GETTING EXCLUSIONARY REASONS RIGHT

GETTING EXCLUSIONARY REASONS RIGHT

By NICOLE AREIAS, H.B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree of Master of Arts

McMaster University MASTER OF ARTS (2022) Hamilton, Ontario (Philosophy)

TITLE: Getting Exclusionary Reasons Right

AUTHOR: Nicole Areias, H.B.Sc. (McMaster University)

SUPERVISOR: Professor S. Sciaraffa

NUMBER OF PAGES: vi, 74

LAY ABSTRACT: Getting Exclusionary Reasons Right investigates the distinctiveness and coherence of the idea of an exclusionary reason—a reason to not act for other reasons (i.e. promises, rules, commands, etc.). It first defends exclusionary reasons as reasons to '*refrain from acting for some reason(s)'.* Understood in this way, exclusionary reasons are relevant where it matters not just what we choose to do, but *how* we choose to do it. Promises, rules, commands, etc. are features of the world that make it valuable to or justify our not acting for otherwise good reasons when they apply. It then considers what '*refraining from acting for some reason(s)'* consists in. While exclusionary reasons are thought to be reasons to have motivations of certain kinds, they are not reasons to choose to be motivated in some way. They are instead, reasons *that determine for us* the reasons we ought to act for. It is argued that while there are some instances where determining what we have reason to do is up to us, exclusion is not one of them. Getting Exclusionary Reasons Right concludes by considering the implications the account offered herein has for rationalist approaches to obligations and authority.

ABSTRACT: Getting Exclusionary Reasons Right offers a defense of exclusionary reasons as originally conceptualised by Joseph Raz. Exclusionary reasons are second-order reasons to refrain from acting for some reasons and are used to explain the ordered nature of practical normativity, and the various normative concepts that are said to follow from it, i.e. mandatory rules, rule-following, authority, and promises to name just a few.

Exclusionary reasons differ from other kinds of defeaters in that they exclude *valid* reasons, i.e. reasons that still justify or make eligible the actions they count for. According to Raz, this is because excluded reasons are defeated not *qua* reasons, but as reasons we can *act for*, or that motivate, which explains why exclusionary reasons are reasons to *refrain from acting for a reason*. However, the coherence and distinctiveness of the idea of an exclusionary reason—understood in this way—has faced serious challenges. I take up these challenges in what follows.

Chapter one presents a coherent account of exclusionary reasons as reasons to *refrain from acting for a reason,* or to 'not-φ-for-*p*'. It both clarifies the sense in which exclusionary reasons concern motivations and motivating reasons, and rejects alternative accounts according to which exclusionary reasons have as their object other normative reasons. It is argued that when they are understood as excluding some considerations as reasons that can rationally motivate, exclusionary reasons confer value on or point to an agent's not acting for otherwise valid reasons. That is, they justify our not being responsive to certain values on some occasions.

Chapter two vindicates the notion of *acting-for-a-reason* on which Raz's account relies. It considers objections which claim that not acting for otherwise valid reasons presupposes a level of control over our reasons and motivations that is incompatible with the rational constraints on attitudes (beliefs, intentions, etc.), and shows how exclusionary reasons, as they are restated in chapter one, avoids them. Perhaps surprisingly, it is argued that instances where exclusionary reasons are relevant, when properly understood, are not instances where reasoning about what we ought to do involves choice.

Getting Exclusionary Reasons Right concludes by considering the implications the account offered herein has for rationalist approaches to obligations and authority. Namely, it makes clear how fully rational agents can ever be moved to act for, or out of an awareness of their obligations.

ACKNOWLEDGEMENTS

I want to begin by expressing my sincerest thanks to my supervisor and mentor, Dr. Stefan Sciaraffa, for his guidance throughout this project. The trust Dr. Sciaraffa placed in my abilities is what allowed me to take on this ambitious topic, while his enthusiasm, patience, knowledge and incisive feedback helped make the project the best that it could be. He has showed unwavering support, and continues to challenge me to grow as a writer and philosopher, and for this I am immensely grateful.

I would also like to express my deepest appreciation for Dr. Wil Waluchow for serving as second reader on this project, but more importantly for inspiring a love for philosophy that I might otherwise have never known. I am indebted to Dr. Waluchow whose support and encouragement is the reason I was able to dedicate two very rewarding years to this discipline.

Thank you to my family, for the many sacrifices that have made the privilege that has been my education possible, and to my partner Tristan for being there always. I share this achievement with all of you, and I cannot imagine these past two years without your love and encouragement.

To my friends, too many to name, for tolerating my thesis misery, and for reminding me to continue to enjoy the life happening outside this project, I thank you. I am especially grateful to have found such a good friend in my cohort, Ayo, without whom I may never have survived grad school during the pandemic. I could not be more grateful for our countless conversations, reading groups, vocational reflections, and your help with coursework.

For all that these individuals have done, I am thankful.

Table of Contents

‘[…] there are dense connections among all parts of Raz’s system. One cannot understand the nature of law unless one understands that it claims authority; one cannot understand authority unless one understands the way it serves reason; one cannot understand reason unless one understands the complexities of action and value, and so on. Approaching Raz’s philosophy is like diving into a pool that has no shallow end and undercurrents everywhere.’

– Leslie Green

Introduction

We should, and normally do, only pursue things in life because they are valuable. I drink coffee because it helps keep me awake through my seminars, and I exercise in order to maintain my cardiovascular health. That "it helps keep me awake through my seminars" is a feature of drinking coffee that makes drinking it valuable, just as "maintaining one's cardiovascular health" is a feature of exercise that makes exercising valuable. When the good-making features an agent recognizes in some action make performing the action intelligible from her point of view, and can guide its performance, they are reasons for so acting. That 'it will help keep me awake' is a reason to drink coffee, and when I drink coffee out of an awareness of or for this reason I can understand why I acted as I did. That is, I take my reaction to this reason to be appropriate. Provided there are no other reasons against doing so (i.e. that grabbing one would make me late for my seminar), my reaction is rational or justified, just as a failure (or refusal) to (ever) exercise, in spite of my awareness of the benefits exercise has for my cardiovascular health, is irrational or unjustified.

According to Joseph Raz, to understand how reasons for acting make claims on us as rational agents is to understand how they are normative. Our rational capacities enable us (fallibly) to recognize value in some options as reasons for taking them, and to respond appropriately—to take those options for those reasons (Raz, 2011, p. 4). Only when facts that constitute a case for certain actions relate to our rational powers in this way are they normative. His understanding of the nature of normativity undergirds his account of practical reason and reason conflicts. In deliberating about what to do, rational agents consider the cases her reasons present for and against acting, and determines what, in light of them all, she is justified in doing.

When her reasons conflict, such that they recommend incompatible actions, she typically weighs these considerations against each other, and determines which of them present the strongest case for the action they recommend. For example, in deciding whether or not to stop for coffee I might take my reasons for getting to seminar on time to outweigh the reasons I have for drinking coffee. Rational agents also consider the validity of the reasons they take themselves to have. If instead I learn that the only coffee available is decaf, I determine that 'that it would help keep me awake' is no longer a feature of drinking the coffee that makes doing so valuable. In other words, I take the consideration to undercut the reason's status *qua* reason. Our reasons conflict in other ways too. Consider the following commonplace examples:

> *I promise a friend that I will not, and thus have reason to not, move her car from its parking spot in the driveway while she is away, though that I need to use my own car which is blocked in recommends that I move it.*
>
> *Where the law dictates the speed at which I can drive, I have reason to drive at or below the limit, yet that I am running late to work recommends that I go above that limit.*
>
> *My employer orders that I stay past closing, such that I have reason to do so, though the fact that the restaurant at which I was planning to eat will be closed by then recommends leaving early.*

In each of these cases, I have reason to act so as to keep my promise, adhere to the law, and obey the command in spite of the reasons that suggest I do otherwise. That is, in each of these cases what I ought to do is not move the car, drive at or below the speed limit, and stay past closing.

The other reasons that apply, however persuasive a case they make for conforming to them, are in some sense defeated. They are defeated, but not because the promise, law or command presents a stronger case for the conflicting action's non-performance. My promise to my friend, for example, is not an especially weighty reason that, when weighed against the fact that moving the car would allow me to access my own, defeats it. Nor is it because these reasons can no longer do their normal justificatory work. Intuitively, your promise to your friend does not change the fact that your need for your car makes moving your friend's car good or valuable—it just is a good-making characteristic of the action. Instead, the promise constrains the eligible courses of action by making those that would allow me to keep my promise obligatory (i.e. not moving the car).

How should we understand the special normative work these reasons do? Joseph Raz famously suggests that we attribute their effects to exclusionary reasons. These are reasons to *refrain from acting for other reasons.* Raz appeals to the idea of an exclusionary reason to explain various aspects of practical reasoning, including mandatory rules, rule-following, authority, and promises to name just a few. In addition to the reason it gives me to not move her car, my promise presents a reason to, and thus requires that I refrain from acting for reasons (i.e. 'that I need to use my car which is blocked in') when doing so would preclude me from keeping my promise, just as the law requires that I refrain from acting for reasons for going above that limit (i.e. 'that I am running late to work'), and the command requires that I refrain from acting for reasons for leaving early (i.e. 'that the restaurant will be closed'). Put another way, these reasons exclude the otherwise valid reasons with which they conflict insofar as they can be a part of an agent's motivational set. Call this the motivational interpretation. However, the coherence

and distinctiveness of the idea of an exclusionary reason—understood in this way—has faced serious challenges. Razian exclusionary reasons, opponents tell us, are about the motivations an agent ought to have when she acts. Not only does this misconstrue the nature of promises, rules, and commands, which are thought to merely require conforming behaviour, it presupposes a level of control over our reasons and motivations that is generally denied.

This project offers a defense of exclusionary reasons as originally conceptualised by Joseph Raz. It is, in some sense, a response to the call made by Leslie Green as noted above, though far narrower in scope. By appealing to the nature of normative reasons in general, and of practical reasons in particular, as laid out in his later works, it offers a rehabilitation of the idea of an exclusionary reason according to which exclusionary reasons, when properly understood, are normative reasons. Its aims are twofold. The first chapter clarifies the sense in which exclusionary reasons are practical reasons to refrain from acting for some of our reasons, or to 'not-φ-for-*p*'. On this interpretation, they are not about the motivations an agent ought to have, but instead reasons that exclude some considerations as reasons that can rationally motivate. They confer value on or point to an agent's not acting for otherwise valid reasons, or, put another way, her being irresponsive to some of the values on offer. The second chapter considers how rational agents respond to or act for exclusionary reasons and shows that we can do as they require—that we can 'not-φ-for-*p*'. It argues that refraining from acting for reasons does not involve having or bringing about certain intentions, and identifies what, instead, 'not-φ-ing-for-*p*' consists in. It then demonstrates its compatibility with the rational constraints on attitudes (beliefs, intentions, etc.). I conclude by considering the implications the account offered herein has for rationalist approaches to obligations and authority (i.e. where these are understood in

terms of pre-emptive reasons). The account makes sense of the intuitions we have about our obligations, and suggests how it is that rational agents can be moved to treat the reasons they generate pre-emptively.

Chapter One: Exclusionary Reasons as Normative Reasons

1. Introduction

One of Joseph Raz's signal contributions to practical philosophy is his proposal that practical reasons exist at different normative levels (Raz, 1975). While first-order reasons are reasons to take or to refrain from taking some action, second order reasons are reasons to act or refrain from acting on some set of reasons *p*. This tiered nature of practical reason offers a more nuanced picture of reason conflicts, and the various normative situations that arise when reasoning about what to do. In positing a second normative order, Raz makes way for a distinct kind of defeat, namely exclusionary defeat or exclusion. Negative second-order reasons, or exclusionary reasons, are reasons that exclude *valid* reasons, i.e. reasons that still justify or make eligible the actions they count for. According to Raz, this is because excluded reasons are defeated not *qua* reasons, but as reasons we can *act for*, or that motivate, which explains why exclusionary reasons are reasons to *refrain from acting for a reason*. Raz appeals to the idea of an exclusionary reason to explain various aspects of practical reasoning, including mandatory rules, rule-following, commands, authority, and promises to name just a few.

However, theorists have recently argued that the notion of *refraining from acting for reasons*, on this interpretation is incoherent. They argue that exclusionary reasons understood in this way are about the motivations an agent ought to have, which misconstrues the nature of authoritative directives, commands, promises, etc. As the objection goes, Raz's account unnecessarily introduces motivations and motivating reasons into issues better understood in terms of justification and normative reasons. As a remedy to this putative shortcoming, Adams

(2020) proposes a new account of exclusion as a matter of justification, where exclusionary reasons are not reasons that bear on some act, 'not-φ-for-*p*', but instead reasons that bear directly on other normative reasons and determine whether and how they count (Adams, 2020, p. 239).

In this chapter, I offer a defense of Raz's account of exclusionary reason as reasons to 'not-φ-for-*p*'. In section 2, I explain Raz's account of exclusionary reasons. In section 3, I close the gap between the motivational and justificatory interpretations of exclusionary reasons by appealing to Raz's idea of the *normative-explanatory nexus,* and clarify the sense in which Raz's account is concerned with motivation.[1] I argue that exclusionary reasons, when they are understood as excluding some considerations as reasons that can rationally motivate, or as reasons that we can *act for,* are reasons that confer value on or point to an agent's not acting for otherwise valid reasons (i.e. to 'not-φ-ing-for-*p*'). In section 4, I explicate a key implication of this defense, namely that exclusionary reasons are relevant where it matters not just what we choose to do, but *how* we choose to do it. I conclude by considering the role both accounts attribute to exclusion in explaining value conflicts, and suggest that the construal of Razian exclusionary reasons as reasons that justify our not being responsive to certain values on some occasions is the more attractive account.

2. Raz's Account of Exclusionary Reasons

[1] According to the 'nexus,' normative reasons can explain a rational agent's behaviour when the agent acts out of an awareness of (i.e. *acts for*) those reasons. As such, the normativity of reasons derives in part from their ability to justify or make a case for an action, and from their ability to guide or motivate agents who are reason-responsive in taking the action. See 'Reasons: Explanatory and Normative' in Raz's (2011), *From Normativity to Responsibility*. Oxford University Press.

Raz defines a practical reason as a reason for some agent, A, to take or to refrain from taking some action, φ. Any such reason is a fact that identifies some point or value in taking the action, φ, for A and therefore counts in favour of φ-ing (Raz, 2011, p. 13). Exclusionary reasons and their differences from first-order practical reasons are best understood in the context of conflicts of reasons and their resolution, to which we now turn.

