviser, the best answer for why I went to the movies could be that I was burnt out or forgot a deadline. This is the case because being burnt out differentiates, roughly, between not doing a student-related activity and the contrast class full of student-related activities. For my friend, these descriptions (being burnt out, forgetting a deadline) may not make a difference between the contrast class, for perhaps I went to the pool last time I forgot a deadline (perhaps a lot of my weekend activities happen while burnt out).

Thus reasons can be better and worse given a context depending on how well they distinguish the event from the contrast class. In the above case, the proposed reason didn't distinguish amongst *all* of the alternatives in the contrast class. Recall that another way for a reason to be poor, or less strong, is for it not to play a *strong role* in distinguishing between the event and the contrasts. This was the case in the previous chapter in the instance of the patient with syphilis who developed paresis: though having syphilis did distinguish the patient from the contrast class of other patients who didn't develop paresis, it wasn't a strong indicator. We can see this strength factor most strongly comparatively, perhaps, as when we considered the connection between a patient receiving chemotherapy and experiencing hair loss. The strength of syphilis as an explanation in the first case is weaker than chemotherapy in the second.

Thus when we consider explanations in the practical domain, our interests and background understanding can be brought to bear not only to judge whether the feature

successfully distinguishes the event from the contrast class. We also can make judgments about how well the consideration does at promoting the event.

Consider the case where Dave, a lactose intolerant student, takes a break when an ice cream truck arrives. As his study buddy, with plans to push each other on such days until dinner, we can seek an explanation for his behavior. The contrast class is made up of the other occasions when we have studied straight through until dinner. We would thus be asking “Why did he take a break rather than study through til dinner, as he has done on X, Y, Z other occasions?” The distinguishing feature that made today different than those other occasions was the arrival of the ice cream truck and his desire for ice cream. The feature may be the correct explanation, but, as in the case of the syphilitic patient, doesn’t strongly predict taking a break given the interest and background understanding – as a lactose intolerant student, we wouldn’t expect Dave to take a break from work on the basis of seeking out ice cream.

Thus there are two main ways for a reason to be weak, or bad. First, it can fail to distinguish the event from the contrast class. For instance, being burnt out fails to distinguish between my weekend activities given my friend’s interest above.

Second, it can succeed at distinguishing amongst the contrast class but not do so strongly, given the context. Thus, while the arrival of the ice cream truck is perhaps the best explanation for Dave’s behavior, it does not fare as well as a reason as the arrival of the ice cream truck would be for Jonah, his non-lactose intolerant friend. In both cases, the feature could succeed in the sense that it did the best job of differentiating the action from the context-determined contrast class. However, because of Dave’s lactose intolerance in

the background understanding we see that this doesn't raise the likelihood of taking an ice cream break as much and thus is a less strong reason than in Jonah's case.

Consider another example of an agent acting on a bad reason. Monique, driving from San Diego to Riverside, takes the 5 and the 91 freeways (thus taking a strange route to get to our campus – basically going through LA). We scheduled her to give an academic talk to the department and thus have an interest in her behavior as a visiting scholar coming in from San Diego. When we hear of her route, our interest and background understanding of the routes in Southern California and our scheduling make us wonder what her reasoning was for taking the unconventional route to get to our campus. Thus, we can inquire for an explanation for her action *qua visiting academic*: “Why did Monique drive on the 5 and 91 instead of the 215 (the route that all reasonable visiting academics take)?”

Upon the standard investigation common to most of our examples thus far (we ask Monique), we find that Monique's reason for taking the route she did was that she thought it was fastest. This is the correct explanation – it distinguishes Monique's behavior from that of the other visiting academics, who did not think that taking the 5 and 91 were fastest. However, such a reason is a bad one because it conflicts with our background understanding. (Recall this as a factor from examples in Chapter 4.) Our background understanding includes information about the traffic in LA and the directness of the route that involves getting on the 215 and staying on it for over an hour. Though Monique's reason succeeds at explaining her behavior, the mechanics of van Fraassen's view allows for us to capture the way in which it is a bad explanation: the consideration conflicts with our background understanding.

Conclusion

In this chapter, I have focused on articulating how we can bring in a theory of explanation from the philosophy of science literature to fill in our picture of the practical domain. In the next chapter, I will turn to evaluating this picture. I will show that this articulation stacks up against the virtues of the standard view discussed in Chapter 1: it illuminates the support relation in a way that also sheds light on the role reasons play in deliberation and justification. I also show how this view meets the desiderata from Chapter 3.

Chapter 6: A Genuine Alternative to the Standard View

In the previous chapter, I laid out how we can apply van Fraassen's pragmatic theory of explanation to the practical domain. Here I will show that my view of practical reasons meets the desiderata laid out in Chapter 3 and address some concerns. I conclude by pointing out some virtues of this explanation-based view of practical reasons.

Section One: Virtues of the Standard View

The main concern with an explanation-based view of practical reasons will be how well it stacks up with the virtues of the standard view. Recall that one of the main reasons the standard view has been the standard for so long is that it tidily accounts for how considerations play their role in deliberation and in justification. Therefore, in this section, I will address possible objections regarding these roles that reasons play.

There are points of interest in capturing both the deliberative and justificatory roles. In the justificatory, we can be concerned with whether I have truly captured what it means to be unjustified, and we can have a related worry regarding being able to capture the relationship between justification and morality. As far as capturing how reasons play their role in deliberation, we can be concerned with modeling how agents decide how to act and how we relate to the reasons we have.

Section 1.1 Justification

Justification concerns good reasons. We use the quality of our reasons to defend our behavior to one another. I claim that reasons fundamentally make sense of performing actions, while the standard view of practical reasons has value play this role. For the standard view, a consideration will be a good reason when there is value in the action, and this easily tracks our intuitions surrounding justification: you are justified in performed actions that are good to do.

Further, the standard view adapts well to the way that justification fits with our moral intuitions: most find it noteworthy that a prime domain in which we find actions important to defend or criticize is the moral domain. When I am concerned that someone acted in an unjustified fashion, often it is because I think they acted immorally. Consider the sorts of behaviors we ask “Why would you *do* that!?” in a justificatory mode. Such cases typically are when someone has been disrespected, or some duty or obligation has been neglected.

