Responsibility as Reasons-Possession

KENTA SEKINE

UCL
Philosophical Studies MPhil Stud
2020

I, Kenta Sekine, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

_____

Kenta Sekine
September 8th, 2020

ABSTRACT

In this dissertation I defend a theory of moral responsibility I call the *Reasons View*, according to which our responsibility for things is principally a matter of our possessing normative reasons bearing on those things. More specifically, my claim is that S is morally responsible for φ-ing if and only if S has a normative reason to φ that is derivatively a normative reason to effectively try to φ, where φ includes omissions. My argument for this claim develops over three Chapters. In Chapter 1 I offer an interpretation, or perhaps reconstruction, of Angela Smith's theory of 'responsibility as answerability' as expressing a Reasons View in the above sense. I defend the Reasons View against possible confusions concerning its highly permissive character, its distinction between judgments of blame- and praiseworthiness and judgments of responsibility, and what it implies about the grounds of moral response. Chapter 2 inquires after the conditions of reasons-possession using the *Reasons Implies Can* principle. I consider metaphysical, physical and powers-based readings of 'can,' and eventually settle on a powers-based reading according to which the principle states a capacity-plus-opportunity condition. In Chapter 3 I suggest Joseph Raz's work exhibits structural features of the Reasons View, and develop my own Reasons View from recognisably Razian materials. Through a consideration of disability, I argue that the responsibility-relevant normative reasons to φ are those that are derivatively normative reasons to effectively try to φ. I go on to consider how, in the presence of such normative reasons (and thus responsibility), privative manifestations of rational powers relate to fault.

IMPACT STATEMENT

This dissertation is intended to contribute to ongoing discussions in the philosophical literature on moral responsibility, including the proper place of accidents and moral luck in the lives of responsible agents, the relationship between disability and moral responsibility, and the interpretation of influential thinkers in the field such as Joseph Raz and Angela Smith.

ACKNOWLEDGEMENTS

CONTENTS

INTRODUCTION

The philosophical literature on moral responsibility is vast, constantly evolving, and terminologically and conceptually complex. It can be difficult to know where to start. Even when a starting point has been decided, what governs the philosophical relevance of considerations drawn upon to justify the decision can be elusive. As a result of this, starting points for discussions of moral responsibility quite often have an air of arbitrariness about them, and I am not sure that mine is any different.[1]

One upside to the vastness of the literature is that there is no shortage of excellent accounts on which to draw and to build. In this dissertation I have decided that Angela Smith's *answerability* view of moral responsibility is as good a starting point as any. My justification is that, to my mind at least, it is self-consciously amongst the sparser and more minimal of the available views on moral responsibility. As such, it makes efforts to assume very little about the topic *ex ante*, aspiring not to say any more than it is entitled to say, with a view to keeping separate concepts separate. In a literature that often dives into thought experiments with seemingly more moving parts than have been earned, the philosophical benefits of taking Smith's answerability view as a starting point are hopefully discernible.

Ultimately, as will become clear, it is this starting point that explains why I called my dissertation *Responsibility as Reasons-Possession*. As this title suggests, my general claim is that our moral responsibility for φ is principally a matter of possession of normative reasons bearing upon φ. Since a great many φ are like this, the resultant theory of moral responsibility, which I dub the *Reasons View*, is highly permissive about what we are morally responsible for. This permissiveness is what allows the Reasons View to handle our moral responsibility for accidents in a satis-

---

[1] For a recent and admirably direct expression of the sort of arbitrariness I have in mind, see Gideon Rosen (2015, 66).

fying way. Much of the labour of this dissertation is to show that the Reasons View, although highly permissive, is not onerously so. By the dissertation's end, my specific claim will emerge as the claim that it is necessary and sufficient for moral responsibility for φ that we possess a normative reason bearing upon φ that is derivatively a normative reason to effectively try to φ, or to effectively try to omit to φ (as the case may be).

Here is a broad outline of the structure of the dissertation. Each Chapter comes with an Introduction and Conclusion to orient the reader on the finer details of each Section. As suggested, Chapter 1, 'Responsibility as Answerability,' argues that the core insight of Smith's view is that our moral responsibility for φ is principally determined by our possessing normative reasons bearing on φ. My argument proceeds via close reading of certain passages in Smith. Admittedly the results may come as a surprise, so I leave it open as to whether they are Smith's own results or my own Smith-inspired results.

The Reasons View that emerges analyses moral responsibility in terms of reasons-possession. But unless the Reasons View can help us say what people are morally responsible for in specific contexts, it is arguably unsatisfying from a richer theoretical perspective. It would be more satisfying if it were supplemented by a workable account of the conditions of reasons-possession. The task of Chapter 2, 'The Conditions of Reasons-Possession,' is to develop one such account. My strategy is to assume the validity of the *Reasons Implies Can* principle, and to shift through metaphysical, physical and powers-based interpretations of 'can,' so as to supply the maximally non-trivial conditions of reasons-possession. I conclude that the *best* Reasons View (i.e. the one most satisfying from the richer theoretical perspective) operates with an *actualist two-way powers*-interpretation of 'can.' Someone 'can' do something, in this sense of 'can,' when they have both the power to do that thing and the opportunity to do it.