First, consider the case of first-order practical reasons. These are reasons for or against taking some action, say, φ. An agent acts correctly (i.e. rationally), or as the first-order reason requires, when she takes the action for which it is a reason. When the actions her reasons recommend are incompatible, her reasons conflict. Consider an agent who has some reason *p* to φ, and some reason *q* to not-φ. The actions for which *q* and *p* are reasons, not-φ-ing and φ-ing, are not compossible such that by conforming to one an agent fails to, or cannot, conform with the other. Here, her first-order reasons conflict completely, and determining what she ought to do involves assessing which reason of those in conflict prevails. If her reason to not φ, *q*, presents a stronger case for not φ-ing than *p* does for φ-ing, and there is no possibility of conforming with both reasons (i.e. she cannot both φ and not-φ), reason *q* can be said to *outweigh* reason *p*. But this is not to say that the defeated reason is any less of a reason for performing the action it supports. That is, reason *p* in our example remains a relevant consideration for the agent in support of φ-ing.

The term ‘outweighs’ naturally invites talk of ‘the balance of reasons’ or reasons competing in weight as a metaphor for first-order conflict. As the metaphor goes, all of the applicable first-order reasons taken together constitute the balance of reasons for a given action,

and establishing what we *rationally* ought to do involves assessing the relative strengths of various conflicting reasons and determining 'what there is most reason to do' (Gert 2007). The idea of weight, or strength, however, plays no explanatory role in Raz's account. Rather, it is but one useful way of illustrating the notion of direct conflict between first-order reasons and the idea that in some such cases of direct conflict one reason defeats the other.

By contrast, second-order practical reasons are reasons for or against φ-ing on the basis of certain first-order reasons. An exclusionary reason is a negative second-order reason: it is a reason to not φ for some set of reasons *p*. Like any first-order reason, an exclusionary reason has as its object an action, namely 'not-φ-for-*p*', such that an agent acts correctly, or as the exclusionary reason requires—when she acts, but not for reason *p* (i.e. 'not-φ-for-*p*'s). These reasons, being considerations additional to the set of relevant first-order reasons, do not participate in the figurative balance of reasons. They instead operate on a different normative level; they constrain which of those first-order reasons we are justified in acting for. Take, as an example, Mary who has promised Dave that she will decide where to send their daughter Jane on the basis of Jane's educational needs alone.[2] Her promise does not bar Mary from considering tuition costs or how her daughter's attendance at some schools might further Mary's own career. It requires only that she not base her decision on these considerations, valid though they may be. Perhaps Mary finds that on the balance of all first order reasons—that school A is more affordable, that school A furthers her own career, and that school be better satisfies her daughter's needs—school A prevails. However, because of her promise, Mary is not justified in acting for finance-related or career-related considerations; these reasons are excluded by her

[2] I borrow this example from Whiting (2017). Against Second-Order Reasons. *Noûs*, *51*(2), 398-420.

promise not to take those reasons into account. In excluding finance-related and career-related considerations, as her promise to Dave requires, school B prevails.

How do we explain the exclusionary effect her promise has? Her promise does not outweigh her finance-related and career-related reasons. That is, its addition does not tip the scales against choosing school A by presenting a stronger case for its non-performance. Because any exclusionary reason *e* is a reason to not-φ-for-*p*, and not a reason to not-φ, it is possible to conform with an exclusionary reason *e* and still φ—namely by φ-ing for some other reason. Thus it defeats certain reasons for φ-ing only partially.[3] Mary's promise defeats the finance-related and career-related considerations that favor her choosing school B only to the extent that she cannot both conform with these reasons (i.e. by choosing school B) and conform with her promise not to act for these reasons.[4]

Note that excluded reasons (like reasons that are outweighed) are not defeated *qua* reasons. The distinctive feature of exclusion is that defeated reasons remain reasons that count in favor of an action. Mary's promise, for example, does not make it the case that her financial- and career-related reasons no longer favour choosing school B since school B really does make the most financial sense or is the more strategic choice for her career. This explains why her

[3] Conflicts are partial when at least one of the reasons in conflict need not be completely frustrated. A reason is completely frustrated if the good that conforming with it would secure will not be achieved. Often, however, the good that conformity with the reason would achieve can be secured, completely or partly, in some other way; in the case of exclusionary reasons, by φ-ing for, say, q or r. If there is an equally effective alternative way of realising the good, the reason need not be frustrated at all. See Postscript to the Second Edition: Rethinking Exclusionary Reasons in *Practical Reasons and Norms* (Raz, 1999).

[4] If the non-excluded reasons favor choosing school B, her finance- and career-related reason that also favor choosing school B are not defeated as reasons for choosing school B. They are defeated when the non-excluded reasons favor choosing some other school instead. While there are some exceptions, in general reasons are excluded to the extent that they prevent the agent from conforming with what she ought to do *all things considered.*

excluded reasons can justify regret, guilt, and other responses that Mary might have when she instead chooses school A because of her promise: they remain reasons that count in favour or recommend choosing school B even when excluded, albeit not ones that she is justified in acting for.

There is at least one other conflict-type: undercutting. While exclusion and outweighing represent conflicts between practical reasons, undercutting involves a distinct species of considerations that *affect* reasons conflicts.[5] When an undercutting consideration conflicts with some reason *p* to φ, they attack *p*'s status *qua* reason such that it no longer counts for or even recommends the action it is a reason for—such that it is no longer a reason to φ. This kind of conflict is sometimes also understood in terms of weight: for a reason to be undercut is for some other consideration to reduce its weight. Where its weight is reduced to zero, the reason is disabled, to use Dancy's (2004) terminology (see also Pollock 1986). But, this is again by way of analogy. They are not (practical) reasons in the Razian sense, but mere facts that negatively affect, or undermine, the ability of other reasons to prevail in cases of conflict by either fully or partly undercutting a fact's status as a reason. While exclusionary reasons and undercutters are often conflated, they are not the same. Unlike exclusionary reasons, undercutting considerations *do* affect a putative reason's status as a reason. If, instead of her promise to Dave, Mary were to discover that school A has several hidden fees such that it is no longer more affordable than school B, she no longer takes 'that it is more affordable' to count in favour of school A.

[5] Raz does not take undercutting considerations to be practical reasons, since they do not identify some point or value in taking an action. As such, defeat takes on a different meaning in this context. A reason that is undercut is defeated *as a reason*, not as the reason that fails to prevail in a conflict. Undercutters defeat reasons without recommending an alternative option thus there need be no conflict with another reason.

To say that an exclusionary reason defeats a reason *p* for φ-ing is not to say that it outweighs, undercuts, or disables that reason but only to say that it requires that you not-φ-for-p. In positing a second-order of practical reasons, Raz's suggestion is that what ought to be done is not always determined by the balance of first-order reasons *alone*. At least some of the time, in addition to the first-order reasons for action that apply to us, we also have reason to satisfy, that is, not frustrate, other considerations which require that we act for only some of those reasons. This explains a more capacious sense of the idea of what ought to be done *all things considered*, and it justifies the deviation from first-order reasons where they require it (Raz, 1999, p. 185). In acting from the point of view of the second-order considerations and the unexcluded first-order considerations, the agent is said to act in accordance with *all* the reasons that apply.

3. Exclusion and the Normative-Explanatory Nexus

We can posit a third kind of defeat only if an exclusionary reason to not-φ-for-*p* and a fact that entails that not-*p* (i.e. an undercutter) defeat *p* in distinct ways; that is, if they generate distinct normative situations. As the discussion above shows, the distinction rests on the fact that, unlike undercutters, exclusionary reasons do not affect another reason's ability to count for the action it favours. That they exclude *valid* reasons for action is the very idea of exclusion. Yet it is unclear how exclusionary reasons do so. How can we explain the distinct normative situation they create?

In his 'Authority, Law, and Razian Reasons' (1989), Moore offers three distinct ways of understanding exclusionary reasons as reasons *'to refrain from acting for some reason',* and exclusionary defeat. He suggests that there is a sense in which an exclusionary reason

(i) changes the reasons (motives) for which we may act (i.e. 'the motivational interpretation'),

(ii) changes what counts as a right-making characteristic of an action (i.e. the 'justificatory interpretation'), and

(iii) changes which of our reasons we can consider or contemplate (i.e. the 'decision-procedure interpretation').

Returning to our example, Mary's promise does not require that she refrain from thinking of these considerations at all as (iii) seems to suggest, but only that she does not act for those reasons (Raz, 1999, p. 185). Mary is welcome to consider how the various options will affect her career, so long as she does not base her decision on these reflections. The 'decision-procedure interpretation', then, can be easily dismissed.[6]

On the justificatory interpretation, reasons are excluded from counting, or from doing their normal justificatory work. However, exclusionary reasons on this reading look a lot like undercutting conditions and threatens to muddle the distinction between them (Moore, 1989, p. 857). As the previous discussion shows, an exclusionary reason does not defeat a conflicting

[6] For a comprehensive account of this version of exclusionary reasons see Anthony Hatzistavrou's (2012) "Motivation, Reconsideration and Exclusionary Reasons."

reason *qua* reason. Because it is unclear, how, on the justificatory interpretation, excluded reasons remain reasons, Raz denies it.

To be sure, the interpretation that can sustain the difference between exclusionary reasons on the one hand and undercutting and first-order reasons on the other is the motivational interpretation. According to this reading, exclusionary reasons are reasons that change the reasons we are justified in *acting for* (Raz, 1989, p. 1156). They are reasons "for not being motivated in one's action by certain (valid) considerations" (Raz, 1999b, p. 185). The solution to the puzzle, then, is as follows: because they are defeated as reasons that can motivate, and not as reasons that can justify, excluded reasons still count for the action they recommend. Mary's financial and career-related reasons are still reasons even when she cannot make her decision on their basis.

Much criticism has been leveled against exclusionary reasons understood in this way. N.P. Adams (2020) argues that an account of exclusionary reasons concerned with how we should be motivated "is at best odd" (Adams, 2020, p. 238). According to Adams, the motivational account of exclusionary reasons seems to require that when, say, φ-ing we be motivated in certain ways—that we ought to have certain motivations and not others. Raz's appeal to motivating reasons suggests that where exclusionary reasons are relevant compliance with one's reasons is required, yet this misconstrues the nature of practical reasons in general. All practical reasons are reasons to conform. Adams notes, as Raz himself does, that there is no blemish, no defect, in acting for any motivation that would lead the agent to do what she ought to, or what she is justified in doing (Raz, 1999b, p. 182). Instances where we have reason to

comply are uncommon, and certainly do not include those involving authoritative directives, commands, promises, etc. Taking Mary as an example, she does all that she has reason to do when she chooses a school on education-related grounds alone even if she does so merely on accident, by coincidence, etc. There is no additional demands on how she should be motivated when she does so—only that she does it. In a similar vein, those in authority who issue rules are concerned not with the motives with which citizens conform their behavior to the rules, but with the conforming behavior itself.

Because instances where exclusionary reasons are relevant are instances involving demands over how an agent ought to act and not how she ought to be motivated in doing so, motivating reasons, says Adams, are relevant only insofar as they are related to or derivative of changes in the underlying normative reasons, i.e. derivative of changes in how an agent ought to act (Adams, 2020, p. 239). It follows that exclusionary reasons defeat reasons as motivating reasons not merely because they ought not serve as her motive in φ-ing, but because it is a reason that no longer justifies φ-ing at a particular time or in a particular context. She ought not φ for that reason because it no longer justifies φ-ing.

Adams also notes that an account of exclusionary reasons that appeals to motivating reasons must clear additional conceptual hurdles. Following those who make similar claims (Moore, 1988; Essert, 2011; Whiting, 2017), Adams asserts that the notion of acting-for-a-reason (i.e. φ-for-*p* or not-φ-for-*p*) on which Raz's account relies is untenable. Acting for a reason seems to require an ability to bring about a motivational set consisting of some reasons and not others, however, our motivations, and more precisely our intentions, do not seem to be under our

control in this way. If acting-for-a-reason is in fact something we cannot choose to do, we cannot have reasons to act-for-a-reason. We will return to objections of this kind in chapter two. Setting them aside for now, let's accept Adam's claim that these objections are decisive against the Razian account. If we cannot have reasons to act-for-a-reason, then Raz's explanation of how excluded reasons remain reasons that count for the action they favour fails. This calls for an alternative account of what makes exclusionary defeat distinct—either one that can sustain this feature or one that concedes that excluded reasons no longer count for the action they favour, and that marks the distinction in some other way.

In his "In Defense of Exclusionary Reasons," Adams takes the latter route, and suggests a new way forward. He proposes a new justificatory account according to which exclusionary reasons do in fact change what counts as a right-making characteristic of an action. Like reasons that are undercut, reasons that are excluded lose their normative force such that they no longer count for the option they are reasons for, though they do so in a distinct manner. Adams suggests that both exclusionary reasons and undercutters are negative second-order reasons, though his understanding of second-order reasons differs from Raz's own account. On both accounts, normative reasons share the same basic structure: they consist of a fact (F1) bearing on an object (i.e. φ). However, while Razian second-order reasons take as their object acting-for-a-reason (i.e. φ-for-*p* and not-φ-for-*p*), Adams' second-order reasons bear directly on other normative reasons. They consist of facts that bear on whether and how another reason (i.e. fact) counts for its object. As such, they represent indirect conflicts between reasons. Negative second-order reasons undermine the reason on which it bears' ability to count for its object—that is, they attack that reason's status *qua* reason. However, undercutters and exclusionary reasons are about different

elements of the target reason and thus represent distinct kinds of indirect defeat. Adams' account of second-order defeat, both exclusion and undercutting, is as follows:

> …There are different kinds of indirect conflict and so different kinds of indirect defeat. Indirect conflict attacks a reason's status qua reason: it attacks the counting relation between fact and object. The higher-order reason (R2) consists of a fact (F2) bearing on another reason (R1/O2), which itself consists in a fact (F1) bearing on an object (O1). We can taxonomize types of indirect defeat according to how F2 bears on R1: is it a fact about F1 or φ? These correspond respectively to undercutting and exclusionary defeat. R2 undercuts R1 when it identifies a feature of F1 that explains why F1 does not count when φ-ing. R2 excludes R1 when it identifies a feature of φ that explains why F1 does not count when φ-ing.
>
> Put another way, facts about the context of choice can undercut while facts about the choice itself can exclude. A reason is undercut when it doesn't count in these circumstances; a reason is excluded when it doesn't count for this kind of choice. Both undercutting and exclusion thus depend on the idea that in a different context, the target reason (R1) counted or would have counted. Then something changes or another reason is brought to bear such that the target reason no longer counts in the same way. We can understand this idea of change in different ways, for example temporally, modally, or as a matter of prima facie and pro tanto reasons. I think all can be relevant but don't have the space to pursue this in detail. What's important is that undercutting and exclusion differ based on what part of the target reason their fact is about. A reason is undercut when the context

> of choice, or the state of affairs, changes; a reason is excluded when the choice under consideration changes. (2020, p. 241)

When the second-order defeater identifies a feature of F1 that explains why F1 does not count when φ-ing, it undercuts the reason. In other words, a reason is undercut when facts about the state of affairs, or the context of the choice, make it the case that the reason no longer counts. Borrowing Adams' own example, in considering her reasons for eating a burger, Kelly recognizes that the fact 'that it is tasty' (F1) counts in favour of eating the burger. Call this reason 'tastiness'. If it is also true that the burger was made with rotten ingredients, this consideration undercuts 'tastiness' as a reason—it renders the burger less tasty or not tasty such that 'that it is tasty' no longer counts for choosing the burger. By contrast, when the second-order defeater identifies a feature of φ that explains why F1 does not count when φ-ing it excludes the reason. That is, a reason is excluded when facts about the choice itself tell us that the reason does not count for this kind of choice (Adams, 2020, p. 241). If instead it is true that Kelly is a vegetarian, this consideration excludes 'tastiness'. It defeats tastiness as a basis for choosing whether to eat the burger without rendering the burger less or not tasty since it is a fact about the sort of choice Kelly is trying to make and not about the burger. Consider how the state of affairs might change while the choice remains the same. If Kelly were instead deciding whether to eat a hot dog or steak, her vegetarian commitment remains a fact about the choice she is making in these instances as well. She also ought not eat the hot dog or steak. It is in this way that excluded reasons lose their normative force without amounting to reasons that have been undercut: facts concerning *the kind of choice* being made make it the case that excluded reasons no longer count without undermining F1 (i.e. the contents of the reason). 'Tastiness' is not excluded insofar as it

supports choosing the burger, but instead insofar as it would make her φ-ing (i.e. her eating) non-vegetarian in character. Thus, Adams offers a version of the justificatory interpretation that, unlike Moore's, can maintain the distinction between exclusionary reasons and undercutting conditions even when both undermine the conflicting reason's ability to justify, and so its status *qua* reason.