Because the standard view has value front and center, it can easily accommodate those cases where moral values are at stake: moral values are simply a subset of the value that is at play in all reasons talk. Agents deal with reasons by taking things to be valuable according to the standard view and justification deals with value. Moral justification deals with moral values, and thus is tidily a subset of the domain of considerations that agents work with, and could be criticized for working with.

In this subsection, I consider how my view answers the burden of accounting for bad reasons and dealing with the domain of moral reasons. I’ll take the concerns in turn.

On my view, an agent takes a consideration to be a reason when she takes it to make an option more likely than a set of alternatives, selected by context. We would say an agent is unjustified when her reason is a bad one; I suggest the way in which reasons are bad reasons is according to standards of explanation. Proponents of the standard view of practical reasons could be concerned with whether I have captured what it means to be a bad reason. Advocates of the standard view can be concerned with whether the explanatory view I propose here captures the ways in which we criticize one another's reasoning for being unjustified.

I suggest that my view can capture when an agent's reasons are bad, and thus when an agent is behaving in an unjustified fashion. Recall the example of Suzanne, who was deciding what career to pursue. She considers reasons to accept her job offer at the Cordon Bleu in France. When considering how Suzanne is making her choice of future career, you could think she is deciding for bad reasons and worry that I seem to be putting too many considerations on relatively equal footing. For instance, some of these considerations (that she likes funny hats), can seem intuitively like bad reasons on which to ground job choices, yet I don't seem to have the machinery to deal with that here: if Suzanne ends up moved by her liking of funny hats, then that is the good explanation for her career choice, yet it doesn't seem like a sensible choice. If the chef path is the only career option with funny hats, this will do good explanatory work because it will be a good distinguisher amongst the options she was considering, and therefore be a good reason. This can seem counter-intuitive, because isn't an account that evaluates reasons supposed to capture when choices fail to be sensible?

Yet what kind of criticism can I lob here with the pragmatic account's machinery? In other words, any silly aspect of being a chef could be seen as a reason to be a chef, yet we would imagine criticizing this person for making her choice on such grounds. Where in my picture can I capture this phenomenon?

On the pragmatic account I'm offering, the prime kind of rational criticism my picture structures is that the reason *doesn't do a good job differentiating* between the action and a class of alternatives. This, therefore, is the way that the account will handle less-than-reasonable choices: if a feature does a good job explaining, as chef-dom's funny-hattedness does in the case above, it is because it did a good job differentiating between becoming a chef and the alternatives that were determined by the agent's interests and background understanding. If you are evaluating Suzanne's reason for pursuing the culinary career and you think making such a choice on hat-grounds is unjustified (you think it is unreasonable, silly, etc.), the way that the pragmatic account will capture this is relative to a context where hat-grounds doesn't differentiate between being a chef and a contrast class.

I think this standard for criticism **does** capture something relevant in our intuitions regarding justification. Consider what you could say to this character - your friend, sister, or daughter - who made their life choice based on hat-potential: "What, would you have gone to *clown college* if they'd accepted you!?" Note that this is suggesting that the consideration would fail to distinguish between being a chef and pursuing other careers that you think the job-seeker should not pursue: it is pressing on the standard, (or interest/context, in the vocabulary of the pragmatic account) that Suzanne was using to decide on her career.

Using Suzanne's reason (the career involves funny hats), you press that other careers that she wouldn't adopt would also be supported: perhaps it was a contingent feature of the contrast class Suzanne was considering that something reasonable was distinguished by funny hatted potential, but to see that this is a silly reason, consider other contexts to see what the consideration of careers with funny hats would lead you to.

Thus your reasons-evaluation can be captured with the pragmatic model: you are bringing to bear a contrast class where "I'll get to wear a funny hat" doesn't distinguish pursuing the culinary career well: a contrast class that includes clowns and other funny-hatted options (priests? Cowboys?) brings out what you see as problematic about Suzanne's reason.

This brings us to the next relevant issue regarding justification: how will different perspectives relate? I noted above that often we consider morality to be the (most?) relevant domain when we are concerned with justification. For my pragmatic view, there are a number of contexts that can be applied to any action scenario. The standard view of reasons connects reasons to value, which intuitively may come in degrees and typically is an objectively framed concept. With a pragmatic framework, how are we to understand the relationship between different contexts of evaluation when I've stressed that these different contexts can result in different evaluations (one context can result in a reason being strong, while another weak)? When we are concerned with justification, we often take some perspective to be privileged, as in the above case with Suzanne's evaluator. Also, this is how morality is often considered. The demands of morality are taken to trump other possible demands or values.

Thus, another potential worry with the pragmatic view is how to handle cases where different interests or standards for evaluation seem to conflict with one another. Consider the following case: Marjorie is an employee on a luxury yacht out for a long cruise.¹ During a deck party where her duties include remaining in a particular area serving beverages, she alone notices one of the guest's children has fallen overboard. In such a case, it seems that there are contrasting standards – that of her job and that of “morality” in some sense – for what to do. It seems like using pragmatic standards for good and bad reasons will produce the wrong result, because we would likely want to say that Marjorie should save the child, yet there are some interests here (the interests of her career, say) that may not produce this result. The pragmatic model leads to a contextual relativism that will make the trumping picture of morality elusive.

If Marjorie swims out to save the child, from the perspective of her job, she is unjustified and this can seem like an unintuitive result. However, recall the case of the Florida lifeguard who was fired for saving a drowning man in 2012.² In life there are cases where interests do conflict – where there are reasons that exist from the perspective of one set of interests that conflict with the reasons that call out from the perspective of another. This account accommodates this feature of our reasoning lives.

This may seem counterintuitive to those that want there to be an objective, absolute stance regarding what reasons demand, or those to consider the demands of morality to be the “final say” on what there is reason to do. Note that there still may be a sense in which

¹ I thank Taylor Cyr for this press and discussion surrounding the example.