Prima facie, the combination of commitments I call the best Reasons View can appear rather outlandish. This understandably lends it to easy dismissal as a serious contender in the theory of moral responsibility. In Chapter 3, 'A Razian Reasons View,' I interpret and build upon Joseph Raz's influential work as just such a Reasons View. Doing so allows a strengthening of the Reasons View, mentioned above, to the effect that our moral responsibility for φ is a matter of possessing normative reasons bearing on φ that are derivatively normative reasons to effectively try to (omit to) φ. It also offers a neat way of introducing familiar practical-philosophical terms of

art, including non-accidents (intentional or reckless), accidents (pure, careless or negligent), fault and blameworthiness, and excuses. The cumulative effect of all this, I hope, is to show the Reasons View more palatable than it might at first seem.

The dissertation concludes with a brief overview of findings, and a statement of potential avenues for future research as recommended by shortcomings of the account.

Responsibility as Answerability

1. Introduction

Though Angela Smith's *answerability* view of moral responsibility has, in my opinion, a great deal going for it, her presentation of it is not always as clear as could be desired. My aim in this Chapter is to closely examine what she says, so as to display the view in its barest bones. Hopefully this will leave adherents and detractors alike in no two minds about what exactly accepting Smith's view amounts to. Doing so will inevitably involve a degree of selective hearing. Some of Smith's looser comments will be discarded, in a principled way, so as to better present what I take to be the core insight of her view. The results may come as a surprise. Given this fact, I leave it to the reader to decide whether the Chapter as a whole is better understood as an interpretation of Smith's view or a reconstruction heavily indebted to it.

In Section 1.1, I refine the condition on moral responsibility Smith presents under the rubric of 'intelligibility.' The major upshot of this Section is that the possession of normative rather than motivating reasons is of primary importance to moral responsibility. After spelling this out, I consider and reject three objections to the effect that the picture of moral responsibility that emerges is overly permissive. In Section 1.2, I go on to refine the apparently quite separate condition Smith presents under the rubric of 'eligibility.' The major upshot of this Section is that on closer inspection the two apparently separate conditions amount to one and the same, namely, that possession of normative reasons is what principally determines moral responsibility. I go on to present and deflate a potential confusion about the 'eligibility' condition that arises from its slightly different outward appearance to the 'intelligibility' condition. Throughout the Chapter, recurring themes include the vicissitudes of the language of 'answerability,' the pervasive

urge to conflate responsibility with blame- and praiseworthiness, and the place of agent-regret and associated rational phenomena within human life.

1.1. The Intelligibility Condition

In the most recent and authoritative statement of her views on moral responsibility, Smith writes,

"[T]o say that an agent is morally responsible for something is to say that **(A)** it would be *intelligible* to ask her to 'answer for' that thing—to give her (justificatory) reasons for thinking, feeling, or acting in the way she has […]."

She continues,

"[…] and that **(B)** she is *eligible for* certain moral responses based on the nature of the thing in question and the quality of the justificatory reasons she is able to offer in support of it." (2015, 103)[2]

Though Smith calls her view the *answerability* view, in my opinion this label tends to prejudge too much about moral responsibility, so I shall refrain from using this terminology. My reasons will hopefully become apparent as we go along. The passage above, reflecting its topic, is densely packed with the kind of concept whose edges seem instantly to lose definition under closer scrutiny. It makes sense to unpack the passage slowly, in the smaller chunks I have marked **(A)** and **(B)**.

Taken independently of chunk **(B)**, what appears under the rubric of 'intelligibility' in chunk **(A)** reads most naturally as a necessary and sufficient condition on moral responsibility. At a

---

[2] Other important papers where Smith establishes this view are her 2005, 2008 and 2012.

first pass, the idea is something like this. For agents S and things (thoughts, feelings or actions) φ,[3]

INTELLIGIBILITY$_1$: S is morally responsible for φ iff (i) S φs, and (ii) S is intelligibly asked to give their (justificatory) motivating reasons for φ.