But why posit a third kind of defeat at all? As Adams' novel account makes clear, it's because we sometimes want our choices to have "a certain character or instantiate a certain value," and this can play a role in determining what we ought to do (Adams, 2020, p. 236). As his example shows, certain of Kelly's commitments determines *how* she ought to φ (i.e. how she ought to eat) thus determining the choices she can make, and disqualifying some kinds of considerations from mattering. Because exclusionary reasons arise from an intrinsic feature of the choice in question, exclusion plays a constitutive role in the practice, decision, or value in question. This can limit the set of relevant factors without implying a change in the state of affairs. 'Tastiness' may no longer count even when the burger really is tasty since it no longer counts in what the agent ought to do with respect to that particular choice. Part of the appeal of positing exclusion as a distinct kind of defeat, then, is that it captures how, on some occasions, the character of the choice being made matters, and this aspect of exclusion is only clear when we take it to be a matter of justification.[7]

[7] This is how Scanlon understands exclusionary effect that our obligations have. On his competing rationalist framework, obligation-generating reasons exclude reasons such that they lose their justificatory force. Fulfilling our obligations, and 'being moral,' he tells us "involves seeing certain considerations as providing no justification for action in some situations even though they involve elements which, in other contexts, would be relevant" (Scanlon, 1998, p. 157). See his *What We Owe to Each Other*. Belknap Press.

3.1 *Motivating and Normative Reasons*

Adams' novel account rests on the distinction that he argues exists between the motivational and justificatory interpretations. On the motivational interpretation, says Adams, exclusionary reasons involve demands that when we act, we be motivated in certain ways but not others. The suggestion here is that we satisfy exclusionary reasons when we act with certain motives. By contrast, exclusionary reasons on the justificatory interpretation involve the right-making characteristics of the relevant action, and merely require that certain ordinarily right-making characteristics "no longer count in what the agent ought to do" (Adams, 2020, p. 245). The former concerns motivating reasons while the latter normative reasons, and Adams argues for the justificatory interpretation on this basis. However, he appears to miss how motivating and normative reasons are closely related.

His suggestion that "Raz change[s] the subject by introducing motivating reasons into issues that are better understood at the level of normative reasons," is somewhat misleading as it implies that motivating and normative reasons are of distinct kinds (2020, p. 238). Generally speaking, normative reasons are reasons by virtue of their being part of the case for a certain response, by making said response *pro tanto* worthwhile (Raz, 1999a, p. 22); see also Alvarez (2010, p. 13); Dancy (2000, p. 1); Parfit (2011, p. 31); Scanlon (1998, p. 17). Motivating reasons, on the other hand, are those reasons that motivate and are thus responsible for, or that can explain why an agent acted as she did. Misunderstandings of one's reasons, unchecked desires, impulses, and habits are motivating reasons in this sense. An impulse, for example, is the reason for which I purchased a book that I have no real interest in reading, and explains why I did it, just

as a habit, being the reason for my doing so, explains why I habitually bite my nails when I write. However, reasons of this kind do not *rationally* justify the actions they explain. Neither the impulse nor the habit make it good or valuable—and thus worthwhile—to purchase the book, or bite my nails, and so neither presents a case for taking either action. Reasons that justify taking some action, by contrast, do just this. If instead I plan on using the book in my research, there is now a case for purchasing it. That I will use it in my research makes purchasing the book a valuable thing to do—it justifies doing so. Yet, these reasons can, and often do, also motivate a rational agent in taking the action, and thus explain why she acted as she did. When I instead purchase a book *because* I must read it for my research, I take the action out of an awareness of the reason that counts for it. Here, a normative reason serves as my motivation. Differentiating between kinds of reasons, then, is more complex than the normative-motivating dichotomy allows since our motivations can often include normative reasons.[8] In contemporary discussions, the distinctions drawn between types of reasons track the different questions we can ask about them (Dancy 2000; Baier 1958), the contexts in which they are referred to (Raz, 1999a, p. 15) or the varying roles they can play (Alvarez 2010). Again limiting our discussion to practical reasons, a reason can explain an action, justify an action, and guide an agent in taking an action.

By definition, all normative reasons are those that justify the action they count for, though they can play additional roles as well. In addition to justifying an action, some normative reasons can explain the action. Consider the reasons Lauren might have for ordering pizza for

[8] We can think of at least one other class of considerations that are motivating, though not fully normative. These are considerations that an agent mistakenly takes to be related to some value or point, and acts on because she takes it as such. As an example, an agent may act out of a sense of duty when she donates most or all of her life savings to a religious institution while being mistaken about either the value to be had in upholding her duty, the connection that exists between this duty and squandering her savings, or both.

lunch. She can order pizza because she is hungry (R1), and she can order it because it is close by, and thus would allow her to resume her work sooner (R2). Both are normative reasons because both favour her ordering pizza. In this case, Lauren ordered the pizza *because* it is close by—that is, it is the reason that counts in favour of ordering pizza and that lead her to order it, the one that she acted for in ordering pizza. While the fact that Lauren is hungry is a reason that can certainly make sense of her action—since it identifies some value in her ordering pizza—it is not a reason that explains why she does so.[9] R2, in being the reason for which she did it, is the reason that, in Lauren's eyes, counted in favour of her ordering the pizza; it guided her in doing so and so explains why she did it. As such, it is both a normative reason for ordering pizza and the reason that served as her motivation. Contrasting motivating and normative reasons, then, misleadingly suggests that there are two sorts of reasons when there are not (Dancy, 2000, p. 2). As this example shows, at least some motivating reasons just are normative reasons—they are reasons that justify *and* motivate.

The connection between the normative and the motivating, however, runs deeper. It is not just that some normative reasons *do* motivate, but that all normative reasons, by definition, *can*.[10] For how else would reasons, which are mere facts about ourselves and the world around us, move (fully) rational agents to act? Many conceptions of normative reasons and intentional action require that normative reasons at least be capable of guiding the agent in taking the action it counts for (Raz, 2011, p. 26; See also Gibbons 2013; Korsgaard 2008; Lord 2015; Parfit 2011; Kolodny 2005; Kelly 2002; Williams 1981). Facts that are reasons, says Raz, are normative by

[9] It does not, for example, explain why she ordered pizza as opposed to a sandwich or salad.
[10] To say that all reasons can motivate is not to assert that all reasons are reasons for complying. As noted earlier, since all practical reasons are reasons for conforming it is certainly not required that an agent acts for a reason, though at least some of the time they do.

virtue of their connection to Reason, or an agent's rational capacities, which enable her (fallibly) to identify value in some options *and* to respond to them. She recognizes that those aspects of the option that make it valuable are reasons for taking it, and she is then able to do so, to take that option *for* those reasons (Raz, 2011, p. 4). This points to the connection between how normative reasons can favour an action and explain an agent's taking it. Raz, following Bernard Williams (1981), refers to this connection as the *normative/explanatory nexus* ('the nexus'). It states,

> Every normative reason can feature in an explanation of the action for which it is a reason, as a fact that, being recognized for what it is, motivated the agent to perform the action, so that the agent guided its performance in light of that fact. (2011, p. 28)

In requiring that an awareness of the reason includes recognition of it as a reason for the action (i.e. belief in its existence as a normative reason) and that agents are motivated by this awareness, the nexus demonstrates the role normative reasons can play in explaining actions done with an intention or performed for a reason—those that are non-coincidental and non-accidental. Normative reasons can feature in an explanation of the action when an agent capable of recognizing and responding to these facts as reasons is (1) fully rational; and given that she is fully rational, (2) when she is otherwise capable of performing the action. A fully rational agent who recognizes the reasons for φ-ing φs *because* it is justified, and even required by the reason. Being motivated by belief in the existence of a (normative) reason, she explains or justifies the action correctly when she cites the reasons for which she acted.[11]

[11] This is the central case of intentional action. Acting intentionally in this sense is acting for a reason, which involves answering for oneself a sort of "why-question" of justification, i.e. the question of whether so to act by reference to the reasons she considers. See G.E.M. Anscombe's (1957) *Intention,* Oxford: Blackwell. For Raz, we can also act intentionally in a broader sense that is not mediated by reason-responsiveness or justificatory in nature.

Consider Riya, who is running late for an interview. In this case, she takes 'being on time (so as to make a good impression)' to be a reason to run to catch the bus because 'being on time' justifies doing so. She explains her action, running, *in the right way* by citing the value she sees in being on time, and evaluates her response as rational because 'being on time' required it. This is the sense in which, for Raz, the normative has to do with both justification and motivation: reasons that justify explain our actions *as rational agents* when we respond to them.

As such, the nexus explains how normative reasons are made for rational beings like us. Features of the world (i.e. the facts that constitute reasons) make some responses (i.e. actions) appropriate. They make certain actions eligible by making them intelligible from the point of view of the agent (Raz, 1999, p. 26). In being facts that confer value or a point to acting as they recommend, reasons explain, and thus allow agents to understand why they acted as they did.[12] When an agent is guided by the (believed) value she identifies in φ-ing, she is responsive to reasons to φ. That is, she is responsive to good-making features or values in the world and can *act for* them. The motivating role that normative reasons (can) play captures the sense in which we can offer explanations of our actions in terms of value that are rational and not mere rationalizations, and this is the mark of rational agency—at least where an agent's rational powers function well.[13]

[12] Raz labels this the 'classical conception' of human agency which has its origins in Plato and Aristotle's writings. On this tradition, Reason is explained by invoking value. Since valuable aspects of the world constitute reasons, agents acts with an intention if and because she believes that the action is good in some respect and is guided by the value of the action as her reason for performing it (i.e. the Guise of the Good Thesis).

[13] If Riya were to instead explain herself as having run in order to get in some exercise, this would be a mere rationalization. While there might be other justificatory reasons, like the value of exercising, that make running an appropriate or eligible response, they are not features of the state of affairs that moved her to action, and so not one she can cite. An explanation that does so is, in some sense, faulty.

In accounting for the normativity of reasons, then, it is not enough that they justify. Facts that confer value or a point to, and thus require, the action for which they are reasons can explain or figure in an explanation of actions performed with a reason-based intention or purpose, and they must be capable of providing such an explanation (Raz, 2011, p. 26).[14] A reason can do so when it guides or moves an agent to take the action it counts for. Only when reasons meet both of these conditions—when they are reason-giving *and* they can move us to respond as they require—can they "make a difference to agents," for this is how our rational powers help us make sense of and act in light of features of the world around us (Raz, 30).

3.2 *'Refraining from Acting for the Right Reasons'*

The preceding discussion casts Raz's and Adams' respective accounts of exclusionary reasons in a new light. We can now return to the central challenge of exclusionary reasons—explaining the distinct normative situation they create and whether and how excluded reasons remain reasons when defeated. We can also formulate both accounts, the features they share and where they diverge, more precisely.

As the above discussion demonstrated, some motivating reasons are normative reasons that motivate. This helps close the gap between the two interpretations. It suggests that because motivating reasons can be normative, Raz's motivational account is about normative reasons and

[14] While I limit this discussion to practical reasons, the nexus extends to normative reasons for belief as well. Epistemic reasons can also guide and thus figure in explanations of beliefs (Raz, 2011, p. 26).

justification too. But how? The answer lies in Raz's own understanding of normativity. His suggestion is that reasons are normative insofar as they (1) identify value in, thereby making eligible and thus *requiring,* certain responses, and (2) they guide rational agents in, or move them to, respond as the reason requires. To say that a reason requires a response is to say that it makes claims on us as rational agents. An agent acts rationally when she responds to eligible options and irrationally when she fails to do so.[15] And this is how Raz understands normative force: reasons affect us when failure to conform is irrational or a fault on our behalf (Raz, 1999, p. 28). As such, in addition to identifying value in the actions they make eligible, reasons are normative because they move rational agents to act accordingly—to respond to the belief that one has reason to do, or that there is value in so doing.

We can now understand the normative consequences that exclusionary reasons on Adams' and Raz's respective accounts have. On Adams' account, exclusionary reasons are about whether and how certain reasons make actions eligible. Reasons that are excluded lose their normative force because—like reasons that are undercut—they no longer make eligible the action they favour, but not because of a change in the state of affairs. Rather, it is because Mary's promise makes it the case 'that it would be strategic for Mary's career' is no longer a reason not because it is no longer true that it would be the strategic choice, but instead because it is no longer a feature that makes choosing school B valuable, and thus eligible. 'That it would be strategic for Mary's career' remains a fact about choosing school B, albeit not one that can justify choosing it were she to keep her promise. This is tantamount to saying that an excluded

[15] Furthermore, as it has been argued elsewhere, an agent who φs for a valid, undefeated reason which favours φing, is creditworthy (Arpaly 2002; Arpaly and Scroeder 2013; Markovits 2010). Conversely, an agent who fails to do so is blameworthy.

reason to φ is defeated by an exclusionary reason, as a reason that justifies φ-ing, because it is no longer a feature of the response (i.e. φ-ing) that identifies value in it, or that it is no longer valuable for the agent. On Raz's account, exclusionary reasons are about whether and how agents ought to *respond* to certain reasons. An excluded reason's normative force is affected insofar as the exclusionary reason makes it the case that it is no longer a fault that the agent fails to respond when the response is otherwise eligible or justified—it is no longer irrational in spite of the response's eligibility. Here, 'that it would be strategic for Mary's career' *is* a feature that makes choosing school B valuable, and thus eligible. However, even when an agent recognizes it as such, the exclusionary reason makes it the case that this reason should no longer move her to act on its basis. It makes it the case that she ought not be responsive to this value.