² Tomas Lopez was patrolling Hallandale Beach, north of Miami, in 2012.

morality is the “final say” or the trump-card regarding what we have reason to do. With the pragmatic model of evaluating reasons for action, the way we understand the support of reasons is in terms of the interests and concerns we bring to bear, so when we bring to bear interest and concern with the morally right thing to do (however conceived), this may be tacitly an attempt to try to take on a broad perspective, or a perspective that includes or transcends others in a distinctive way.

Agents often disagree not about whether a reason counts in favor of an action given a context (which action is supported given an interest and background understanding), but they rather disagree about which context is the appropriate one to use to determine what to do in a particular scenario. In cases where my interest in friendship is part of a context where one action is supported (spending time supporting our friend, say) but my interest in my career supports another (investing time on a project for my job), the individual contexts may not settle the matter of what to do. Deciding how to decide between these two options is often difficult exactly because we see the force of support that each of the perspectives brings to bear.³ In order to decide between them, another higher order interest or context will be required. When agents disagree about what to do, often it is because one person thinks that one context is more important and the other person favors the other context. However, note that in order to deciding how to decide, we need to take up a further context in order to adjudicate between reasons.

³ See Velleman’s “Deciding How to Decide” in Velleman (2000) for a discussion of using practical reason to determine the principles of practical reason. (He, of course, does not have a model of reasoning based around “Why?”-questions as I present here.)

Section 1.2: Deliberation

When we consider practical reasons, typically this is in a context of deciding what to do. Reasons count in favor of performing actions, and this is how we deal with them – determining which actions to perform. The standard view of practical reason is an articulation of a picture of determining what to do that has us tracking what would be good to do – it articulates a vision of agents as value-trackers. My view is that agents have interests and background understanding that constitute the context of deliberation and that reasons distinguish amongst the context-determined possibilities: reasons are difference-makers, and in determining what to do, we fit our potential actions into our understanding of ourselves and the world. Similarly to how we located the best explanation in the scientific cases by investigating given our understanding of how the world works, in our deliberative perspective, our own agency is part of our background understanding, and we fit the action we pursue into our understanding of how the world makes sense to unfold.

A further question in the domain of the deliberative perspective would be a view's ability to capture the scenario when an agent is moved by a reason she knew to be a bad one. This is a puzzle common to all views in moral psychology, for the agent is behaving in a way that doesn't make sense by her own lights. To take something to be the thing to do, yet do something else, is a bit mysterious.

On the standard view, an agent takes a consideration to count in favor of performing an action in virtue of the good in performing the action. It is thus difficult to make sense of an agent who takes there to be more value in one action than another yet still perform the less valuable action: the weight of reason is in favor of one action in virtue of the value.

On my view, we can make sense of this phenomenon a little better, at least. When a reason is stronger than another, it is often relative to a context. The agent may be acting against reason, but we have the mechanisms to characterize what is happening here: when an agent acts on a reason that was weaker, she was acting on a reason relative to a different context. Thus, she was aligning herself with the interests that made up that context. So, though she still was practically irrational in the sense of acting for a weaker reason, we can say more than this about the case on my view. We can make sense of the agent's behavior a bit more, and I take this to be a virtue. There is more to the agent's moral psychology here.

Section 1.3 Support

One concern could be that throughout this discussion of how my explanation-based view of practical reason stacks up against the standard view, I have moved too quickly past the first virtue that the standard view exhibits: it illuminates the "support" relation. To some, it may seem that I have pushed the question of support back from how a consideration supports performing an action to how a consideration distinguishes between an option and a contrast class. The worry, then, is whether I have addressed this virtue of the standard view of reasons in my alternative.

In taking a consideration to support performing an action, the standard view of practical reasons claims that the agent is taking the consideration to point to what is good in performing the action. I have claimed that this is illuminating this relation in a way that sheds light also on how reasons play their deliberative and justificatory roles.

My own view claims that the agent takes considerations to point to the intelligibility in performing the action, a view that is shared by others that reject the standard view⁴. In applying a view of explanation to the practical realm, I fill in what it means to take a confirmation to point to the intelligibility of performing an action: the consideration makes the action more likely than a set of alternatives, constructed by context. Providing an account of explanation illuminates the support relation – it is the distinguishing between the action and the alternatives that are context sensitive. In addressing the above concerns, I hope to have addressed the question of how the view stacks up to how well the support relation sheds light on the deliberative and justificatory roles that reason plays.

Section Two: Meeting the Desiderata

In Chapter 3, I engaged with two theorists who reject the standard view of reasons by suggesting that practical reasons are primarily explanatory, Setiya and Velleman. These accounts were instructive because they begin to fill in how we can reject the standard view of practical reason by having an explanatory picture of taking a consideration as a reason for action. The key difference between the two views is how they see their explanatory pictures fit with the justificatory role that we take reasons to play. In the previous chapter, I evaluated the different approaches to develop desiderata for a successful explanatory view, and here I will show how my view meets this standard.

⁴ Such thinkers include Setiya and Velleman.

For Setiya, there are two ways an agent can take a consideration to count in favor of performing an action. It is not necessary to take the consideration to be pointing towards something valuable in performing the action in order to take it to be counting in favor or supporting, and thus, he rejects the standard view of reasons. His view is that when an agent takes there to be good reason (justificatory reason) to perform an action, this will be pointing towards the good in the action. Setiya thereby has the view that the standard view of practical reason applies to justification, but that taking a consideration to add justificatory support is not a necessary aspect of taking a consideration to be a reason, which is to take the consideration to add explanatory support.

For Velleman, these different functions cannot be separated in this way. When an agent takes a consideration to support performing an action, she is taking it to add explanatory support to performing the action. He takes it that this is just what it is to take a consideration to add justificatory support. On Velleman's view, this doesn't involve taking the action to be good and he rejects the standard view of practical reason.

We can understand the difference between these two accounts as follows: for Setiya, an agent can take a consideration to be a reason without taking it to add to the justification for acting at all, whereas for Velleman, there is no difference in the explanatory and justificatory support an agent takes a reason to provide. From this characterization, I argued that both accounts had too strong a relationship between the way that we take considerations explain and justify performing actions: taking a consideration to render performing an action intelligible is not completely independent of taking it to justify

performing the action, as Setiya claims, but nor does it say all there is to say about justification as Velleman's view may contend.