This condition says that agents are morally responsible for all and only those things to which they stand in the relations specified in clauses (i) and (ii) of its right-hand side. Clause (i) is satisfied when the relevant agent does the thing in question, and clause (ii) is satisfied when they are intelligibly asked to justify doing that thing. Although (i) may look so obvious as to be threatened with redundancy, it is in fact very important. It is what allows INTELLIGIBILITY$_1$ to handle cases where having intelligibly been asked to justify φ, S rightly denies that φ is attributable to them. This happens for example in legal proceedings when the defendant, having received a court summons for their alleged legal wrongdoing, successfully pleads a denial to the effect that their conduct did not satisfy the elements of the legal wrongdoing they have been summoned to court for. That they were summoned shows it was intelligible, in some sense of 'intelligible,' to ask the defendant to justify conduct alleged to amount to the relevant legal wrongdoing, though their successful denial shows that no such wrongdoing is attributable to them. This illustrates the presence of two senses of 'intelligibility,' one weak and one strong. In our example, though it is intelligible in the weak sense of (ii) that S is asked to justify φ, it is not intelligible in the strong sense of INTELLIGIBILITY$_1$, since (i) is not satisfied. Smith's view therefore allows that agents and things do not satisfy INTELLIGIBILITY$_1$ despite there being a genuine sense in which it is intelligible to ask them to justify those things.

There is another sort of case that it makes sense to introduce here, if only to register it for more detailed consideration much later on. Those who experience forms of what, for better or

---

[3] In this dissertation, 'things' is often used as a catch-all term for whatever objects of moral responsibility Smith takes her view to cover (viz. ways of "thinking, feeling, or acting"), and should not be read as ontologically loaded. 'Things' also covers *omissions*, i.e., absences of thought, feeling or action. Though this may be controversial, for sustained critique of the opposite tendency see John Martin Fischer and Mark Ravizza (1998, 145-50). Where talk of 'things' is strained and reference to 'thoughts, feelings or actions' too cumbersome, I will simplify by speaking in terms of 'actions,' and concomitantly of 'doing φ'—but this should be read as standing in for any type of object of moral responsibility unless otherwise qualified. Whether an instance of φ is an omission or not should be clear enough from the context.

for worse, gets labelled 'disability,' are not commonly considered morally responsible for things falling within the range of their form of disability. In Smith's terms, this is to say that for the relevant φ it is not intelligible in the strong sense of INTELLIGIBILITY$_1$ to ask those with disability to justify φ. It is, however, intelligible in the weak sense of clause (ii) to ask those with disability to justify φ. Consider for example someone who φs, for whom φ falls within the range of their form of disability, and whose form of disability is invisible. Ordinarily it is perfectly intelligible, in the weak sense, for someone ignorant of the disability to request a justification for φ. Just like those S who rightly deny that the φ in question is attributable to them, those S who are disabled with respect to φ satisfy (ii) but do not satisfy INTELLIGIBILITY$_1$. But unlike them, those with disability satisfy (i), at least in its present formulation. This situation forces an *ad hoc* amendment to clause (i) of INTELLIGIBILITY$_1$ along the following lines.

> INTELLIGIBILITY$_1$´: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S is intelligibly asked to give their (justificatory) motivating reasons for φ.

The amendment is *ad hoc* because 'disability' is left unanalysed and intuitive. As indicated, this will be remedied later on in the discussion of Section 3.1. For the time being, however, this will have to do. In any case, what matters to the present discussion is simply that in order to give chunk **(A)** a fighting chance of being the necessary and sufficient condition it appears to be, something like this amendment needs to be made, and it is best rolled into clause (i) of the condition.

Let us now consider clause (ii). Under what conditions are agents (weakly) intelligibly asked to give justificatory motivating reasons for things? Notice that *what* must be intelligible is the asking of a question, addressed to S, that asks S to give their justificatory motivating reasons for whatever φ is under consideration. What must be intelligible will be familiar as the special sense of the 'Why?'-question famously studied by Elizabeth Anscombe (1957, §5). Let me recast INTELLIGIBILITY$_1$´ to reflect this observation.

> INTELLIGIBILITY$_2$: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S is intelligibly asked the 'Why?'-question about φ.

This constrains what can be substituted into S and φ such as to satisfy (ii). Say we substitute *standing there* into φ and hold it fixed. If we now substitute into S a chair, or a tree, then it will not be intelligible to ask S why they are standing there, even if they are in fact standing there. This is not because there are no reasons why a chair, or a tree, might be standing there. There certainly are always facts (whether efficient-causal or teleological) that explain why a chair, or a tree, is standing there, if it is indeed standing there. Neither is it because chairs and trees are incapable of offering true explanations of their states or movements by reporting the relevant facts. Even if they somehow became endowed with this capacity (and this capacity alone), this would not make it intelligible, in the sense at hand, to ask them why they are standing there. Even for an S that is plausibly intelligibly asked 'Why?' about some things, like an adult human (say), if the only true explanation of φ is merely efficient-causal or teleological, their capacity to report these facts hardly makes it intelligible to ask them why they φ, in the relevant sense of the question. The genetic makeup or peristaltic movements of an adult human are both cases in point.