As such, on Raz's account, exclusionary reasons are reasons not to be motivated by one's rational reaction to certain reasons or values, and this explains the distinct normative situation they create.[16] While Adams' account implies otherwise, it is not always the case that where an agent ought not act for a reason this is because the option is ineligible (i.e. not justified). Instead, it is sometimes the case that an agent ought not respond because not responding is itself an eligible response. This is the true insight of Razian exclusionary reasons and exclusionary defeat: on some occasions we are justified in not responding to some of our reasons even when those reasons justify, or make eligible the response (i.e. the action) they require.[17] This is not because the exclusionary reason merely shields agents from blame, but because not responding to certain reasons is itself what the agent ought to or has reason to do. What Mary ought do is not respond

[16] Put another way, to say that the excluded reason ought not rationally motivate is to say that the reason ought not explain her φ-ing in this instance. While the response it recommends is eligible, the explanation an agent who acts for this reason would have to offer is not.
[17] This is what is meant by the phrase '*to refrain from acting for some reason,*' or to 'not-φ-for-*p*'.

to the career-related reasons that favour school B, and she acts rationally when she does so—when she refrains from responding to these values even when she recognizes them as such.

This is how Raz brings exclusionary reasons into the justificatory fold. Exclusionary reasons are normative reasons that justify not acting for other normative reasons. That is, they are features of the world that identify value in not acting for some reasons, and thus require it. Promises, rules, and commands are all examples of these features that make not responding to some reasons eligible, and that can guide rational agents in doing so—in not acting for some reasons *because* she promised, *because* of the rule, etc. And an agent acts rationally, or as the exclusionary reason requires, when she does this—when she does not respond as the excluded reason requires. Accordingly, they play a role in determining what we ought to do not by changing the reasons that can justify our acting, but by prescribing not acting for some reasons as what full reason responsiveness requires of us.

Importantly, Raz's account captures the same insight that Adams makes about exclusion: that that our choices should have "a certain character or instantiate a certain value," can in part determine what we ought to do. Like Adams, Raz can maintain that exclusionary reasons are relevant where it matters not just what we choose to do, but how we choose to do it. Consequently, they are about what we are justified in doing *all things considered*, where this *includes* how we do it. In addition to having reason(s) for choosing a school, Mary has reason to do something more. Her promise is a reason to refrain from choosing for some reasons—that is, to refrain from choosing a school *because* it is strategic for her career. She does what she has reason to do when she chooses a school *and* refrains from choosing a school for some reasons.

For Mary, then, it matters not just which school she chooses, but how she goes about choosing the school. She does the right thing when she *acts for* some reasons and not others, but this is not because exclusionary reasons require that we have the 'right' motivations.[18] Nor is it because, as Adams suggests, exclusionary reasons make it the case that some of her reasons are no longer features that make choosing school B valuable or good. Instead, it is because where exclusionary reasons are relevant acting only for some reasons is constitutive of doing that which we ought to do. It is constitutive of *how* Mary goes about choosing a school.

4. Exclusionary Reasons and Value Pluralism

As I have shown, the motivational interpretation, as Adams and others understand it, misses much of what Raz's own account has to offer. Most significantly, it fails to capture the coherent (and compelling) story it can tell about the normative consequences that Razian exclusionary reasons have. But what of Adams' account of exclusionary reasons as reasons that have as their object other normative reasons? At first blush, this appears to be just another way of telling the same story. It seems that both Raz and Adams can explain (1) how, when they are added to the set of reasons that apply, they alter what the agent ought to do all things considered; and (2) why where exclusionary reasons are relevant, it matters not just what we choose to do, but *how* we choose to do it. So, who tells it better?

[18] Instances in which exclusionary reasons are relevant are not instances in which acting on certain reasons and not others is the right thing to do. In instances of this kind, while you may have a variety of reasons (i.e. *p, q, r*) that, if followed, will allow you to conform with what you have reason to do, you are asked to *comply* with certain of those reasons, *p*, even if conforming with the other reasons, *q* and *r*, would satisfy what you have reason to do all the same. Here the right-making feature, is a feature of the motivational set. Think of cases where 'it's the thought that counts'. An agent who has reason to get their loved one a Valentine's day gift does the right thing when, for example, they get flowers because they have reason to gift them (*p*), not when they do so to avoid a fight (*q*) or get into their loved one's good graces (*r*). While giving flowers to avoid a fight or win their favour satisfies the reason you have to get a Valentine's day gift, it is one's compliance with *p* that is praiseworthy. Exclusionary reasons, however, are not instances where compliance is required.

Recall there is a point to positing a third kind of defeat. We want our choices to have certain character or instantiate certain values, but this often involves some loss or trade-off between values. How do we go about choosing some values and rejecting others? On both accounts, exclusion is thought to be indispensable to this more nuanced picture of value pluralism, according to which morality and practical reasoning necessarily involves value conflicts and trade-offs (Scheffler, 2004). However, the role exclusion plays differs on both accounts. In what follows, I make some suggestive comments about the additional nuance that exclusion promises, and why it is only available on Raz's account.

To see this let's contrast the two accounts again, this time using legitimate law as an example. Legitimate laws are supposed to exclude considerations in favor of law-breaking. I am driving on the highway and see an animal suffering on the side of the road. The suffering gives me a reason to stop but it is illegal for me to pull over here. How does the law exclude the suffering? For Adams that there exists a traffic law is not a fact that identifies a feature of the fact 'that it would relieve the animal's suffering' that explains why it does not count when pulling over. Instead it identifies a feature of pulling over—that it is illegal—that explains why relieving the animal's suffering does not count when pulling over. Just as in the burger example the choice changes from being a choice about what to eat to being a choice about what to eat *as a vegetarian*, the choice is no longer whether or not to pull over, but whether or not to pull over *as a law-bound citizen*.[19] The agent's commitment to being a law-abiding citizen disqualifies it as

[19] The exclusionary effect that particular choices have is narrow in scope. It can still be a reason for other things, i.e. calling animal control, pulling over up ahead, etc. and thus leaves a remainder (Adams, 2020, p. 251).

grounds for *legally* pulling over—that is, as something that makes pulling over valuable in the context of one's commitments. To be sure, it is no longer a reason not because the state of affairs have changed, for it is true that the animal is still suffering, but because pulling over for the sake of relieving the animal's suffering is no longer valuable. It follows that, on this account, the agent's commitments remove certain values from the picture.

Raz, however, understands things differently. He maintains that that pulling over would relieve the animal's suffering *just is* a feature of pulling over at this time that makes it valuable. The law does make it the case that you ought not pull over, yet this is not because relieving the animal's suffering no longer justifies doing so. It is because what, on this occasion, you are justified in doing is not acting for the animal's suffering in a way that were it not for the law's existence you wouldn't be. Further still, Raz suggests that the law, as a reason for us, indicates that there is something valuable in doing as it requires. It is a feature of the agent's state of affairs that makes not pulling over to relieve the animal's suffering valuable or good. As such, while Adams' account is about removing certain values from the picture, Raz's account is about adding one into the mix. The law may be grounded in any number of values, such as pragmatic values (i.e. coordination), democratic values (i.e. respecting democratic procedure), or simply the value in being the kind of person that respects rules. Whatever it is, exclusionary reasons on Raz's account capture the sense in which laws, promises, commands, etc. make not acting for some reasons valuable, and the sense in which agents choose (i.e. intend) to instantiate these values when they follow the law, keep promises, and obey commands.

He allows that our commitments do in fact alter what we have reason to do without it being the case that they change what is or isn't valuable. They instead give us reasons by making it the case that being unresponsive to certain values on some occasions is itself valuable, and so something that rational agents can be moved to do. In this way, exclusion adds more nuance to our understanding of value pluralism and value conflicts: it creates space for instances where we have reason to not pursue certain values. They give our choices and actions character (i.e. their fairness, autonomy, altruism, etc.) precisely because they involve not acting for otherwise valid reasons, or eligible options, not because our choices or actions having a certain character make it the case that certain options are no longer eligible or valuable.

Adams seems to miss this entirely. The notion that his account relies on, namely that agents can make choices that have a certain character or that instantiate a particular value just is the idea that agents can 'φ-for-*p*'.Whereas Raz pulls an agent's acting (i.e. φ-ing) and their acting for certain reasons (i.e. φ-for-*p*) apart, Adams suggests that agent who chooses to *φ in a certain way* (i.e. fairly, autonomously, etc.) when they choose to φ, and this is because they recognize that some reasons no longer count. However, on an account where exclusionary reasons are not reasons to act, and thus not reasons to intend to do as they require, it is unclear whether agents can ever choose to *φ in a certain way*. Even if on his account it matters *how* we do things, he is unable to explain (or at least has yet to do so) how agents meet the demands of exclusionary reasons intentionally, and not merely coincidently or accidentally. As a conceptual matter, then, his account seems to be squarely about what we do, and not how we do it. This discussion does not represent a conclusive case for Raz's approach, but it does present some challenges for Adams' account. At the very least, what is missing from his account is an explanation as to how

exclusionary reasons make it the case that features of the world are no longer valuable even if only on some occasions or in certain contexts (i.e. the metaphysical question); and, how and why agents are ever responsive to these kinds of reasons (i.e. the normative question). By making exclusionary reasons, and so promises, rules, commands, etc., out to be based in values that agents can recognize and choose to instantiate, Raz avoids these concerns.

However, the account of exclusionary reasons offered herein is contingent on whether it can meet concerns about Raz's notion of acting-for-a-reason—i.e. the concern that motivations and intentions do not seem to be under our control in the way this account requires. We turn now to the objections of this kind.

Chapter Two: Acting for Exclusionary Reasons

1. Introduction

The idea of an exclusionary reason is coherent and distinct only when they are understood as reasons to not be motivated by certain considerations. As chapter one suggested, on this reading, exclusionary reasons are reasons to not be motivated by one's rational reaction to certain reasons or values. This explains the distinct normative situation they create: they call on agents to not act for reasons even when they are recognized as such. Understood in this way, exclusionary reasons are normative reasons that have as their object an action, namely 'not-φ-for-*p*'. In asserting that exclusionary reasons are practical reasons, Raz commits them to both of the existence conditions the nexus sets for normative reasons. Exclusionary reasons are facts that confer value on and that (can) motivate or guide rational agents in not acting for some reasons (i.e. in 'not-φ-ing-for-*p*'). If they really are normative reasons for acting, it follows from the first condition that they confer value on 'not-φing-for-*p*'. Moreover, as the second condition of the nexus requires, at least some of the time, rational agents are able to *act for* these reasons. That is, that the rational agent can choose or intend to do as they require.[20] However, as several theorists have recently argued, it is unclear that we can do just this—that we can intend to *refrain from acting for reasons,* or 'not-φ-for-*p*'.

[20] This presupposes a familiar constraint on normative reasons that maintains that if they are to count as a reason for some action, it must be at least possible that someone could perform the action (*i.e. the ought-implies-can principle*). The suggestion is that exclusionary reasons fail the motivational constraint *because* our not being able to 'not-φ-for-*p*' precludes our being guided by exclusionary reasons in 'not-φ-ing-for-*p*'.

Doing as exclusionary reasons require, or 'not-φ-ing-for-*p*', seems to involve an ability to (1) act for some reasons and not others while (2) holding constant our judgments that excluded reasons are in fact reasons. The former concerns the role that agents play in bringing about intentions to φ for reasons other than *p*. The latter concerns our ability to be irresponsive to excluded reasons (i.e. reason *p*) as rational agents. Yet, it seems both (1) and (2) require a level of control over our attitudes and reasons that a number of theorists deny (Moore 1988; Essert 2012; Brown 1977; Scanlon 1998; Ross 2002). The idea shared by these theorists is that we cannot choose to bring intentions about or maintain certain of our judgements because our attitudes are not 'up to us'. If this is in fact what 'not-φ-ing-for-*p*' involves, then it seems we cannot act for exclusionary reasons. An inability to account for how we can 'not-φ-for-*p*' poses a serious challenge for Raz' idea of exclusionary reasons. This chapter seeks to explain what refraining from acting for reasons, or 'not-φ-ing-for-*p*,' is and how we can (on some occasions) choose to do it. It considers objections which claim that the notion of *refraining from acting for reasons* Raz's account presupposes is incompatible with the rational constraints on attitudes (i.e. beliefs and intentions). In response, I show that the restated account of exclusionary reasons avoids them. Perhaps surprisingly, it is argued that instances where exclusionary reasons are relevant, when properly understood, are not instances where reasoning about what we ought to do involves choosing the reasons for which we act.

In section 2, I describe how our attitudes, both beliefs and intentions are rationally constrained, and show how acting for a reason, or intending, in the ordinary sense, differs from intending for reasons. In section 3, I consider an objection that characterizes exclusionary reasons as reasons to intend for reasons—that is, to *have* motivations of certain kinds, or to

choose to be motivated in some way. I then offer a reply. Using the satisfaction conditions of exclusionary reasons, I identify what it is to 'not-φ-for-*p*', and show that acting for an exclusionary reason involves intending in the ordinary sense. In section 4, I consider an argument that denies our ability to hold constant our judgment that excluded reasons remain normative reasons irrespective of whether they are defeated by an exclusionary reason. I argue that while we cannot maintain conflicting beliefs about what we ought to do, the distinctly pluralistic nature of value-constituted reasons for action allow for conflicting beliefs about the reasons we have. I conclude by noting the role that exclusionary reasons play in our deliberations about what to do. I suggest that the account of 'not-φ-for-*p*' offered herein demonstrates how exclusionary reasons are reasons that constrain the ways in which a rational agent can conform to some of her reasons so that she may come closer to satisfying all of the reasons that apply. In doing so they are reasons that determine for us—and thus removes choice over—the reasons for which we ought to act.

2. Rational Constraints on Attitudes

As noted in chapter 1, Raz's account of exclusionary reasons relies on the notion of *refraining from acting for reasons* or 'not-φ-ing-for-*p*'. Yet, several theorists have raised concerns about the managerial role that it seems agents must play in bringing about intentions to φ for reasons other than *p* or in maintaining one's judgement that an excluded reason remains a normative reason despite its exclusion. To understand these charges, we must first consider what it is to believe and act for a reason (i.e. intend) in the ordinary sense, how these attitudes are rationally constrained, and why this precludes an ability to believe or intend voluntarily.