From the conclusion of my discussion in Chapter 3, therefore, I have desiderata for an account of how agents take considerations to be reasons for action. Taking a consideration to be a reason is a matter of rendering the action intelligible. This means taking it to have some justificatory force, but the agent can take the justificatory force to differ in degrees. The agent must be able to take the degree of justification for performing her action to differ from the degree of intelligibility there is behind performing the action.

By preserving the possibility that the agent takes the degree of justification to differ from the degree that the consideration makes sense of performing the action, I press on Velleman's habit of paraphrasing back and forth between explanation and justification. In order to preserve the intimate relationship between taking a reason to make sense of performing an action but also to allow for the nuanced relationship that resists the simplistic paraphrasing, we need to elaborate on the possibility of how these two forces can come apart from the agent's perspective.

On my view, justification is a matter of being a good or bad reason. When we offer justification and when we are concerned with our own justification we are evaluating the strength of our reasons. Because reasons are explanatory, justification just is a matter of applying the standards of explanation to evaluate whether the reason was a good reason. Returning to a pair of examples of action explanation from the previous chapter can illuminate this relationship. Lactose intolerant Dave took a break from his work because the ice cream truck arrived. This consideration – that the ice cream truck arrived – explains

Dave's behavior. Applying the standards of explanation provides us with the degree of justification in this case. If we had an interest in explaining Dave's behavior, say, we were going to meet up with him later in the evening for dinner and wonder why he's doing things that will make him late, that the ice cream truck arrived successfully distinguished between Dave taking a break and continuing to get work done, or packing up and going home to work out before we meet, or doing other things to be sure to be on time. The arrival of the ice cream truck explains why Dave took a break, but recall that it doesn't receive comparatively high scores according to our standards of explanation. Our background understanding includes Dave's lactose intolerance, which means that an ice cream truck's arrival, though distinguishes taking a break from the other activities in the contrast class, doesn't make a strong distinction. Especially when we think of the distinction that it would make for lactose-tolerant and ice cream-lover Jonah. According to this model, Jonah has better reason for taking a break than Dave. The strength of the reason is how well the consideration does according to these standards of explanation. In Dave's case, the reason doesn't justify as well because it conflicts with our background understanding.

So, in a sense, justification just is explanation, and my view parallels Velleman's above. Crucially, on my view, I outline how an action can make sense to different extents depending on the context. Recall the differing contexts of explanation that make an explanation stronger or weaker. The agent can appreciate that this can be the case by considering different contexts from within her perspective. An agent can note that

performing an action can make more sense relative to one interest than relative to another,⁵ and thus consider the reason to make sense relative to some context but be a bad reason relative to another.

To see how this might work, we can consider some cases where the agent acts on reasons she considers to be bad reasons yet nevertheless add intelligibility to her performing the action. Recall the cases from Chapter 3, characterized in Setiya's terms:

Pete the Puritan – Pete thinks sexual desire is the work of the devil and considers any pursuit of such behavior to be unjustified. However, he does experience the average spectrum of human urges and sees some considerations (the attractiveness or availability of some potential partner, say, or how good he has been in depriving himself up to this point) as reasons to indulge in these desires. He sees these considerations as reason to indulge, but he does not believe them to justify indulging. (Watson (1975):210)

Academic Al – Al takes great pleasure in the life of a philosopher, constructing arguments about abstract positions and engaging in debates with other sharp minds of his day. It gives him a sense of power he very much enjoys, though he sees the life of a philosopher and the pursuits that make it up to be empty of meaning and not worthwhile in the least. He sees considerations as reasons to engage in

⁵ This is technically true relative to a context, which includes interest and background understanding. However, when we consider the agent's perspective, it would be striking if she had different background understanding when considering the reason for acting in different contexts.

philosophy while at the same time, he doesn't believe these considerations do anything to justify participating in the activities of philosophers. (Setiya (2010): 27 adapting Burnyeat (1980): 76.)

Setiya the Smoker – Setiya, though he has smoked for years up to this point, considers smoking to be completely unjustified. He is going to quit at midnight tonight, and while he sees this as a reason to smoke a pack of cigarettes before then, he does not take this consideration (i.e., that he is quitting at midnight) to do anything to justify smoking the pack of cigarettes.

In each of these cases, what makes them difficult for both Setiya's and Velleman's views to accommodate is that the agent takes there to be a consideration that makes sense of performing the action, and the justification that the consideration adds seems to be out of sync with the degree of sense. In Chapter 3 I suggested that the behaviors make more sense than the degree to which they are justified, and charged that neither view could accommodate two assessments of degree of support. I now claim that this is because to make these two assessments, different contexts are brought to bear in assessing the consideration's support.

Take the case of Setiya the Smoker. In Setiya's articulation, he takes smoking to be "completely unjustified". This is a common way of speaking about behaviors we take there to be a serious lack of support for, often because they invoke a self-conception that we don't endorse. On my view, this scenario would be accounted for by a context where that he is going to quit at midnight does a horrible job of differentiating between an action and

a contrast class. Setiya the Smoker sees having a cigarette on the basis of the fact that he's quitting at midnight as unjustified: that he's quitting is a bad reason to smoke.

Setiya the Smoker, in his awareness of health risks, seems like he'd be pretty likely to adopt some deliberative perspectives that would set up contrast classes that would result in this reason being bad. He could, for instance, be interested in deciding what to do *qua role model for impressionable progeny*, or *qua agent with health risks*, or *qua self-controlled adult*. We could model his thinking in terms of "Why?"-questions of explanation and see the weakness of the reason in question. In order to ask a well-formed "Why?"-question, we would need to include a contrast class against which the potential action would be explained. In the case of Setiya seeking an explanation for why he, *qua agent with impressionable progeny*, would smoke a cigarette, say, the relevantly similar cases of actions would be those that an agent concerned with being a role model may adopt in these circumstances, and those would make up the contrast class.