These constraints gesture towards the conditions under which clause (ii) of INTELLIGIBILITY$_2$ is satisfied. In order for agents to be intelligibly asked the 'Why?'-question about certain things, those things need to have more than just a true explanation, and indeed more than just a true teleological explanation in virtue of which they are evaluable. Rather they must be evaluable in the light of explanatory facts that are *normative* reasons, and the relevant agents must *have* those normative reasons. In other words, for a given S and φ, clause (ii) is satisfied when S has at least one normative reason in the light of which φ is rationally evaluable, where this includes both normative reasons for φ and normative reasons against φ. Collecting together all such normative reasons and calling them 'φ-reasons' for short, the whole condition can be recast as follows. For agents S such as to have normative reasons and rationally evaluable things φ,

INTELLIGIBILITY$_3$: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S has a φ-reason.

Let me clarify what it means to say S 'has' or 'possesses' φ-reasons. Since φ-reasons are normative reasons either for or against φ, to say S has a φ-reason means S has a normative reason speaking either for or against S's φ. That S is said to 'have' such a reason does not imply that S stands in an epistemic relation to it. For arbitrarily high n, it may be that S has n φ-reasons, yet also that S lacks awareness of every single one of them. Suppose that in some such situation, S comes to φ. In that case, it cannot be that S's φ was motivated by one of their φ-reasons, since this would require that S takes at least one of their normative reasons to speak in favour of S's φ, which implies, contrary to hypothesis, that S stands in an epistemic relation to that φ-reason. In other words, to say that S has φ-reasons does not imply, if S actually φs, that S φs for one of their φ-reasons. Therefore to say S 'has' or 'possesses' φ-reasons does not collapse the distinction between reasons in their normative and motivating roles. Having said that, it is perfectly possible that reasons figure in both roles. Indeed this possibility is essential to *justification*. When S φs for a genuine φ-reason, so that the φ-reason is both normative and motivating, and the φ-reason is moreover *undefeated* (i.e. it is normatively sufficient), then S's φ is justified by that φ-reason. When S has an undefeated φ-reason, but is not moved to φ for that φ-reason, I will say that S's φ is justifi*able* but not justified by that φ-reason.

These remarks are consonant with Anscombe. If S's having a φ-reason is what makes the 'Why?'-question intelligible—or as Anscombe likes to say, gives it 'application'—then it should be obvious why she thinks it "is not refused application because the answer to it says that there is *no* reason" (1957, §17). Its applicability is indifferent to the content of actual answers given because, minimally, actual answers supply information about motivating rather than normative reasons.[4] The 'no reason' response that Anscombe imagines reports the absence of motivating

---

[4] This is arguably controversial as an interpretation of Anscombe. To see why, consider that she also says "This question is refused application by the answer: 'I was not aware I was doing that'" (1957, §6). Though S's (non-observational) awareness of φ-ing does not entail S's having a motivating reason to φ, S's lacking it does entail S's lacking a motivating reason to φ. The passage therefore tends to support the claim that lacking motivating reasons is at least sufficient for non-applicability. Here, I agree with Roger Teichmann that Anscombe is not wholly clear on this issue (2008, 34-41). But whilst Teichmann takes the §6 case as firm and tries to show how the §17 case differs from it, I take my reading of the §17 case to be firm and question the §6 case as an instance of non-applicability. Admittedly my reading is less in the spirit of Anscombe, though in my defence what I am proposing is not a reading of Anscombe but an independent view about the intelligibility of asking the 'Why?'-question.

reasons.[5] This is why it is misleading when Smith says the addressability of the 'Why?'-question to someone for some thing "suggests that the thing in question must in some way reflect her own judgment or assessment of reasons" (2015, 103). Without further qualification, this gives the impression that reasons bearing on the things in question must enter into the content of the addressee's deliberations. As we have seen, however, for the 'Why?'-question to be intelligibly asked about some thing, reasons bearing on that thing need not have entered into deliberation at all. The reasons need only be normative, not motivating. For Smith to remain consistent, she must be using 'judgment' in the sense of its second occurrence in the following: that S has no judgment about a given φ-reason reflects S's judgment about that very φ-reason. Needless to say this is an idiosyncratic use of the word.

Since clause (ii) of INTELLIGIBILITY₃ appeals to normative as opposed to motivating reasons, a legitimate worry would be that the condition is far too permissive. The condition states that any thing someone does that they have a normative reason for or against is something for which they are morally responsible, quite irrespective of whether they are moved to it *for* those normative reasons. Very many things are like this, for instance, all those mundane things we do justifiably but without justification. Can it really be right to say, with Smith, that we are answerable for all these mundane things? Here we arrive at one reason for my suggestion that we exercise caution with the language of 'answerability.' True enough, it sounds wrong to say we are answerable for our tea-sippings, daydreams and omitted handstands, at least where they are justifiable. But this speaks less to the limits of our moral responsibility than to the propriety of the language of 'answerability' as applied to it. Granting this, it could perhaps still be pressed that it sounds wrong to say we are morally responsible for these things. But if 'responsible' is the problem, then I would ask what else could distinguish them categorially from things like our genetic makeup or peristalsis. Conversely, if 'moral' is the problem then I would diagnose this