Consider Pamela Heironymi's argument that canvasses the general constraints on our ability to form attitudes. Heironymi argues that both believing and intending are less than voluntary.[21] To be sure, there is a sense in which we control our attitudes—the sense in which forming attitudes is part of the active side of life (Heironymi, 2006, p. 46; Raz, 1999a, p. 10).[22] For example, one believes *p* insofar as one takes *p* to be true. The believer controls her beliefs in the sense that she "decides" whether or not the case for the truth of the belief's contents is conclusive. Heironymi calls this evaluative control. When we find the case for the truth of the belief's contents conclusive we acquire the belief. In doing so, there is no gap, no extra step in reasoning, between believing that the case for the truth of the proposition is conclusive, and believing the proposition where the content of the proposition is immediately available. An agent who takes a (truth-related) consideration to bear on whether *p* to be sufficient in showing *that p*, believes *that p*, and does so immediately (Raz, 2011, p. 38; Heironymi, p. 54). Belief's immediacy is not temporal or causal, but rather a consequence of the constitutive relation between the commitment to *p* as true and the belief. A well-functioning believer who believes *p* is committed to *p* as true, where the belief *just is* this commitment.

Two things follow from Heironymi's argument. First, our beliefs represent rational commitments in that they are answerable to requests for rational justification. The beliefs we can

[21] This is understood in terms of responsiveness to practical reasons, where practical reasons, according to Jonathan Bennett, are reasons to make it the case that something is true as opposed to reasons that bear on what is true. See his (1990) "Why Belief is Involuntary?," *Analysis* 50, p. 89.

[22] Raz suggests that we are active when our mental life displays sensitivity to reasons, and passive when mental events occur in such a way which is not sensitive to reasons. This distinction, while related, does not correspond directly with the voluntary/non-voluntary distinction. See his (1999) "When We are Ourselves: The Active and the Passive" in *Engaging reason: On the theory of value and action*. Oxford University Press.

form depend on the availability of reasons that can justify having these beliefs. Because belief is said to "aim at" or is "answerable to" the truth, a rational agent can form a belief only insofar as she takes the contents of the belief to be true (Williams, 1973, p. 148). Yet, because they purport to represent reality, determining whether this standard is met is not up to us. Believing in the ordinary sense, then, is less than voluntary. Second, agents can believe in the ordinary sense in response to only those reasons (i.e. facts) that are part of the case for the truth of the belief's contents. Raz calls these epistemic reasons, though they are also described as content-related or truth-related reasons.[23]

Believing voluntarily, by contrast, requires an ability to decide to believe, perhaps because it would be good, useful, rewarding, etc. That is, it seems to require an ability to believe for non-truth-related reasons, or *extrinsic* reasons. These are reasons that favour *having* the belief by showing that the possession of the belief would be good in some sense independent of whether or not it is true. In short, non-truth related reasons are reasons that favour *p* without showing *that p*.

A common thought is that non-truth related reasons for belief are reason of the wrong kind for belief (WKRs) (see also Kelly 2002, Shah 2006) that differ from truth-related reasons in the following respect. Whereas finding some epistemic reason for belief *p* to be convincing or sufficient leads immediately to the belief '*that p*,' finding an extrinsic reason for believing *p* convincing leads (immediately) to the belief that 'belief *p* is good to have'. From this difference,

[23] See Christian Piller's (2001) "Normative Practical Reasoning," Proceedings of the Aristotelian Society, Supplementary Vol. 75, 1, pp. 195–216. It is worth noting that Piller denies exclusion as a distinct kind of defeat by appeal to this distinction in his (2006) "Kinds of Practical Reasons: Attitude-Related Reasons and Exclusionary Reasons" *Analyses.* pp. 98-105.

a further key difference follows. Unlike with conclusive epistemic reasons for belief, the adoption of a belief on the basis of the WKRs Is voluntary. That is, upon consideration of the WKRs for a belief, the agent's adoption of the belief requires a further act of the will—an intent or decision to adopt the belief. Assuming that it is possible to adopt a belief on the basis of WKRs, the key point is that any such adopted belief is mediated by an act of the will. To adopt such a belief is to *act upon oneself* and bring it about that one has the belief *p* (Hieronymi, 2006, p. 55).

Reasons for intending are similarly divided into content-related and extrinsic reasons—those that bear on the contents of the intention, namely the action, and those that make the intention itself valuable or good to have. Content-related reasons for intending are just those practical reasons that favour the relevant action—a reason to intend to φ just is a practical reason for φ-ing. Just as truth-related reasons, in being reasons for the belief, justify having the belief, practical reasons, in being reasons for the action justify executing the intention, or, in other words, acting. Intending in the ordinary sense, like believing, involves evaluating one's reasons for acting and deciding which of them present a conclusive case for taking the action. This is again an exercise of evaluative control.[24] An agent intends when she answers for herself the question of whether so to act (however implicitly, unreflectively, unconsciously, or spontaneously; Hieronymi, 2006, p. 56). However, intending in this sense is also less than voluntary. An agent who has resolved that she ought to, say, φ, commits herself to doing so by forming the intention to φ. Like beliefs, then, intentions are rational commitments that are likewise constrained by the availability of reasons that justify executing the intention, or acting.

[24] Here, 'deciding' marks the conclusion of one's deliberations not an intervention of will. That is, it is not an act that constitutes choosing.

Consider why this is so. As the nexus suggested, to be guided by or act for a reason (i.e. to intend) is to be motivated by the mediating belief that recognizes the reason as such—as a reason for, or that justifies, so acting (Raz, 2011, p. 29).[25] An agent who is aware of the reason *believes* in the fact that is the reason, and in its character as a reason, and in doing so she believes that the action is one she ought to perform—that it is an act that she has reason to perform (Raz, 1999b, p. 212). Practical reasoning, or determining what we ought to do, then, should be understood as an inferential process whereby we modify our beliefs, including our normative beliefs about what we ought to do. We conclude our reasoning when we arrive at a belief about what we have reason to do, and adjust our intentions accordingly. Forming intentions, then, depends on the availability of reasons for intending because our intentions are governed by beliefs about the reasons we take ourselves to have, and what they give us reason to do. They are, in this sense, rationally constrained such that an agent's commitments to acting in certain ways (i.e. her intentions) reflect her commitments to certain beliefs about the reasons she takes herself to have. As such, we control our intentions only to the extent that, in response to reasons for acting, we resolve or decide, rationally—that is, based on what we believe they give us reason to do—that we ought to do as they require.

Consider an example. In deciding between two otherwise identical pairs of new loafers, where I have need for only one pair, I consider the fact that pair A is substantially cheaper, whereas pair B is designer and thus far less affordable. If we accept that the lower price point is

[25] Practical reasons are also reasons to believe that one has reason to perform the act (Raz, 1999, p. 212). As such, reasons explain actions through the mediation of belief in the reasons *as reasons.* Since the belief is explained by the existence of the reason the reason is part of the explanation of the belief, and thus figures in the explanation of the action (Raz, 2011, p. 32).

the weightier reason (perhaps due to my current financial situation), I evaluate (correctly) that I had better choose pair A, and I believe as much. Insofar as I take my reason for choosing pair A to defeat my reason for choosing pair B, namely that it is designer, I believe or decide that I ought to choose pair A, and therein intend to do so (i.e. to buy them). Yet, while I evaluate pair A as what I ought to do, it is my reasons that make this the case. That 'pair A is substantially cheaper,' as the undefeated, available reason, presents a conclusive case for choosing pair A, and my intention to buy pair A is rational when it reflects the judgment that choosing pair A is what I have reason to do. It is rational because it reflects a commitment to taking the action that my reasons support.[26] When agents are fully rational, there is some requirement to intend to do what one believes one ought to do. When she does so her actions are intelligible to both her and those around her. Our intentions, then, being rationally constrained, are in the same sense not up to us. Because they reflect (fallibly) the reasons that justify acting as we intend to, they are not the product of our own choosing.

Intending voluntarily (i.e. at will), by contrast, just *is* choosing to have particular intentions. Like believing at will, it involves intending for extrinsic reasons. These are reasons that show the intention good, useful, desirable, appropriate, or important to have independently of whether it would be good to φ, or in other words without bearing on whether to so act.[27] Whereas deciding that a reason presents a conclusive case for φ-ing leads an agent to form an intention to φ, an agent that finds an extrinsic reason convincing fails to do so. She cannot, in

[26] An intention to φ is rational insofar as φ-ing is rational because it represents a commitment to acting accordingly. When I executes the intention—when I choose pair A—I act rationally, because I does what I ought to or have reason to do.

[27] These are not just reasons that make the intention good to have for any and all practical reasons do just this. A practical reason that makes φ-ing good or valuable will also make the intention to φ good or valuable insofar as the intention is instrumental to the action's performance.

response to these reasons, form the intention to φ (i.e. commit herself to φ-ing) because an extrinsic reason fails to give her reason to φ. In not being reasons that justify φ-ing, they are not reasons that suggest that φ-ing is what she ought to do. Instead, extrinsic reasons give agents reason to merely *have* the intention. In response, an agent evaluates having or forming an intention to φ as what she ought to do, and forms an intention to have the intention to φ (i.e. to intend to φ). An agent can, in principle, and however irrationally commit herself to *having* the intention to φ and thus form an intention to act upon oneself by taking steps designed to bring this intention about (e.g. by fabricating reasons to φ). This, however, is not intending in the ordinary sense. In the absence of reasons that justify taking some action, an agent cannot rationally commit herself to acting accordingly. Our intentions, then, are not amongst the items that we can will.

3. The Wrong Kinds of Reasons Objection

That our intentions rely on the availability of reasons that can justify acting accordingly points to the difference between acting for reasons, or intending in the ordinary sense, and intending for reasons. A rational agent intends in the ordinary sense when she is responsive to reasons that can justify executing the intention, or acting such that her intentions reflect what she takes herself to have reason to do. An agent who instead intends for extrinsic reasons, or reasons of the wrong kind, is committed to reasons that make having certain intentions worthwhile. With this distinction in mind, consider the first objection against the notion of refraining from acting for reasons (i.e. 'not-φ-ing-for-*p*').

If exclusionary reasons are to count as a reason for some action, namely 'not-φ-for-*p*', it must be at least possible that an agent is capable of 'not-φ-ing-for-*p*' and choose or intend to do so. An agent that does so (i.e. that complies) acts for or complies with the exclusionary reason, and thus conforms with the rule, promise, command, etc. non-accidentally or coincidentally. If *acting for reasons* has to do with how an agent is motivated, or the reasons for which she intends, then not acting for reasons requires than an agent choose, or "single out," certain of the reasons she has as the ones she acts for or that moves her to act. Choosing to 'not-φ-for-*p*', then, seems to require an ability to choose which of the applicable reasons we act for, or to form, by our own choosing, intentions to act on some reasons and not others. In other words, they seem to be reasons to *have* certain intentions or form intentions in a particular fashion (i.e. for some reasons and not others; See Adams, 2020, p. 239; Moore, 1988, p. 878; Essert, 2011, p. 480). If exclusionary reasons are practical reasons to 'not-φ-for-*p*', they are features of 'not-φ-ing-for-*p*' that make doing so good or valuable. Yet, if 'not-φ-for-*p*' consists in having certain intentions or forming intentions in a particular fashion, then are exclusionary reasons not reasons of the wrong kind for intending?

Daniel Whiting (2017) makes a version of this argument against second-order reasons in general. The objection is as follows. An agent can intend in the ordinary sense—or in the right way—only in response to reasons that justify or count for the action the agent intends to take. If, for example, Mary were to intend to choose school B, she could do so for the reason that 'school B has small class sizes'. However, by their very definition, second-order reasons (and so exclusionary reasons) cannot make acting (or not acting) good or valuable. They cannot be reasons to φ or not-φ, at least not without amounting to mere first-order reasons that participate

in the balance of reasons. This seems right. After all, her promise to Dave is not a reason that favours choosing school B in the sense that it is not a feature of school B that makes it valuable. Instead, a second-order reason seems to be a respect in which φ-ing on a certain basis (i.e. for a certain reason) without also being a respect in which what on that basis is decided, namely φ-ing, is good (Whiting, p. 409). For example, the fact that 'in choosing school B because it has small class sizes Mary keeps her promise' is a respect in which deciding on educational grounds is good, not a respect in which choosing school B is good.

In being reasons for the attitude (i.e. φ-ing for *q*), without also being reasons for the relevant action (i.e. φ), they are reasons to *choose* to have certain intentions, or reasons to intend to intend in some way. As a conceptual matter, then, exclusionary reasons are extrinsic reasons, or reasons of the wrong kind for intending. They are reasons to act upon ourselves to bring about an attitude that reflects a commitment to one reason (i.e. *q*) and not another (i.e. *p*). This is again, in principle, possible, however, it is not what intending in the ordinary sense involves. We cannot intend to have certain intentions because our intending to have intentions to, say, φ does not lead to a genuine intention to φ. And if this is what acting for exclusionary reasons, or intending to 'not-φ-for-*p*' amounts to, namely intending to form an intention to φ in response to reason *q* and not reason *p*, then it seems we cannot act for exclusionary reasons either.[28] Because they fail to meet the nexus' second condition, it follows that exclusionary reasons are not normative reasons at all.[29]

[28] This presupposes a familiar constraint on normative reasons that maintains that if they are to count as a reason for some action, it must be at least possible that someone could perform the action (*i.e. the ought-implies-can principle*). The suggestion is that exclusionary reasons fail the motivational constraint *because* our not being able to 'not-φ-for-*p*' precludes our being guided by exclusionary reasons in 'not-φ-for-*p*'-ing).

[29] In the case he makes against exclusionary reasons, Christian Piller also takes exclusionary reasons to be an attitude-related reason—in being features of the relevant attitude. However, whereas this objection denies that exclusionary reasons are genuine normative reasons by denying that attitude-related reasons are genuine normative

3.1. A Reply

The objection, while forceful, misconstrues the nature of exclusionary reasons and the situations in which they tend to occur. In particular, this misconstrues what it is to 'not-φ-for-*p*'. To act for an exclusionary reason is not to intend to have certain intentions, or to be motivated in certain ways because, as chapter 1 endeavoured to show, exclusionary reasons are not relevant where it matters what motivations we have. Instead, exclusionary reasons are relevant where it matters how we act (i.e. how we φ). To see the difference, consider what it means to satisfy a reason. An agent satisfies a reason when she does all that it gives her reason to do. An agent who has reason to buy lunch does what she has reason to do when she buys lunch, and, where she has reason to not buy lunch, when she fails to do so (i.e. when she eats the lunch she has prepared, or refrains from having lunch at all). Where it it matters what motivations we have, then, an agent satisfies her reasons, or does what she has reason to do, when she merely has a certain motivational set. If, for example, all that an agent needs to do in order to win a million dollar cash prize, is have the intention of drinking a toxin, the prospect of winning the cash prize merely gives her reason to have the intention.[30] She satisfies this reason, and thus is awarded the prize, when she forms the intention without having to execute it, i.e. drink the toxin. And because she can win the cash prize without drinking the toxin, the cash prize fails to give her reason to, or to justify drinking it. In other words, there is no reason to execute the intention, or

reasons, Piller argues that attitude-related reasons play a distinct normative role in our practical deliberations. He argues that the distinct normative role how exclusionary reasons are said to play can also be captured by employing the concept of an attitude-related reason.