Thus, the question would be "Why smoke a cigarette rather than (chew nicotine gum), (family games night), (go out with non-smoking friends), (have an indulgent dinner), etc.?" The contrast class is made up of actions that Setiya the Smoker (stipulatively) considers someone who is concerned with the impression he is making on his kids would do that evening. The reason Setiya the Smoker takes to be bad (to do nothing to add to the justification of smoking) is that he's quitting at midnight.

However, he also does see reasons to indulge. He's a classic conflicted character, in the sense that even though he has strong commitments that lead him to disvalue smoking as a part of a healthy and fulfilling life, he still is a smoker, and smokes for reasons.

Smoking occurs to him as something to pursue, and he does take the consideration that he is quitting at midnight to support smoking. I suggest that this is relative to a different context: a different context of evaluation has this reason as a good reason than had the reason as a bad reason above.

While Setiya the Smoker *qua role model* saw the lack of good reason to smoke, Setiya the Smoker *qua smoker* sees reason to smoke, appreciates how intelligible it is. Considering what to do *qua smoker*, Setiya's contrast class will be options that a smoker would perform, and his deliberation may be modeled, "Why smoke a cigarette rather than (chew nicotine gum), (chew regular gum), (go for a run to put off a craving?), other things smokers would do in an evening?" The consideration that he is quitting at midnight would point towards smoking a cigarette rather than these other options, which could be done anytime but smoking can only be done now. This consideration does a better job explaining relative to this interest than to the previous interest.

The agent can see that a consideration does a good job at making sense of a behavior (which is language we typically use liberally, no matter the interest we use to implicitly construct a context), while doing a bad job at justifying (which is language we typically use specially, in contexts that are "at stake", or relevant to defend our behavior at any particular time). In these cases, there are contexts where the considerations do a fine job of explaining the agent's behavior. Those contexts happen to be those that we typically wouldn't press for the agent to provide her reason for evaluation. Relative to contexts where we have an intuitive understanding of the action being *at stake* – contexts where smoking is bad (where agents are role models), sex is undesirable (where agents are ascetic

puritans), and philosophy is useless (narrow-minded academia) – the consideration doesn't do a good job at distinguishing the behavior from others the interest suggests.

From the agent's perspective, when she takes a consideration to be a reason for performing a behavior, it will have import for justification. Taking the consideration to make sense of the behavior will be taking the consideration to justify the behavior. Thus, it won't be possible to take there to be reason to perform an action and take the action to be completely without justificatory support. From within that context, it has justificatory support. This doesn't mean that there isn't another context where the consideration *doesn't* do a good job at explaining, and thereby leaves the agent open to justificatory presses.

In this way, my account meets the desiderata laid out in Chapter 3. When an agent takes there to be a reason to perform an action, she takes it to justify performing the action. However, on my view, the agent can recognize multiple contexts, and thus multiple strengths of the consideration, and this captures how we can take our reasons to make more sense of performing the action than we would take them to justify performing the action.

Section Three: Virtues of Reasons as Explanation

Besides meeting the desiderata of the previous section and thereby capturing the way that agents can take their actions to have different degrees of support, the explanatory view of practical reason has other benefits over the standard view.

First, this model can capture the commonsense intelligibility of a jerk's unjustified behavior. A virtue of having a model of reasons for action that captures the justificatory and explanatory function of reasons in terms of the sense-making "Why?"-questions of van

Fraassen's pragmatic account is that it captures the tension we feel with most unjustified behavior. When someone behaves in a way that is unjustified, often we feel that it doesn't make sense, yet there's a way in which it is frustratingly intelligible. Take, for instance, cases where someone is a jerk, but where we can easily predict that they would be a jerk in that situation.

When someone flouts norms, it feels comfortable to say "that just doesn't make sense", or "why would they *do* that?!" – these are common enough phrases we use in response to such behaviors. However, consider also, in cases like this where we could have predicted the behavior, it *does* make some sense. I think this is the right result – it both makes sense and doesn't, and in the way/for the reasons I am attempting to model. It makes sense because we have a good explanation for it, flat out – the perspective of the agent makes pursuing the jerk-y action the thing to expect. It doesn't make sense because we (as reasoners) have a good understanding of what agents would do, and this wasn't it. So, yes, it makes sense that this person was a jerk (they're a jerk), but in a sense it doesn't make sense to be a jerk because being a jerk isn't supported by reasons, and our expectations of one another. Because we do often reflect on what to do, and ask one another for our reasons for action, we frequently act according to reasons: we tend to make sense. We expect one another to behave according to shared reasons, and we can sometimes see and predict when people will behave out of sync with those expectations. Both responses are best understood in terms of explanation – in terms of sense-making.

A second virtue of an explanation-based view is that it better captures the first person perspective on perverse cases. Because on my view, we take considerations to count

in favor by explaining rather than pointing to the good in performing an action, I can accommodate a broader range of human behavior without categorizing it as pathological *qua agency*. When agents pursue behaviors for reasons, it is not necessary that they take the behavior to be good. In the cases where it seems that the agents take their behaviors to be bad, or evaluatively neutral, this is not a misfire, mistake, or deviation for the agent as a reasoner.

As reflective and social creatures, we often adopt aims that are complex and sometimes at odds with what we take to be good. An explanation-based view of taking considerations as reasons can more accurately capture these complexities without attempting to force them into a positive evaluative framework.

Consider the nuances that develop in our self-destructive behavior. It is common enough to experience the impulse step out over a precipice or to swerve the wheel dangerously while driving. That impulse often feels disconnected from our evaluative frameworks, an interesting motive perhaps leftover from early in our evolutionary history, but in our motivational experience nevertheless. At times when we behave in complex ways – perhaps picking a fight with a friend, or burning a bridge with a colleague or yelling at an employer – it is at least plausible that our reflection on our behavior as we perform it is not in terms of catharsis or the good that will come from the behavior. In other words, when we behave in self-destructive ways at times, is it not plausible that we are not taking there to be something good in what we are doing? Indeed, at points we may be a bit horrified or disappointed in our behavior even as we are responsive and acting

intentionally.⁶ A non-evaluative view of reasons helps to capture that this behavior can indeed be performed for reasons and that the strangeness or pathological aspect is in the higher order qualities such as the evaluative commitments, not the intentionality.