---

[5] Ulrike Heuer suggested to me in conversation the possibility that Anscombe's 'no reason' response reports the absence of a normative reason. Her example: 'Why did you, mid-walk, break out into a run?'—'No reason'; and this means: the agent has no normative reason. This would be troublesome for my reading of Smith because the agent here is presumably responsible for breaking out into a run. To my mind, however, in cases like this there will always be normative reasons for or against breaking out into a run, however mundane, and whether known to the agent or not. *This* is what secures the intelligibility of 'Why?'-questioning. The agent's ignorance, which they reveal in answering 'no reason,' is neither here nor there.

as symptomatic of a distorting tendency to conflate moral responsibility with being an appropriate target of moral responses, in this case *blame*—more on which in Section 1.2.[6]

Even granting these points, the picture still might be thought overly permissive on the following grounds. Whenever someone does something justifiable (with justification or otherwise), they leave a trail of other things behind which they had normative reasons bearing upon that are left hanging in doing the justifiable thing. Is it really the case that they are morally responsible for all these other things too? There are two points to make about this. Firstly, many of these normative reasons will have been *defeated* by whichever normative reason renders the thing they in fact did justifiable. Reasons are 'defeated' when they are losers in the metaphorical battle of rational strength. It is not generally worrisome to claim the people in question nonetheless have those defeated reasons, nor that they are morally responsible for omitting to do the things those defeated reasons bear upon. So long as satisfaction of the undefeated reason does not destroy the conditions for satisfying the defeated ones, the defeated reasons assert themselves as soon as the undefeated reason is satisfied. Secondly, the remainder of the trail of hanging normative reasons will be undefeated reasons—the justifiable thing actually done will be one among many alternative things justifiably done. When this happens, it is not that the normative reasons bearing on these things come out joint winners in the metaphorical battle. It is rather that the values of what they bear upon are incommensurable, so that they are not such as to compete at all. Value incommensurability tends to call to mind matters of gravity in people's lives. Love or career? Security or adventure? But equally, and perhaps less obviously, it applies just as much to the mundane. Doritos or Pringles? Scenic route or fast route? Whether defeated or undefeated, then, pointing to the trail of hanging reasons is a poor ground for objection. Without it what remains is an unfamiliar picture of human life, swashbuckling to a fault, such that the question of agent-regret cannot be heard when all goes well.

Moving on, a similar and equally legitimate worry about INTELLIGIBILITY₃ is raised by what I call 'deviant cases.'[7] Deviant cases resemble mundane cases insofar as clauses (i) and (ii) of the

---

[6] This is a diagnosis made in the good company of Bernard Williams (1993, 67-70) and Joseph Raz (2011, 251). Smith is for the most part clear-eyed about the relevant distinctions between 'basic' moral responsibility, liability to certain kinds of moral response, and our practices of actual moral response (2015, 107-8). I say 'for the most part' because Smith occasionally errs on this point, which, as we will see in Section 1.2, tends to obscure what adopting her proposal would actually commit us to.

[7] For more on deviant cases, see the discussion of moral luck in Section 3.2.

right-hand side are satisfied. But with deviant cases, though S φs (not from disability), there is a deviance in the connection between S's φ and the relevant φ-reason. Deviant cases come in two kinds. Cases of 'happy' deviance involve someone doing the right thing for the wrong reasons. More precisely, though S φs justifiably, the φ-reason that actually motivates S's φ is defeated—it is normatively insufficient. Again, if we impart philosophical significance onto the odd ring of calling someone 'answerable' for doing the right thing, then here, as before, the prejudicing language of 'answerability' has led us astray. Similarly, the infelicity may be perceived as lying in judging someone 'morally responsible' for what is effectively an accident. But once again: to judge someone morally responsible for a happy accident falls far short of judging them *praiseworthy* for it. To see this, suppose we posed the 'Why?'-question and S's dodgy motivating reasons came tumbling out. Whilst on the one hand this would reveal S as unworthy of praise for S's φ, on the other, it would concede φ as something S is morally responsible for. In a concession to Smith's terminology, we could say that S's 'answerability for' φ—however misleading this may be where φ is the right thing—presupposes S's *answerability to reasons*, where this means φ-reasons S possesses (here a normative reason for φ). To my mind, this represents the best reason for Smith to call her view the 'answerability' view, though she by no means foregrounds this line of thought.[8] That Smith's choice of terminology allows for this concession is therefore itself a happy accident.