[30] See Kavka, G. (1983) 'The Toxin Puzzle', Analysis, 43, 33–36.

act. As such, the agent has an extrinsic reason to intend because reasons that justify acting accordingly are absent.

By contrast, where it matters how we act, intending in some way or having certain motivations is not enough. Instances where it matters how we act are instances where acting for some reasons and not others is constitutive of the action's right performance—that is, of *φ-ing in the right way*. Quite unlike the toxin case, then, an agent *φs in the right way* when she also *φs* (i.e. when she executes the intention). This is because in addition to reasons that favour φ-ing in some way (i.e. for *q* instead of *p*) there exists independent reasons to φ (i.e. that justify φ-ing). An agent can φ if she has reason to φ, and she can φ in a certain way if she has certain reasons for φ-ing. For example, she can φ-for-*p* if there also exists reason *p* for φ-ing, and she can φ but not for *p* if there also exists reasons *q, r, s,* etc. for φ-ing.[31] And since an agent's reasons for forming and executing the intention to φ (i.e. for φ-ing), exist independent of her reasons to *φ in some way*, she can do so—*φ in some way*—without having to intend to have the requisite motive or intention. An agent can φ such that she keeps a promise, obeys a command, or follows a rule even when she fails to consider that φ-ing that way, or for those reasons, would allow her to do so.

To see this, consider what Mary's promise to Dave (to choose a school not for finance- or career-related reasons), calls on her to do. Her promise calls on her to keep her promise when she chooses a school for her daughter. Because she can keep her promise when she chooses the school that is best on educational grounds alone. It follows that her promise also calls on her to

[31] Where there are no independent reasons to φ she ought not φ at all.

refrain from acting for considerations that are not education-related.[32] Here, Mary satisfies her promise, and does all that she has reason to do, when she chooses the school that the education-related considerations—at least those she is aware of—determine to be the best school. If school B is the better school on educational grounds alone (i.e. because it has small class sizes), she does all that her promise gives her reason to do—she keeps her promise—when she chooses school B, and this is so whether or not she is attentive to the reasons for which she chooses school B. That is, Mary still keeps her promise even when, for example, she picks the right school by some coincidence or accident (e.g. because she is aware of only education-related grounds or when she forgets about her promise). In this same vein, she can fail to keep her promise to Dave by merely being attentive to her reasons for choosing school B. If Mary were to simply intend to choose a school for education-related grounds, but fails to choose a school for considerations of this kind (perhaps due to a weakness of will in the presence of more attractive finance-related reasons), she fails to keep her promise to Dave and thus satisfy it. It matters, then, not just how she is motivated when she chooses a school for her daughter, but that she actually does so, and does so for some reasons.

Mary's normative situation differs from the toxin case in that her promise to Dave is relevant *because* she has reason(s) to choose a school for her daughter. Her reasons for choosing her school *and* her promise to Dave, as a reason to refrain from choosing on certain grounds, make it the case that she ought to or has reason to choose a school (i.e. φ), *and* to choose a school in some way. She can choose a school not for career or finance-related reasons, if there

[32] Her promise calls on her to do two things. It gives her reason to keep her promise, as well as reason to exclude considerations that recommend doing otherwise. This just is the combinatory concept of a protected reason or pre-emptive reason that Raz appeals to in explaining obligations and authority. See Raz, J. (1990). Authority and justification. In J. Raz (Ed.), Authority. Oxford: Basil Blackwell.

exists other grounds for which she can choose a school, and when she chooses a school for those reasons, she satisfies her promise even when she is otherwise inattentive to the reasons other than the education-related consideration for which she chooses the school (i.e. that it has small class sizes) that apply. My suggestion, however, is not that exclusionary reasons are reasons that we satisfy only when we are inattentive to them, such that we can never act for exclusionary reasons intentionally. I am instead suggesting that if we can understand when exclusionary reasons are satisfied, we can understand what it is to 'not-φ-for-*p*'. Exclusionary reasons are not reasons to intend to act or not act for some reasons because they are not satisfied when an agent does just this. Instead, an agent does all that an exclusionary reason requires, or satisfies it, and thus 'not-φs-for-p' when she φs for other undefeated reasons to φ (i.e. reasons *q, r, s,* etc.) instead, or, in the absence of other undefeated reasons to φ, when she does not φ at all.

It follows that 'not-φ-for-*p*' represents an action-type, namely φ-ings that are instances of not-φ-ing-for-*p*.[33] When an agent has an exclusionary reason, *e*, and she acts for a non-excluded reason by, say, φ-ing for *q* rather than excluded reason *p*, the action that satisfies both reason *q* and *e* just is the agent's φ-ing. Yet, to assert this is not to reduce exclusionary reasons to reasons for φ-ing, nor, for that matter, reasons to φ-for-x (where x is some other undefeated reason to φ) since an exclusionary reason does not explain an agent's φ-ing or her φ-ing-for-*x*. Invoking the nexus here helps make this clear.

The nexus tells us that a normative reason can feature in an explanation of the action for which it is a reason, as a fact that, being recognized for what it is, motivated the agent to

[33] Note that non-action might also be an instance of 'not-φ-ing-for-*p*'.

perform, and so guides the agent in performing the action. In doing so, normative reasons make the rational agent's action intelligible to both her and those around her. A rational agent that is responsive to her reasons, and thus acts out of an awareness of them, and who 'not-φs-for-*p*' by φ-ing-for-*q* makes sense of her φ-ing, and so her φ-ing-for-*q*, by reference to reason *q*. If asked why she φ-ed, or whether she was justified in φ-ing, she points to reason *q*—the feature of φ-ing that makes doing so good or valuable. She might respond, "I φ-ed because of *q*" or "I φ-ed for *q* because *q* makes φ-ing valuable". Mary, for example, can explain why she chose school B by citing its small class sizes as a feature of that school that makes it good or valuable: "I choose school B because it has small class sizes, and that makes it's a good choice for my daughter's education." A rational agent that 'not-φs-for-*p*' in response to an exclusionary reason will instead cite the exclusionary reason in an explanation of her failing to respond to or act for the excluded reason(s). Her exclusionary reason makes sense of her being unresponsive to otherwise valid reasons because it identifies value in, and so justifies, her doing so. If instead asked why or whether she was justified in not responding to reason *p,* she refers to reason *e*—the feature of 'not-φ-for-*p*' that makes doing so good or valuable. She might respond, "I did not φ-for-*p* because of *e*". For example, if asked why she didn't choose the more affordable option (i.e. school A), she cites her promise to Dave as her reason for not choosing a school on the basis of finance-related considerations: "While school A is far more affordable, I promised Dave I would only consider schools on the basis of education-related considerations, and so I didn't factor in the cost."

An agent that can do this, it seems, can act for an exclusionary reason intentionally. Insofar as she can (1) recognize that *e* is a reason that justifies or gives her reason to 'not-φ-for-

p', and that she can 'not-φ-for-*p*' by φ-ing for, say, *q*, and (2) explain her 'not-φ-ing-for-*p*' by reference to *e*, she can refrain from acting for reasons on the basis of the exclusionary reason (i.e. 'not-φ-for-*p*-for-*e*') by acting for other undefeated reasons instead. In other words, she can intend to do as the exclusionary reason requires insofar as she can resolve that 'not-φ-ing-for-*p*' is part of what she ought to or has reason to do as she does when she answers for herself whether so to act (i.e. whether *e* presents a conclusive case for doing what it recommends, and thus requires), and commits herself to acting accordingly. Notice, however, while an agent's (undefeated) exclusionary reason gives her reason to conform to it, it leaves undetermined how she should do so. An undefeated reason to keep her promise to 'not-φ-for-*p*,' for example, gives Mary reason to act in a manner that is consistent with her promise, yet what this entails will depend on the other reasons for acting that are available to her. For example, Mary might have another undefeated reason, say *q*, to φ, i.e. that school B has small class sizes. Choosing school B for reason *q* would allow her to keep her promise. Recognizing as much, she can resolve that what her undefeated reasons—both her promise (reason *e*) and that school B has smaller classes (reason *q*)—require of her is that she choose school B for the fact that it has smaller classes (i.e. to φ-for-*q*). However, where other undefeated reasons for φ-ing are not available, she acts in a manner that is consistent with her promise when she does not φ at all. If Mary were only able to identify reasons for choose among the candidate schools that are career and finance-related (perhaps because the schools she has shortlisted are not great in terms of the education they provide), she keeps her promise when she does not choose any of them. She might instead expand her search so as to find schools worth choosing on educational grounds.

Let's say Mary does have reason *q* for φ-ing, such that what she ought do is φ-for-*q*. In resolving what to do, and committing herself to φ-ing, she is moved by both *e* and *q* because both reasons supply answers to this question. Reasons *e* and *q* suggest that she is justified in φ-ing insofar as she has reason to φ, and insofar as her φ-ing is an instance of not-φ-ing-for-*p*. So, for example, in recognizing that by dint of her promise to Dave she has reason to be unresponsive to finance- and career-related reasons, she recognizes that she ought to choose school B because of the education-related consideration that favours it, and because doing so would allow her to keep her promise. When Mary chooses school B for reason *q* (i.e. small class sizes) and remains unresponsive to reason *p* (i.e. finance- or career-related reason) that favor some other school, she acts for *q* and she does not act for *p*. In so doing, she conforms to the exclusionary reason. She might comply with that exclusionary reason as well insofar as she remains unresponsive to her finance- and career-related reasons for the reason that she has promised not to choose the school on the basis of those reasons. In this case, she is guided by both the first-order reason for school B and the second-order reason not to act for reasons that would preclude her from keeping her promise. In summary, exclusionary reasons are reasons for intending in the ordinary sense. They are not extrinsic reasons to have certain intentions because they are not instances where what we ought to do is have or bring about some intention. Instead, they are instances where we ought to act and act in some way, and only some of our reasons allow us to do both.

4. The Normative Beliefs Objection

The previous section suggests that we can 'not-φ-for-*p*,' and we can do so intentionally, because acting for an exclusionary reason involves intending in the ordinary sense. But can this

account of *refraining from acting for reasons* explain how we hold constant the judgment that excluded reasons still are reasons for us? The following objection suggests otherwise.

As we saw in section 2, as rational agents who engage with features of the world via our rational powers, we reason about how we should act. There is, then, there is some requirement to intend to do what one believes one ought to do (irrespective of whether agents are acting fully rationally).[34] If acting for exclusionary reasons does involve determining or resolving which of our reasons would allow us to do as we ought—to both φ and φ in some way—then this seems to affect a change in our beliefs about what we ought to or have reason to do, and so our beliefs about what reasons we have. If, for example, I have an exclusionary reason to 'not-φ-for-*p*', such that what I ought to do is φ-for-*q* instead, my intention to φ must reflect only my commitment to or belief in *q* as a reason, even when I believe that both reasons *p* and *q* are reasons to φ. That I continue to take reason *p* to be a reason is, after all, the very idea of exclusion. In requiring that we both continue to recognize excluded reasons as reasons—or believe in their character as reasons—acting for exclusionary reasons, it would seem, requires the ability to form intentions on the basis of some beliefs about the reasons we have and not others. I am rationally required, for example, to respond only to the non-excluded reasons, and so 'single out' *q* as the reason I act on, or as the commitment that guides my φ-ing.

To do this, an agent must, in some sense, exclude her belief that *p* is a reason to φ while remaining committed to her belief in *p* as a reason. Consider an example ('Prize'). [35] As a

[34] Even when agents are mistaken about what they ought to do, they intend to act as they believe they ought to.
[35] I borrow this example from Christopher Essert. Examples like 'Prize' represent cases in which it is clear that it is an exclusionary reason that would be doing the work of guiding action. When the excluded reason favours φ-ing and the non-excluded reason favours ζ-ing, showing that one merely conforms with the non-excluded reasons (i.e. by ζ-

member of a prize committee, you have reason to award the prize to Tristan because he demonstrates financial need (reason *p*), and because he demonstrates the greatest academic achievement (reason *q*). However, the terms of the prize stipulate that you must select a recipient on the basis of academic merit alone (reason *e*). Acting for *e*—the prize's terms—requires that you not act for *p*, leaving only *q* as an eligible basis for your action. In other words, it requires that you choose Tristan for the reason that he is the most accomplished alone while remaining committed to the fact that he demonstrates financial need as a valid reason for awarding him the prize (albeit not one stipulated in the terms of the prize). On the basis of exclusionary reasons, then, we must exclude our belief in, or commitment to reason *q*. However, it is unclear that we can do just this.

An argument like this is made by Christopher Essert in his "From Raz's Nexus to Legal Normativity" (2012), in which he explains what exclusion at the level of epistemic reasons would require, and why it is not possible (See also Owens, 2008, p. 416-7).[36] Excluding a belief would require some middle ground between either accepting a piece of evidence as sufficient to justify a belief and not accepting the evidence as sufficient (Essert, 2012, p. 478). This middle ground would allow an agent, who believes that both reason *p* and reason *q* are reasons for φ-ing, to remain committed to the truth of proposition *p* without taking it to be sufficient for believing that (at least) part of her reason for φ-ing is *p*. Only then would her intention to φ reflect her

ing) cannot conclusively prove that the agent did in fact exclude some reason as opposed to, say, take the reason to have been outweighed or undercut. Cases where you have two (or more) reasons for φ, but should not act for some of them are instances where it seems the agent is choosing to not act for a reason as the exclusionary reason requires.

[36] I have extended Essert's original argument in order to capture the full strength of this objection. While he notes that the ability to recognize and respond to practical reasons is the same ability we have for recognizing and responding to reasons, he does not consider the ways in which exclusionary reasons exclude the normative beliefs that accompany reasons for action. He also does not consider the difference between akrasia and exclusion. However, these insights linger just below the surface of, and so follow from his original account.

commitment to reason *q*, and not her commitment to *p*. However, because belief is immediate, there can be no middle ground between either accepting a piece of evidence as sufficient to justify a belief and believing. A commitment to *p* as a reason that is sufficient for, or warrants, belief—a commitment to *p* as true—just is the belief *that p*. Upon becoming aware of the reason, then, a rational agent either does or does not acquire the belief, and when conflicting reasons become available her beliefs self-correct. Failing to form the belief or abandoning the belief *just is* taking the reason to be insufficient for that belief. As such, we cannot exclude reason *p* as a reason to believe that we ought to φ without taking it to be insufficient for believing that *p* is a reason to φ.