Finally, a virtue in having an account of agency that doesn't necessarily involve positive evaluation is that it could allow for a more continuous understanding of the spectrum of "rational" agents. While the standard view has it that in order to take a consideration to be a reason, an agent needs to have the capacity to take actions as good, an explanation-based view of practical reasons does not require evaluative capacities for reasoning. The capacity to take a feature to distinguish amongst options is the fundamental shape of this view, as opposed to taking a feature to be good.

Conclusion

The explanation-based model of practical reason allows for a distinction between taking a consideration as a reason for action and taking the action to be good. Having explanation as the basis of support allows for more nuance in evaluating pathological cases such as self-destructive behavior and perversity, and it allows for a more continuous picture of reasoning between human and non-human animals. It also allows for multiple frameworks of justificatory evaluation for a single action, thereby accommodating our

⁶ Something like this may be what Freud discusses as significant about the human condition in Freud (1990). He posits a tendency in human psychology to will ends beyond our pleasure, broadly speaking, incorporating the possibility of perversity. By allowing for more complex motivational structures in his picture of human psychology, he captures behaviors that go beyond the standard view. He attributes most of these tendencies to what may be characterized as pathological agency, so classifying Freud's picture of moral psychology is a difficult task, but adding to the complexity of our self-understanding was a leap forward for his time.

intuitions regarding the intelligibility of some unjustified behavior and the tensions involved in the desiderata cases from previous chapters.

Bibliography

- Alvarez, M. (2009). Acting intentionally and acting for a reason. *Inquiry* 52 (3):293-305.
- Anscombe, G. E. M. (2000). *Intention*. Cambridge, MA: Harvard University Press.
- Aquinas, St. T. (1964). *Summa Theologiae*. London: Black Friars.
- Aristotle (1984). *Complete Works of Aristotle*. (J. Barnes, trans.) Princeton, NJ: Princeton University Press.
- Aristotle. (1999). *Nicomachean Ethics*. (T. Irwin, trans.) Indianapolis, IN: Hackett Publishing Company, Inc.
- Armstrong, D. (1983). *What is a Law of Nature?* Cambridge, UK: Cambridge University Press.
- Augustine. (2008). *The Confessions*. Oxford, UK, Oxford University Press UK.
- Bagnoli, C., (2002). Moral constructivism: A phenomenological argument. *Topoi* 21: 125–138.
- Bagnoli, C. (2009). Practical necessity: The subjective experience. In *Value and Ontology*. Edited by B. Centi & W. Huemer. Frankfurt: Ontos-Verlag: 23-43.
- Baker, Derek (2015). Akrasia and the Problem of the Unity of Reason. *Ratio* 28 (1):65-80
- Benn, S. I. (1985). Wickedness. *Ethics* 95 (4):795-810.
- Brady, Michael S. (2003) Valuing, Desiring, and Normative Priority. *Philosophical Quarterly* 53 (211): 231-242.
- Bratman, Michael (1987) Intention and Evaluation. *Midwest Studies in Philosophy* 10 (1):185-189.
- Bratman, Michael (2007). *Structures of Agency*. New York, NY: Oxford University Press.
- Bratman, Michael (2014) *Shared Agency: A Planning Theory of Acting Together*. Oxford, UK: Oxford University Press.
- Bromberger (1966) Questions. *Journal of Philosophy* 63 (20):597-606.

- Broome, John (2007). Wide or narrow scope? *Mind* 116 (462):359-370.
- Broome, John. (2013). *Rationality Through Reasoning*. Chichester: Wiley-Blackwell Publishing.
- Burge, Tyler. (2003). Perceptual Entitlement. *Philosophy and Phenomenological Research* 67 (3): 503-48.
- Burnyeat, Myles F. (1980). Aristotle on learning to be good. In Amélie Oksenberg Rorty (ed.), *Essays on Aristotle's Ethics*. University of California Press: 69-92.
- Carroll John, W. (2008) *Laws of Nature*. Cambridge, UK: Cambridge University Press.
- Carroll, J. W. (2016.) Laws of nature. *Stanford Encyclopedia of Philosophy*
- Carroll, Lewis (1993) *Alice's Adventures in Wonderland*. Dover Publications.
- Chan, David K. (1995). Non-Intentional Actions. *American Philosophical Quarterly* 32 (2):139 - 151
- Chandler, J. (2013). Contrastive confirmation: Some competing accounts. *Synthese*, 190(1):129–138
- Chang, R. (2005.) Parity, internal value, and choice. *Ethics* 115 (2): 331-350.
- Chislenko, Eugene (draft 2014) Akkratic Action under the Guise of the Good. Academia.edu
- Copp, David (1995) *Morality, Normativity, and Society*. Oxford, UK: Oxford University Press
- Dancy, Jonathan (2004). Enticing Reasons. In R. Jay Wallace, Philip Pettit, Samuel Scheffler & Michael Smith (eds.), *Reason and Value: Themes From the Moral Philosophy of Joseph Raz*. Oxford, UK: Clarendon Press.
- David, M. (2005). Truth as the primary epistemic goal: A working hypothesis. In M. Steup & E. Sosa (Eds.), *Contemporary debates in epistemology* Oxford, UK: Blackwell Publishing.
- Davidson, Donald (1980) *Essays on Actions and Events*. Oxford, UK: Oxford University Press.