Similar points apply to cases of 'unhappy' deviance, where someone does the wrong thing for the right reasons. To be more precise, in unhappy cases, though S's φ is motivated by an undefeated φ-reason, φ is not a justifiable thing for S to do. How this is possible becomes apparent when we recall that φ-reasons may not only be normative reasons for φ but normative reasons against φ. It can be difficult to imagine how S's φ could be motivated by a normative reason against φ. But it happens: in circumstances best described as unhappy accidents. The language of 'answerability' is admittedly more at home with unhappy accidents, since the relevant φ are negatively valenced. Nevertheless it may be thought that any account that makes us 'answerable for' unhappy accidents—or indeed 'morally responsible' for them—sounds intolerably mean-spirited and harsh. To allay these concerns, suppose that under 'Why?'-questioning S

---

[8] By contrast, John Gardner brings answerability to reasons to the fore (2017, 16). Note that Gardner considers his position to have strong affinities with Smith's (*op. cit.*, 7n16).

reveals due respect for the normative sufficiency of the relevant undefeated φ-reason. Indeed, with unhappy accidents this is always so. By an inversion of cases of happy accident, this time S answers for φ in a manner that absolves their blameworthiness, and that it can have this effect presupposes S's answerability to, their possession of, the relevant φ-reasons. So whilst there need be nothing obnoxiously blame-like going on, this fact carries no suggestion that S's moral responsibility for φ is in any way diminished.

It may be thought that I have considered only half the cases with the character of what I am calling 'unhappy' deviance. Wrong things are done for right reasons not only when someone accidentally does the wrong thing in a sincere bid to avoid doing that very thing, but also when someone accidentally does the wrong thing in a sincere bid to do something else deemed right. In the latter sort of case, S unjustifiably ψs, despite being motivated to ψ by an undefeated normative reason for φ. Such cases may be thought problematic for INTELLIGIBILITY₃ insofar as the condition makes mention only of φ, saying nothing about ψ. In other words, it appears that in the cases just described, S will not come out as morally responsible for ψ according to INTELLIGIBILITY₃. On closer inspection, however, INTELLIGIBILITY₃ covers these other cases too. Wherever ψ is genuinely wrong (i.e. unjustifiable) and there is an undefeated normative reason in favour of something else, φ, this can only be because the facts speaking in favour of φ are the very facts speaking against ψ. Therefore, whenever S's ψ is unjustifiable and yet motivated by a normative reason for φ, S's ψ is unjustifiable and yet motivated by a normative reason against ψ. That is: S ψs (not from a disability), and S has a ψ-reason. And since 'ψ' and 'φ' are merely schematic placeholders, in all such cases clauses (i) and (ii) of INTELLIGIBILITY₃ are satisfied. Therefore, S also comes out morally responsible for the accidental thing in the other half of cases of unhappy deviance.

To recap, in this Section I have attempted to extract chunk **(A)**'s core claim. After some refinement, I ended up with the condition stated by INTELLIGIBILITY₃. Finally I considered and rejected three objections to the effect that INTELLIGIBILITY₃ is overly permissive: from cases of mundane but justifiable things, from the trail of unsatisfied normative reasons left hanging in the wake of things done justifiably, and from deviant cases. It will not have escaped notice that consideration of these objections brought us well within the territory of chunk **(B)**. Let us now

spend some time on chunk **(B)**, before seeing where this leaves Smith's view of moral responsibility.

1.2. The Eligibility Condition

Just as with 'intelligibility,' what appears under the rubric of 'eligibility' in chunk **(B)** also appears to be a necessary and sufficient condition of moral responsibility. The gist of what it says is straightforward enough. At a first pass, and incorporating the *ad hoc* amendment to clause (i) from before, for agents S and things φ,

> ELIGIBILITY$_1$: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S is eligible for moral responses based on φ's nature and the quality of S's (justificatory) motivating reason for φ.

Though the gist of ELIGIBILITY$_1$ is clear, I want to suggest that clause (ii) harbours an ambiguity that makes it susceptible to two readings, both of which, in their own way, are problematic for Smith. The general shape of the issue is as follows. Whereas on the first reading, ELIGIBILITY$_1$ faces the problem of standing in tension with INTELLIGIBILITY$_3$, on the second reading, it faces the opposite problem, that it no longer states a separate condition from INTELLIGIBILITY$_3$ at all. The ambiguity giving rise to the issue is about *what* exactly clause (ii) wants S to be eligible for.

To anticipate a little, in this Section I come down in favour of the second reading. On one level this should be unsurprising, since the conditions stated in chunks **(A)** and **(B)** are both conditions on the same thing (viz. moral responsibility), and both purport to be necessary and sufficient. So of course they cannot state separate conditions. But on another level it should be at least somewhat surprising. It is hard to shake off the impression that Smith is presenting not one but two separate conditions. Before saying too much, let us consider the two readings of clause (ii).