By extension, it would appear that there is no middle ground between taking some reason for action to justify so acting and acting for it (i.e. forming the intent to act). The suggestion here is not that we are compelled to act on a reason upon being convinced that it *is* a valid reason for us. Unlike our attitudes, our actions are not similarly (rationally) constrained. We can and quite often do act irrationally or *akratically* when we fail to act as we believe we have reason to. While our rational faculties commit us (fallibly) to certain beliefs about what there is reason to do we act by our own volition. As such, an agent who has resolved the question of what she ought to do, say φ, has yet to settle the question of what she is going to do. She may, in light of her belief that she ought to, φ, but she may instead γ.[37] I may take myself to have reason to buy the cheaper pair of loafers, and still decide to buy a designer pair of otherwise identical loafers. Here I am not choosing which pair I ought to buy (for in this respect I have already decided it's the cheaper pair), but instead which pair I ultimately will buy. Answering this further question does not

[37] That these are distinct questions leaves room for irrationality. Action is irrational when there is a discrepancy between what the agent judges to be the best (or better) thing to do, and what she does. It is faulted because the agent fails to conform to her better judgment.

involve 'picking out' which of the reasons that apply I should act for, but that I act in spite of what I take these reasons to be (i.e. what I take myself to have reason to do).

Instead the suggestion is that we cannot *rationally* take some reason for action to justify so acting and not act for it. Returning to Prize, I 'not-φ-for-*p*' when I instead 'φ-for-*q*': when I choose Tristan on the basis of his academic achievement. If I really do believe that him demonstrating financial need *is* a reason to award him the prize, it seems I necessarily intend to φ for both reasons. Conversely, if my intention to φ really does reflect only my belief that he should be awarded the prize because he shows the greatest academic achievement, then does this not signal that I no longer believe that *p* is a reason to φ? If taking a reason *p* to justify φ-ing, just is believing 'that *p* is a reason to φ', then it seems there can be no middle ground between either accepting *p* as sufficient for believing that one has reason to φ, and φ-ing for that reason. An agent who 'not-φ-for-*p*'s by acting for another reason necessarily gives up or revises her belief in *p* as a reason for φ-ing, and thus demonstrates a "commitment to the idea that the other reason—the one not acted for—is not really a reason […] at the relevant time" (Essert, 2012, p. 480). We cannot sincerely, or rationally, believe that an excluded reason really is a reason that justifies acting in some way when we do not act for that reason.[38] That is, we cannot take ourselves to have an excluded reason for acting without taking it to be insufficient to justify, and thus to not be a reason for, the action.

[38] This parallels R.M. Hare's argument against the possibility of weakness of will.[38] Hare argues: "just as sincere assent to a statement involves *believing* that statement, sincere assent to an imperative addressed to ourselves involves *doing* the thing in question: provided it is within my power to do *a* now, if I do not do *a* now it follows that I do not genuinely judge that I ought to do *a* now" (Hare, 1952, p. 20). In the context of the Hare-Davidson debate over the possibility of weakness of will, akrasia and weakness of the will are taken to be synonymous.

If we cannot do this, if we cannot hold constant our judgment that the practical reason in question is a reason to φ *and* not act for the reason, then we cannot 'not-φ-for-*p*'. Even if 'not-φ-ing-for-p' involves forming intentions in the ordinary sense, it still seems to involve some choice over the reasons we take ourselves to have.

4.1. A Reply

The objection relies on belief's immediacy, which allows an epistemic reasoner to move from premises (i.e. epistemic reasons) to a conclusion about what she ought to believe (i.e. the belief), and leaves no space for additional reflection about the reasons she takes herself to have. However, things are different when one reasons from practical reasons.

As noted in section 2, both determining what we ought to believe and determining what we ought to do are rationally constrained. Just as there are some beliefs that we ought to have, there are some actions we ought to take by virtue of the availability of reasons that justify or present a case for so believing or acting. Our normative reasons, in both cases, can be thought of as guides to determining what it is that we ought to—what there is reason to—believe or do. In the epistemic case, our reasons are governed by a single concern, namely the determination of whether the belief for which it is a reason is true. Because all of our beliefs purport to represent reality, epistemic reasons are all guides to one and the same end. There is no loss when an agent, who has conflicting (putative) reasons to believe that "it is sunny outside," such as her own observation that it is in fact sunny and a news report claiming that it is raining, dismisses the less reliable guide (i.e. the news report). Since the lesser reason for belief cannot serve a concern that

is not better served by the better reason there is no possibility of preferring to follow the lesser reason. This is why an agent who takes there to be reasons that justify believing *p* just believes 'that *p*,' with no extra step, and in the presence of reasons that suggest otherwise, she revises or abandons this belief, automatically as it were.

Our practical reasons, however, can serve or disserve a number of distinct and independent values, and are typically governed by many concerns.[39] This explains why an action can have several distinct evaluative properties, and why when reasons for action representing independent concerns conflict—such that we have reason to perform several incompatible actions—there is some loss. This is so even when it is clear which action is supported by the better reason. Consider an easy example. While sleeping in serves my interest in avoiding the anxieties of the day, it does not serve my interest in having a productive day. Similarly, while writing my thesis serves would my interest in having a productive day, it does not serve my interest in avoiding the anxieties of the day. Even when I ultimately resolve that what I ought to do is write my thesis, there is some loss in not sleeping in, namely that I do not get to 'avoid the anxieties of the day'. Thus, even when we successfully do our best some of the concerns that generate reasons for us remain unsatisfied (Raz, 2011, p. 42). It follows that in addition to discerning whether one has reason to act, about what we ought to do, at least some of the time, involves adjudicating between the concerns and values that our practical reasons serve (or disserve).

[39] A single action can serve or disserve a number of distinct and independent values and an action that derives from a single value may serve independent concerns. For example, an act can be out of both friendship and justice, and thus instantiate several values, or a single act, such as one that advances the welfare of several individuals, where the interest of each of them is an independent reason to perform it, it can serve several concerns.

Even when our reasons determine what it is that we are justified in doing, and so what we believe we ought to do, they need not determine our beliefs about the concerns at play. While I cannot believe that I ought to nap *and* ought to write my thesis, for I know that I cannot do both, I can believe that it would be good to sleep in (i.e. because I would get to do nothing) and good to write my thesis (i.e. because doing so would make for a productive day). As such, I can believe that both 'that I would get to do nothing' and 'that I would have a productive day' are reasons even when what I ought to do is write my thesis. In fact, later on, while I am writing my thesis, I might grumble over the fact that I could instead be sleeping in or that I am not sleeping in. The distinctly pluralistic nature of practical reasons allows for beliefs about the value of actions that are not self-correcting. An agent can hold several, consistent beliefs about reasons that recommend incompatible actions without revising any of them. This gives rise to special deliberative situations or outcomes.

First, it makes irrational or akratic action possible. Even when values are commensurable such that one's reasons do determine the right option, an agent can both believe that a practical reason is defeated by a better conflicting reason, and that it nonetheless serves a concern that the better reason does not, and thus act irrationally. I can, for example, act in spite of my belief that my reason(s) for working on my thesis defeat my reasons for sleeping in, because I also believe that my reasons for sleeping in serve my interest in doing nothing. Second, it allows for choice in determining what we ought to do when our reasons are incommensurable. If I were to choose instead between grading assignments and writing my thesis, where both have equally distant deadlines, and so either is a justified choice, I must choose what—in light of my goals, plans,

etc.—I ought to do. In other words, I must supply an additional premise.[40] I might consider the concerns my reasons serve (i.e. giving students feedback sooner on the one hand, and making some progress on my project on the other) matters more to me, and choose accordingly.

Lastly, the pluralism characteristic of practical reasons also allows for the possibility of exclusion. To see this, let us return to Prize. Both reason *p*, that Tristan demonstrates financial need, and reason *q*, that he demonstrates the greatest academic achievement, generate the beliefs that (1) you ought to award the prize to Tristan because he demonstrates financial need, and that (2) you ought to award the prize to Tristan because he demonstrates the greatest academic achievement. In other words, they generate beliefs in these reasons as reasons for φ-ing (i.e. choosing Tristan). Reason *e* also generates a belief that (3) you ought to or have reason to follow the committee rules and not choose a recipient on the basis of factors other than academic merit. Notice how beliefs (1) and (3) conflict. They conflict insofar as you cannot both believe that what you ought to do is award the prize because he demonstrates financial need, and follow the rule by not choosing a recipient on the basis of factors other than academic merit. Because these reasons recommend incompatible actions, it cannot be the case that you ought to do both. Beliefs (2) and (3), by contrast, are compatible. Taken together, they determine that what you ought to do is award Tristan the prize because he demonstrates the greatest academic achievement. Yet, this need not affect your belief that (4) awarding him the prize would help meet his financial need. That awarding him the prize would help his financial situation just is a feature of awarding Tristan a monetary prize that makes doing so good or valuable. Thus a belief that this just is a

[40] Raz tells us that where our reasons underdetermine what ought to be done we supply the rest. That is we factor in our goals, projects and desires, and determine what we ought do accordingly. See his (1999) *On the theory of value and action.* Oxford University Press., at 65.

feature of φ that makes φ-ing valuable is consistent with a belief that the reason is excluded by another of her undefeated reasons, namely the committee rule.[41] Exclusionary reasons, then, can defeat other reasons without changing what is or what we take to be reasons. That is, They can, then, play the same role in our practical deliberations as other normative reasons do insofar as they change what we believe we ought to do, and they can do so without changing what reasons we take ourselves to have.

5. Exclusion as a Rational Constraint

The two replies taken together show how our exclusionary reasons, as normative reasons, figure into our deliberations in an ordinary sense. As reasons that justify being irresponsive to some of our reasons, they make a case for not acting for them by identifying something valuable in our being irresponsive to those reasons, and they do so by altering the beliefs we have about what we ought to, or have reason to do. Agents who can recognize exclusionary reasons as such, can be moved to do as they require insofar as they can answer for themselves whether conforming to the exclusionary reason is part of what they ought to do. Understood in this way, exclusionary reasons serve as guides in doing what we have the *most* reason to do.[42] Because we are rationally required to, when possible, come to as close to complete conformity to our reasons as possible, exclusionary reasons constrain the ways in which one can conform to other of our

[41] How excluding remains distinct from undercutting should now be clear. Consider how one's deliberations may have differed if instead of the committee rule, she was informed that Tristan doesn't in fact demonstrate financial need. Unlike the rule, this new consideration generates a belief—that Tristan does not demonstrate financial need—that is inconsistent with another, namely her belief that awarding him the prize would help meet his financial need. In the face of conflict, her belief set is revised automatically, such that she no longer holds a commitment to this reason for choosing Tristan as true. When it comes to exclusion, on the other hand, an agent can hold both her belief that the rule is a reason to exclude certain considerations (and so belief that some reasons preclude her from conforming to the rule), and that the excluded reason still is a reason (i.e. that it counts for awarding him the prize).
[42] Critical to Raz's conception of rationality is the general truth about the nature of reasons that one should come as close to complete conformity to the reasons that apply as possible ('the conformity principle'; Raz, 2011, p. 189).

reasons, in order to allow for (as close to) complete satisfaction of all of them (Raz, 2021, p. 12). and they do so by making a class of actions, not-φ-ings-for-*p*, and so φ-ings for only some reasons, valuable. Our belief in exclusionary reasons *as reasons*—that is, as features of our being irresponsive to some of our reasons that make doing so valuable—help us make sense of our beliefs about the other values on offer. Where we believe that we have reason to not act for some of our reasons, we believe that doing so would be valuable. And we can be motivated by this fact—that there is some other value that we can instantiate only when we reject some other values—in doing so (i.e. in rejecting those values). Only when exclusionary reasons meet both of the nexus' conditions, when they are normative reasons, can we be moved to not act for some of our reasons without having to choose among the reasons available to us the ones for which we will act. Instead, by identifying some value in not acting for certain reasons, they determine what these reasons are for us. As such, the exclusion of some reasons represents an additional rational constraint on the attitudes we can form in determining what it is that we ought to do. That some reasons are excluded constrains what an agent ought to do *all things considered*, and an agent who is fully responsive to her reasons does as the exclusionary reason requires when she evaluates correctly which of the reasons that apply present a *conclusive* case for acting, where a correct evaluation of what she has reason to do just is one that allows for as close to complete conformity to the reasons that apply as possible.

These objections, then miss their mark. While rational agents cannot choose the reasons for which they act, they can, as fully rational agents, recognize the ways in which their reasons for acting are rationally constrained, and they can (and do) form the corresponding attitudes without having to intend or believe voluntarily.

Conclusion and Next Steps

This project's aim has been to understand exclusionary reasons as normative reasons. It took up two tasks along the way: (1) clarifying the sense in which exclusionary reasons are reasons to refrain from acting for some of our reasons and (2) showing that we can do as they require. In vindicating the Razian account of exclusionary reasons, I demonstrated that exclusionary reasons are committed to a set of conceptual truths about normative reasons in general, and practical reasons in particular. In doing so, I appealed to the normative-explanatory nexus, which explains how reasons are made for rational beings like us, and thus sets existence conditions for normative reasons. It tells us that reasons are normative insofar as they identify value in, thereby making eligible and thus *requiring,* certain responses, and can guide rational agents in, or move them to, respond as the reason requires. I used the nexus in (1) in showing that exclusionary reasons are about whether and how agents ought to *respond* to certain reasons, and (2) in showing that exclusionary reasons can meet the conditions it sets. The result is the following restatement of his account.

As negative second-order reasons, exclusionary reasons are reasons to refrain from acting for some reasons (i.e. to 'not-φ-for-*p*'). Put another way, they are features of 'not-φ-for-*p*' that make its performance worthwhile or valuable, and that can guide it's performance. As reasons that exclude, not outweigh or undercut, exclusionary reasons play a distinct normative role in reason conflicts. Because an exclusionary reason to 'not-φ-for-*p*,' *e*, conflicts with some undefeated first-order reason, *p,* to the extent that acting for (i.e. compliance with) *p* precludes conformity with the *e*. An excluded reason's (i.e. *p*) normative force is thus affected insofar as

the exclusionary reason makes it the case that it is no longer a fault that the agent fails to respond when the response is otherwise eligible or justified—it is no longer irrational in spite of the response's eligibility. As such, exclusionary reasons are reasons not to be motivated by one's rational reaction to certain reasons or values. And this explains the distinct normative situation that exclusion creates: where exclusionary reasons are relevant, we are justified in not responding to some of our reasons even when those reasons justify, or make eligible the response (i.e. the action) they require.