- Dressel, Ashley. (2014). Aquinas and Later Scholastics on Willful Wrongdoing. UC Irvine: Philosophy. Retrieved from: <http://escholarship.org/uc/item/8x19s71c>
- Dworkin, Ronald (1993) *Life's Dominion: An argument about abortion and euthanasia*. New York, NY: Knopf
- Enoch, D. (2006). Agency, shmagency: Why normativity won't come from what is constitutive of action. *Philosophical Review*, 115(2), 169–198.
- Evans, Matthew (2010). A Partisan's Guide to Socratic Intellectualism. In Sergio Tenenbaum (ed.), *Desire, Practical Reason, and the Good*. Oxford, UK: Oxford University Press.
- Fischer, John Martin (2012). Responsibility and autonomy: The problem of mission creep. *Philosophical Issues* 22 (1):165-184
- Fitelson, B. (2012). Contrastive Bayesianism. In Blaauw, M., editor, *Contrastivism in Philosophy: New Perspectives*. Routledge.
- Frankfurt, Harry G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy* 68 (1):5-20.
- Frankfurt, Harry G. (1977). Identification and externality. In Amelie Rorty (ed.), *The Identities of Persons*. University of California Press.
- Frankfurt, Harry (1987). Identification and Wholeheartedness. In Ferdinand David Schoeman (ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge University Press.
- Freud, Sigmund (1990) *Beyond the Pleasure Principle*. Norton and Company.
- Graham, Peter J. (2012). "Epistemic Entitlement". *Noûs* 46 (3):449-482.
- Goor, D.G. (Writer). (2010, January 21). Summer Catalog [Television series episode]. In *Parks and Recreation*. NBC.
- Hanson, Norwood Russell (1958). *Patterns of Discovery an Inquiry Into the Conceptual Foundations of Science*. Cambridge University Press
- Harman, Gilbert (1986) *Change in View*. MIT Press.
- Harman, Gilbert (1997) Practical reasoning. In Alfred R. Mele (ed.), *Review of Metaphysics*. Oxford University Press. pp. 431--63.

- Hawkins, Jennifer. (2008). Desiring the bad under the guise of the good. *Philosophical Quarterly* 58 (231):244–264.
- Helm, Bennett W. (2001). *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge University Press.
- Hempel, Carl (1965) *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press
- Heuer, Ulrike (2004) “Raz on Values and Reasons” In R. Jay Wallace, Philipp Pettit, Samuel Scheffler & Michael Smith (eds.), *Reason and Value*. Oxford University Press.
- Hieronymi, Pamela (2005). The Wrong Kind of Reason. *Journal of Philosophy* 102 (9):437 - 457.
- Jaworska (2007a) Caring and Internality. *Philosophy and Phenomenological Research* 74 (3):529-568.
- Jaworska, Agnieszka (2007b). Caring and full moral standing. *Ethics* 117 (3):460-497
- Jaworska, Agnieszka (1999). Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value. *Philosophy and Public Affairs* 28 (2):105-138
- Kant, Immanuel (2011). *Groundwork of the Metaphysics of Morals: A German-English Edition*. Cambridge University Press.
- Knobe, Joshua and Erica Preston-Roedder (2009) The Ordinary Concept of Valuing. In Ernest Sosa & Enrique Villanueva (eds.), *Metaethics*. Wiley Periodicals.
- Kolodny Niko (2003). Love as valuing a relationship. *Philosophical Review* 112 (2):135-189.
- Kolodny, Niko (2005). Why be rational? *Mind* 114 (455):509-563.
- Kolodny, Niko (2007). State or process requirements? *Mind* 116 (462):371-385.
- Korsgaard, Christine M. (1996). *The Sources of Normativity*. Cambridge University Press.
- Korsgaard (2003) The dependence of Value on Humanity. In Jay Wallace (ed.), *The Practice of Value*. Oxford University Press.

- Korsgaard, Christine (2005). Acting for a Reason. *Danish Yearbook of Philosophy* 40.
- Korsgaard, Christine M. (2008). *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford University Press.
- Korsgaard, Christine M. (2009). *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press.
- Lebar, M. (2008) Aristotelian constructivism. *Social Philosophy and Policy*, 25(1): 182–213.
- Lenman, J. (2010) Humean Constructivism in Moral Theory, in *Oxford Studies in Metaethics*, 5: 175–193.
- Lewis, D. (1983.) New work for a theory of universals. *Australian Journal of Philosophy* 61: 343-377.
- Lewis, D. (1986.) *Philosophical papers, volume II*. New York: Oxford University Press.
- Mantel, Susanne (forthcoming). Three Cheers for Dispositions: A Dispositional Approach to Acting for a Normative Reason. *Erkenntnis*:1-22.
- McAninch, Andrew James (2015). Acting for a Reason and Following a Principle. *Ethical Theory and Moral Practice* 18 (3):649-661.
- McDowell, John. (1979). Virtue and Reason. *The Monist* 62 (3):331-350.
- McHugh, C. (2013). Normativism and doxastic deliberation. *Analytic Philosophy*, 54(4), 447–465.
- Mele, Alfred R. (2003). Philosophy of Action. In Kirk Ludwig (ed.), *Donald Davidson*. Cambridge University Press.
- Millar, A. (2009). How reasons for action differ from reasons for belief. In S. Robertson (Ed.), *Spheres of reason* (pp. 140–163). New York, NY: Oxford University Press.
- Milton, J. (2003) *Paradise Lost*, London, England: Penguin Classics
- Nagel, Thomas (1970) *The Possibility of Altruism*. Oxford Clarendon Press.

- Owens, D. J. (2003). Does belief have an aim? *Philosophical Studies*, 115(3), 283–305.
- Pettit, Philip (ed.) (2004). *Reason and Value: Themes from the Work of Joseph Raz.* with R. Jay Wallace Samuel Scheffler and Michael Smith. Oxford Univ. Press.
- Plato. (1997) Apology, in Cooper, John M. & Hutchinson, D. S. (eds.) *Plato: Complete Works*. Hackett Publishing Co.
- Quinn, Warren. (1993) *Morality and Action*. Cambridge University Press
- Rawls, John (1971) *A Theory of Justice*. Harvard University Press.
- Raz, Joseph (1999). *Engaging Reason: On the Theory of Value and Action*. Oxford, UK: Oxford University Press.
- Raz, Joseph (1999) Explaining normativity: On rationality and the justification of reason. *Ratio* 12 (4):354–379.
- Raz, Joseph (2001). *Value, Respect, and Attachment*. Cambridge University Press.
- Raz, Joseph (2004) The role of well-being. *Philosophical Perspectives* 18 (1):269–294
- Raz, Raz (manuscript) “On the Guise of the Good”
- Reath, Andrews (2006), *Agency and Autonomy in Kant's Moral Theory*, Oxford: Clarendon Press
- Reath, Andrews (Forthcoming) Constructivism in *The Encyclopedia of Ethics*, New York: Garland Press.
- Rowling, J.K. (1999) *Harry Potter and the Prisoner of Azkaban*. New York, NY: Random House
- Salmon, Wesley (1971). *Statistical Explanation & Statistical Relevance*. University of Pittsburgh Press.
- Salmon, Wesley, (1984), *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Salmon, Wesley (1997) Causality and Explanation: A Reply to Two Critiques. *Philosophy of Science*, 64: 461–477