On the first and probably more natural reading, (ii) wants S to be eligible for whatever moral responses are deserved given φ's nature and S's motivating reason for φ. So if S's φ is wrong and S's motivating reason for φ is shoddy, then (ii) wants S to be eligible for something like blame. Or if S's φ is, say, supererogatory and S's motivating reason for φ is impeccable, then (ii) wants S to be eligible for something like praise. What does it mean for someone to be 'eligible' for blame or praise? In ordinary parlance we would say they were blameworthy or praiseworthy.[9] To be blame- or praiseworthy is to be an appropriate target, that is, to be deserving or worthy, of actual blaming or praising behaviours. This of course leaves the possibility that though someone is an appropriate target for actual such behaviours, they are withheld or attenuated on the basis of further and relatively independent considerations. Someone is 'eligible' for moral responses, on this reading, in the sense of being deserving of them, whether or not the deserved moral responses ever materialise. Let me recast ELIGIBILITY₁ to make this explicit.

ELIGIBILITY₂: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S is deserving of moral responses based on φ's nature and the quality of S's (justificatory) motivating reason for φ.

Now we are in a position to see how ELIGIBILITY₂ stands in tension with INTELLIGIBILITY₃. Recall that according to INTELLIGIBILITY₃ someone is morally responsible for all and only those things for which they have normative reasons speaking in favour of or against. But according to ELIGIBILITY₂ someone is morally responsible for all and only those things for which they are blameworthy, or praiseworthy, or worthy of some other moral response besides. As we saw in our discussion of mundane and deviant cases (see Section 1.1), the former casts its net considerably more widely than the latter. Since this means that pairs of agents and things satisfying these two conditions are not coextensive, one or the other condition must be discarded, at least as presently formulated. Given the tenor of what Smith says about 'eligibility,' and how clearly she intends for it to cohere with 'intelligibility,' it makes sense to discard ELIGIBILITY₂ for the purposes of understanding her view. This move is straightforwardly supported by claims to the

---

[9] Or in Michael Zimmerman's (2002) less ordinary parlance, that they are the objects of 'hypological' judgments. I will use this terminology where convenient.

effect that "questions of basic moral responsibility need to be distinguished from questions of moral blameworthiness and praiseworthiness" (2015, 108). It would be strange for Smith to offer ELIGIBILITY2 whilst saying things like this.[10] Charity demands that we reject the first reading of clause (ii).

The second reading of what exactly (ii) wants S to be eligible for is arguably less natural than the first. It demands that S is eligible for it to be determined, based on φ's nature and the normative quality of S's motivating reason for φ, what moral responses are deserved. On this reading, the whole condition reads:

ELIGIBILITY3: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S is eligible for it to be determined, based on φ's nature and the quality of S's (justificatory) motivating reason for φ, what moral responses are deserved.

What ELIGIBILITY3 wants S to be eligible for looks like some sort of procedure or inquiry that appeals to S's reason for φ to determine what S deserves. This is admittedly opaque compared to the first reading, at least on an intuitive level. To get a clearer sense of what this might mean, it helps to draw an analogy between ordinary moral practices of responsibility-attribution and criminal law procedure.

For our purposes, criminal law procedure is adequately caricatured as moving through three stages: the *summons*, *trial* and *sentencing* stages. At the summons stage, the defendant is brought before the criminal courts to explain themselves (usually via legal representatives) in the currency of reasons as best they can for the crime alleged. That this happens usually suffices to show it intelligible to request justification for their conduct, legally speaking.[11] It raises the question of whether the defendant deserves sanctions under criminal law. Outside the law, the analogue is this: that someone is morally responsible raises the question of whether they are deserving of

---

[10] Having said that, Smith does sometimes say things more in line with ELIGIBILITY2. For example, "A morally answerable agent is *eligible* for reproach for her moral transgressions, even if she is not eligible for reproach from just anyone" (ibid.). Presumably, it is the italicised occurrence of 'eligible' that Smith thinks captures her meaning. But it is hard to read "*eligible* for reproach" as anything other than 'blameworthy,' so that when it occurs after "a morally answerable agent is…" it carries the suggestion that all judgments of basic moral responsibility are hypological judgments.

[11] Here, 'intelligible' carries the strong sense. I say 'usually' because sometimes it turns out only to have been 'intelligible' in the weak sense. See Section 1.1 for this distinction.

moral responses *tout court*. At the trial stage the defendant is interrogated over their conduct, and the quality of their motivating reasons, if all goes well, is laid bare for all to see. With these in plain view, the question raised during the summons stage can be settled—'Guilty! Legal sanctions are in order' (say). But a new question arises as to what legal sanction is in order, and how much of it. The analogue outside the law is trial by 'Why?'-questioning, which lays bare the reasons required to determine deservingness of a moral response of some kind—'Blameworthy! This person is an appropriate target for blaming behaviour' (say). Finally, at the sentencing stage, legal adjudicators appeal to relatively independent considerations to determine, often to temper, the quantity and quality of the actual legal sanction meted out. This answers the question raised at the trial stage. Similarly, in ordinary morality, we decide in what form and how much blame (say), if any at all, is to be meted out to such-and-such blameworthy person. Relevant considerations include *inter alia* the blameworthy person's background, the social proximity of the judge's life to that of the blameworthy person, and indeed the judge's own moral standing to blame.