Exclusion allows that our commitments can alter what we have reason to do. Our commitments can, at least some of the time, require that our choices and actions have a certain character or instantiate a certain value and they can do so only when they are based on a restricted set of the reasons that apply. On these occasions, it matters not just what we choose to do, but *how* we choose to do it. By prescribing not acting for some reasons as part of what full reason responsiveness requires, exclusionary reasons play a role in determining what we ought to do. They tell us that in addition to the reasons we have for and against acting, i.e. φ-ing, we have reason to φ in some way. An agent φs in some way (i.e. 'not-φ-for-*p*') when she acts for some reasons and not others. Yet, this does not make exclusionary reasons about the motivations an agent should have when acting. Because one fails to satisfy an exclusionary reasons when she merely has the requisite intentions, to refrain from acting for reasons is not to bring it about that we have certain intentions. Since reasons to φ in some way are relevant where an agent also has reason to φ, an agent refrains from acting for reasons, or 'not-φs-for-*p*', and thus satisfies her exclusionary reason when she φs for non-excluded reasons—when her φ-ing is an instance of not-φ-ing-for-*p*. It follows that exclusionary reasons are reasons that determine for us the reasons for which we should act. Because an agent can φ and φ in some way (i.e. 'not-φ-for-*p*') on the

basis of only some reasons, exclusionary reasons constrain the ways in which one can conform to her reasons for acting, in order to allow for (as close to) complete satisfaction of all of them; and they do so by making a class of actions, not-φ-ings-for-*p*, and so φ-ings for only some reasons, valuable. Insofar as an agent can evaluate some of her undefeated reasons as presenting a conclusive case for φ-ing, and commit to φ-ing because it is justified—both by appeal to some feature of φ that makes its performance valuable and to some feature of the class of actions 'not-φ-ings-for-*p*' to which it pertains—an agent can act for an exclusionary reason. That is, she can be guided by the value she identifies in being irresponsive to some reasons insofar as she can identify which of her other undefeated reasons it would instead be appropriate to respond to.

Chapter 2 attributed the possibility of exclusion to value pluralism: the idea that many values in the world are not just incommensurable, but also incompatible since adopting some of them often requires that you neglect others. Our assessment of what we ought to do need not change the reasons we take ourselves to have, because our reasons—even when they are defeated—just are features of an action that make its performance worthwhile. As a consequence, when our reasons conflict such that we cannot satisfy all of them we give up some of the reasons we have for acting, and thus the good or value they represent. Raz tells us that "all people have to choose between different goods," and "that rejection of some goods, abandonment of some reasons for action, is an inevitable part of the life of good, even of ideal people" (Raz, 1999b, p. 199). Sometimes we give up a value because there is a stronger case for pursuing another one. Other times, the possibility of incompatible or incommensurable reasons makes it the case that more than one of our conflicting reasons can be undefeated. Where this is

true, we are able to, in some instances, determine for ourselves what we ought to do, and thus choose the reasons for which we act.

Other times it is not our choice to make. Where exclusionary reasons are relevant we tend to have more than one undefeated reason for acting. Yet, these reasons, in constraining the ways in which one can conform to other of her reasons, determine for us which of our reasons we ought not act for, and thus which of the available goods or values we ought to forgo. As such, exclusionary reasons take the matter of whether to respond to some reasons, and so whether to instantiate some values, out of our hands. We can thus understand their appeal in explaining how a special subset of our reasons, namely our obligations (legal, moral, political, etc.) and other fo our commitments, figure into our deliberations about what we ought to do. The law, as an example, is one such reason. The law purports to not only give its subjects reasons to act in some way (Raz, 1999b, p. 155; Coleman, 2011, p. 78), but to set things straight by telling people

> "this is what you should do and whether you agree that this is so or not, now that it is the law that you should you have the law as a new, special kind of reason to do so." (Raz, 2010, p. 7)

Intuitively, then, the law affects what one has reason to do; it makes a normative difference in the lives of its subjects. A law that prohibits tax evasion, for example, makes it the case that I ought not submit fraudulent tax return forms. And those subject to this law, at least some of the

time, take their legal obligations to be reasons to do (or not do) as the law prescribes.[43] Call this the *Normative difference intuition.* The law can also guide it's subjects to the appropriate response (i.e. action), and thus explain why they acted as they did. It can move me to act as it requires, and when I do so I might, by way of an explanation, say "I claimed the additional benefit on my tax return because it's the law". Call this the *Guiding reasons intuition*. These intuitions tell us that we can recognize our obligations as giving us reason to act in some way, and that they can move us to these actions. The law can do this, but so too can our promises, social rules, commands, rights, and other obligation-generating reasons.

Importantly, these 'special reasons' do more than recommend or present a case for certain actions. That is, they make demands on us (each to varying degrees). Call this *The demandingness intuition*. Our obligations make it the case that it is no longer up to us to judge whether doing the required thing would be best, *all things considered*. It just is, they claim. But how? The explanation that Raz offers involves exclusionary reasons, which he also refers to as pre-emptive reasons. Exclusionary reasons are meant to provide us with a rationalist analysis of this distinctive feature of obligations, and the distinct role they are thought to play in our practical deliberations. As reasons to *refrain from acting for some of our reasons*, they protect some of our reasons and turn them into demands. An obligation, then, is what Raz terms a protected reason—both a first order reasons for acting (i.e. for fulfilling the obligation) and an exclusionary reason to refrain from acting for reasons that recommend otherwise (Raz, 1990, p. 199). This is, Raz tells us, how we should understand the pre-emptive nature of obligations, and

[43] That, at least some of the time, people take their legal obligations to be reasons for which they act is famously argued by Hart in his objection against Austin's sanction theory. See. H.L.A. Hart's (3d ed. 2012) *The Concept of Law*. Oxford University Press.

of some of our commitments; it is also the fundamental mark of authority. Whereas a request leaves it up to the recipient to decide whether or not they will conform, a demand "removes the decision from one person to another" (Raz, 1999b, p. 193). By taking the decision out of the recipient's hands, it pre-empts deliberation over whether or not conforming with the demand is what you ought to or have reason to do. And we now know how: exclusionary reasons represent choices that have been made for us by the exercise of another's (legitimate) normative power.[44]

In the case where I am driving on the highway and see an animal suffering on the side of the road, I have a reason to pull over (i.e. that it would relieve the animal's suffering). The law that prohibits pulling over on the highway also gives me reasons. For one, it gives me reason to obey it. Both my reason for pulling over and for obeying the law are undefeated, yet choosing which of these reason I ought to act for is not up to me. The law also gives me reason to exclude considerations in favor of law-breaking. In doing so, it demands that—whatever reasons I have for acting—I not break the law. It *obliges* me to obey it. It tells me that the choice of whether or not to pull over has already been made for me by legal officials with legitimate legislative powers (Gardner, 2012, p. 30; Raz, 1994, p. 215). By rationally constraining which of the available reasons we ought to act for, they determine which of the available goods or values we ought to forgo. This is the very idea of pre-emptiveness that exclusionary reasons are said to explain, and we can understand it in rationalist terms as follows: exclusionary or pre-emptive reasons make it valuable, and thus intelligible or rational, to be irresponsive to some of our reasons, and so their constituent values, on some occasions. I recognize that in conforming to the exclusionary reason, and forgoing certain of the values available to me, I can instantiate some

[44] Where an 'obligation' just is the product of an exercise of normative power (Owens, 408).

other value—the value that constitutes the exclusionary reason. And I believe as much. I take there to be some good in my being irresponsive to those values such that it is both justified and *rational*, and thus intelligible from my perspective. In figuring into our deliberations about what we ought to do as any normative reason would, these special reasons *rationally* pre-empt our deliberations.

Why understand the pre-emptive nature of obligations, and of some of our commitments, in this way? This is, as David Owens (2006) describes it, *The Question of Rationalism*:

> "Why regard exclusionary considerations as reasons, as contributing to practical deliberation in the way that a reason does? Why not say that an exclusionary consideration shapes deliberation so as to make sense of action but without providing us with reasons?" (421).

This project supplies an answer. Only when exclusionary reasons play the same role in our practical deliberations can they make sense of all three intuitions about obligations (and other of our reasons that are pre-emptive in character). We can regard certain responses as obligatory only when we take them to require that we reject other eligible, but incompatible responses. And we can take them to require that we reject other responses only if we can be moved to do so. In complying with the nexus, exclusionary reasons—as normative reasons—explain how obligations are made for rational beings like us. Just as first-order reasons for action do, exclusionary reasons can move us to do as they require (i.e. *to refrain from acting for some of our reasons*). As such, the nexus completes Raz's rationalist analysis of reasons that are pre-emptive in character: in addition to recognizing reasons to fulfill our obligations, we can be

moved to regard these reasons as being pre-emptive because we can (and can be expected to) refrain from acting for reasons since we, as rational agents, can engage with features of the world that make doing so valuable; that is, we can recognize them, and respond appropriately.

Exclusionary reasons as normative reasons captures the sense in which laws, promises, commands, etc. make not acting for some reasons valuable, and the sense in which we can, as rational agents, choose (i.e. intend) to instantiate these values when we follow the law, keep promises, and obey commands (i.e. when they refrain from acting for reasons that recommend doing otherwise). With a coherent account of their normativity—one that is value-based—in place, this clears the way to consider the values that might constitute exclusionary reasons, and thus justify being responsive to the exercise of another's (i.e. a legal authority's, promisor's, commander's, etc.) normative power. While identifying putative values exceeds the scope of this project, *Getting Exclusionary Reasons Right* makes clear how fully rational agents can ever be moved to act for, or out of an awareness of their obligations. They can do so, I submit, when the reasons our obligations generate are value-based, and when agents can justify their actions by appeal to these values.

References

Adams, N. P. (2021). In defense of exclusionary reasons. *Philosophical Studies*, *178*(1), 235-253.

Alvarez, M. (2010). *Kinds of Reasons: An Essay on the Philosophy of Action*, Oxford: Oxford University Press.

Anscombe, G.E.M. (1957). *Intention*, Oxford: Blackwell.

Arpaly, N. (2002). Moral worth. *The Journal of Philosophy, 99*(5), 223-245.

Arpaly, N. and Schroeder, T. (2013) *In Praise of Desire*. Oxford: Oxford University Press.

Bennett, J. (1990). Why is belief involuntary?. *Analysis, 50*(2), 87-107.

Brown, J. (1977). Moral theory and the ought—can principle. *Mind, 86*(342), 206-223.

Coleman, J. L. (2013). The architecture of jurisprudence. In *Neutrality and Theory of Law* (pp. 61-126). Springer, Dordrecht.

Dancy, J. (2004). *Ethics without principles*. Clarendon Press.

Dancy, J. (2000). *Practical reality*. Oxford University Press, USA.

Davidson, D. (1969). "How Is Weakness of the Will Possible?". In J. Feinberg (Ed.), *Moral Concepts* (pp. 21–42). Oxford University Press.

Essert, C. (2012). From Raz's Nexus to Legal Normativity. *Canadian Journal of Law & Jurisprudence, 25*(2), 465-482.

Gans, C. (1986). Mandatory rules and exclusionary reasons. *Philosophia, 15*(4), 373–394.

Gardner, J. (2012). How law claims, what law claims. In Matthias Klatt (ed.), *Institutionalized Reason: The Jurisprudence of Robert Alexy.* Oxford University Press.

Gert, J. (2007). Normative strength and the balance of reasons. *The Philosophical Review, 116*(4), 533-562.

Gibbons, J. (2010). Things that Make Things Reasonable. *Philosophy and Phenomenological Research, 81*(2), 335–361.

Green, L. (2005). Three Themes from Raz. *Oxford Journal of Legal Studies, 25*(3), 503–523.

Hare, R.M. (1952). *The Language of Morals*, Oxford: Clarendon Press.

Hatzistavrou, A. (2012). Motivation, Reconsideration and Exclusionary Reasons. *Ratio Juris, 25*(3), 318-342.

Hart, H. L. A., Raz, J., & Green, L. (2012). *The concept of law*. Oxford University Press.

Hieronymi, P. (2006). Controlling attitudes. *Pacific Philosophical Quarterly, 87*(1), 45-74.

Jordan, A. (2018). Exclusionary Reasons, Virtuous Motivation, and Legal Authority. *Canadian Journal of Law & Jurisprudence, 31*(2), 347-364.

Kavka, G. (1983). The Toxin Puzzle. *Analysis, 43*(1), 33–36.

Kelly, T. (2002). The rationality of belief and some other propositional attitudes. *Philosophical Studies, 110*(2), 163-196.

Kolodny, N. (2005). Why be rational?. *Mind, 114*(455), 509-563.

Korsgaard, C. M. (2008). *The constitution of agency: Essays on practical reason and moral psychology.* OUP Oxford.

Lord, E. (2015). Acting for the right reasons, abilities, and obligation. *Oxford studies in metaethics*, *10*, 26-52.

Markovits, J. (2010). Acting for the right reasons. *Philosophical Review*, *119*(2), 201-242.

Moore, M.S. (1989). Authority, Law, and Razian Reasons. *Southern California Law Review*, *62*(4), 827-896.

Owens, D. (2008). Rationalism about obligation. *Available at SSRN 1159183.*

Parfit, D. (2011). *On What Matters*, Oxford: Oxford University Press.

Piller, C. (2006). Kinds of Practical Reasons: Attitude-Related Reasons and Exclusionary Reasons. *Analyses*, 98-105.

Pillar, C. (2001, July). Normative practical reasoning: Christian pillar. In *Aristotelian Society Supplementary Volume* (Vol. 75, No. 1, pp. 195-216). University College London: The Aristotelian Society.

Raz, J. (1994). Authority, Law, and Morality. In *Ethics in the public domain: essays in the morality of law and politics* (pp. 210-238). Oxford University Press.

Raz, J. (2010). *Between authority and interpretation: On the theory of law and practical reason.* OUP Oxford.

Raz, J. (1990). Authority and justification. In J. Raz (Ed.), *Authority* (pp. 115-141). Oxford: Basil Blackwell.

Raz, J. (1999a). *Engaging reason: On the theory of value and action.* Oxford University Press.

Raz, J. (1988). Facing up: A reply. *Southern California Law Review*, *62*, 1153.

Raz, J. (2011). *From normativity to responsibility.* Oxford University Press.

Raz, J. (2021). Exclusionary Reasons. Unpublished Manuscript. Retrieved from https://ssrn.com/abstract=3933033 or http://dx.doi.org/10.2139/ssrn.3933033

Raz, J. (1999b). *Practical reason and norms*. OUP Oxford.

Ross, W.D. (2002). *The right and the good.* Oxford: Clarendon Press.

Scanlon, T.M. (2009). *Moral Dimensions: Permissibility, Meaning, Blame.* Cambridge, MA: Harvard University Press.

Scanlon, T. (2000). *What we owe to each other*. Belknap Press.

Shah, N. (2006). A new argument for evidentialism. *The Philosophical Quarterly, 56*(225), 481-498.

Scheffler, S. (2004). Projects, relationships and reasons. In R. J. Wallace, P. Petit, S. Scheffler, & M. Smith (Eds.), *Reason and value: Themes from the moral philosophy of Joseph Raz* (pp. 274-269). Oxford University Press.

Whiting, D. (2017). Against Second-Order Reasons. *Noûs*, *51*(2), 398-420.

Williams, B. (1973). Deciding to Believe. In *Problems of the self: philosophical papers 1956-1972* (pp. 136-151). Cambridge University Press.

Williams, B. (1981). *Moral luck: philosophical papers 1973-1980*. Cambridge University Press.