- Scanlon, T.M. (1998). *What We Owe to Each Other*. Belknap Press of Harvard University Press.
- Scanlon, T.M. (2014). *Being Realistic About Reasons*. Oxford University Press.
- Scheffler, Samuel (ed.) (1988). “Agent-Centered Restrictions, Rationality and the Virtues” in *Consequentialism and its Critics*. Oxford University Press.
- Scheffler, Samuel (2011) “Valuing” in *Reasons and Recognition: Essays in the Philosophy of T.M. Scanlon* edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman, Oxford Scholarship online.
- Schlosser, Markus E. (2012). Taking Something as a Reason for Action. *Philosophical Papers* 41 (2):267-304.
- Schroeder, Mark (2007a). Reasons and agent-neutrality. *Philosophical Studies* 135 (2):279 - 306.
- Schroeder (2007b). Teleology, agent-relative value, and 'good'. *Ethics* 117 (2):265-000.
- Schroeder, Mark. (2007c). *Slaves of the Passions*. Oxford University Press
- Schroeder, Mark. (2008) Having Reasons. *Philosophical Studies* 139 (1):57--71.
- Seidman, Jeffrey (2009a). Caring and the Boundary-Driven Structure of Practical Deliberation. *Journal of Ethics and Social Philosophy* 3 (1).
- Seidman, J. (2009b). Valuing and caring. *Theoria* 75 (4):272-303.
- Setiya, Kieran (2007). *Reasons without Rationalism*. Princeton University Press
- Setiya, Kieran (2010). Sympathy for the devil. In Sergio Tenenbaum (ed.), *Desire, Practical Reason, and the Good*. Oxford University Press.
- Shah, N. (2003). How truth governs belief. *Philosophical Review*, 112(4), 447–482.
- Shah, N., & Velleman, J. D. (2005). Doxastic deliberation. *Philosophical Review*, 114(4), 497–534.
- Shute, Sara (1992). The Idea of a Reason for Acting. *International Studies in Philosophy* 24 (1):133-133.

- Snedegar, Justin (2017). *Contrastive Reasons*. Oxford University Press.
- Stocker, M. (1979) Desiring the Bad: An Essay in Moral Psychology. *The Journal of Philosophy*, 76, 739-40.
- Stocker, Michael. (2004) Raz on the Intelligibility of Bad Acts, in R. Jay Wallace (ed.) *Reason and Value: Themes From the Moral Philosophy of Joseph Raz*. Oxford University Press.
- Street, Sharon (2008) Constructivism about Reasons. *Oxford Studies in Metaethics* 3: 208–245.
- Street, Sharon (2010) What is Constructivism in Ethics and Metaethics? *Philosophy Compass*, 5: 363–384.
- Tenenbaum, Sergio (2007). *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge University Press.
- Tenenbaum, Sergio (2013). Guise of the Good. In Hugh LaFollette (ed.), *The International Encyclopedia of Ethics*. Wiley-Blackwell.
- Tuomela, Raimo (2006). On Acting for a Reason. *Acta Philosophica Fennica* 78:187
- van Fraassen, Bas (1980) *The Scientific Image*. Oxford University Press.
- van Fraassen, Bas (1988) The pragmatic theory of explanation. In Joseph C. Pitt (ed.), *Theories of Explanation*. Oxford University Press.
- van Fraassen, Bas (1989) *Laws and Symmetry*. Oxford University Press.
- Velleman, David (1989). *Practical Reflection*. Princeton University Press.
- Velleman, David (1992a) What Happens When Someone Acts? *Mind* 101 (403):461-481.
- Velleman, J. David (1992b). The Guise of the Good. *Noûs* 26 (1):3 - 26.
- Velleman, J. David (2000). *The possibility of practical reason*. New York, NY: Oxford University Press.
- Velleman, J. David (2003) Narrative explanation. *Philosophical Review* 112 (1):1-25.
- Velleman, J. David (2005). The self as narrator. In *Autonomy and the Challenges to Liberalism: New Essays*. Cambridge: Cambridge University Press.

- Velleman, J. David (2005). *Self to Self: Selected Essays*. Cambridge University Press.
- Velleman, J. David (2009). *How We Get Along*. Cambridge University Press.
- Wallace, R. J. (2010). Reasons, Values and Agent-Relativity. *Dialectica* 64 (4):503-528.
- Wallace R. Jay (2012). Constructivism about normativity: some pitfalls. In Jimmy Lenman & Yonatan Shemmer (eds.), *Constructivism in Practical Philosophy*. Oxford University Press.
- Watson, Gary (1975). Free agency. *Journal of Philosophy* 72 (April):205-20.
- Watson, Gary (1987). Free action and free will. *Mind* 96 (April):154-72.
- Watson, Gary (2013) Psychopathic Agency and Prudential Deficits. *Proceedings of the Aristotelian Society* 113 (3pt3):269-292.
- Wedgwood, R. (2002). The aim of belief. *Philosophical Perspectives*, 16(s16), 267-97.
- Woodward, J., (2000) Explanation and Invariance in the Special Sciences. *British Journal for the Philosophy of Science*, 51: 197-254.
- Zalabardo, J. L. (2010). Why believe the truth? Shah and Velleman on the aim of belief. *Philosophical Explorations*, 13(1), 1-21.
- Zaragoza, Kevin (2006). What happens when someone acts compulsively? *Philosophical Studies* 131 (2): 251-268.