My suggestion is that the difference between the two readings of clause (ii) is best understood in terms of the difference between the moral analogues of the trial and summons stages in criminal law procedure, respectively. Whether someone is morally responsible for something according to the first reading (ELIGIBILITY₂) corresponds to the question raised by the moral analogue of the trial stage. Whether someone is morally responsible for something according to the second reading (ELIGIBILITY₃) corresponds to the question raised by the moral analogue of the summons stage. As we saw, when the latter question is rightly raised, it must be settled by the very 'Why?'-questioning whose criminal law analogue is the interrogative format of the trial stage. Reasoning by analogy, it follows that 'eligibility' as it figures in clause (ii) of ELIGIBILITY₃ means eligibility for 'Why?'-questioning, quite apart from whatever kind of moral response is determined to be deserved in accordance with the motivating reasons thereby elicited. In short, and to be clear: *what* ELIGIBILITY₃ wants S to be eligible for is 'Why?'-questioning.

At this point, it is natural to think any difference between 'eligibility for' and 'intelligibility of' 'Why?'-questioning must be merely verbal. This is nearly right, but strictly speaking inaccurate. Criminal law procedure admits the possibility that the defendant turns out ineligible for the summons stage. When the defendant successfully pleads a denial by offering facts showing their conduct not to satisfy the elements of the crime alleged, they prove themselves not to be rightly

interrogated over their conduct, that is, to have been strictly speaking ineligible for a court summons at all. Someone may therefore be 'ineligible' on clause (ii)'s second reading whilst being intelligibly summoned to court, in the weak sense of 'intelligibility' (see Section 1.1). By analogy, the same point holds in ordinary moral practices of responsibility-attribution. It is because the expressions are not intersubstitutible in all contexts that it is strictly speaking inaccurate to claim a mere verbal difference. Nevertheless it is clear that on its second reading, 'eligibility for' 'Why?'-questioning is in all contexts intersubstitutible with the 'intelligibility of' it, in the strong sense of INTELLIGIBILITY2. Since INTELLIGIBILITY3 does no more than unpack INTELLIGIBILITY2, it must therefore be, despite appearances, that ELIGIBILITY3 and INTELLIGIBILITY3 in fact state one and the same condition.

Before concluding, let me deflate a potential confusion surrounding ELIGIBILITY3 that may appear to undermine its validity. ELIGIBILITY3 says that someone is morally responsible for all and only those things such that that person's deservingness of moral responses can be determined by 'Why?'-questioning about those things. This carries the hint of a suggestion about the grounds of the appropriateness-conditions of moral responses in general, namely, that they are always based in an assessment of motivating reasons, the reasons *for* which the person did the thing in question. First and foremost, it is important to acknowledge that this suggestion is in fact false. Once the falsity of the suggestion has been rightly acknowledged, however, the appearance that it is entailed by ELIGIBILITY3 will in turn cast doubt on the condition itself. So it is important also to see that this second step is fallacious, since no such entailment actually holds. The potential confusion lies in the failure to recognise either or both of these points: firstly that the suggestion is false, and secondly that it is in the end merely a suggestion.

The suggestion is false because it quantifies over *all* kinds of moral response. This is easiest to see in cases where S has an undefeated φ-reason, but the circumstances conspire to make it impossible for S to φ without breaching a duty owed to another, who is thereby wronged by S's φ even though S φs for the undefeated φ-reason. In cases like this, the reason for which S φs is impeccable, and so S is not appropriately blamed for φ. Nevertheless it might be thought that S's φ is an appropriate object of other natural inhabitants of the 'moral response' category with

a similar (in this case negative) valence.[12] Agent-regret in its dyadic form is commonly considered an appropriate moral response under the circumstances described.[13] The appropriateness-conditions of agent-regret are grounded in those of S's φ-reasons that are left hanging without conformity, such as, in our example, the φ-reason constituted by the fact of S's duty to another. But though the appropriateness-conditions of agent-regret are grounded in S's reasons, this does not mean S's *motivating* reasons. So the counterexample of agent-regret therefore shows up the suggestion as false. Here, as before, I think that any allure the suggestion may have is symptomatic of a wider tendency to see the study of moral responsibility through the lens of blame- and praiseworthiness.

It is characteristic of agent-regret cases that the appropriate moral response to S's negatively valenced φ comes apart from the quality of whatever φ-reasons motivate S to φ. Note however that the inverse case, where S's φ is positively valenced, is also possible. Suppose someone has the following normative reasons: an undefeated ψ-reason, and a supererogatory φ-reason with respect of another. Now imagine this person acts for their ψ-reason, i.e., they are motivated to act merely justifiably, but accidentally ends up doing the supererogatory φ. Though this person is hardly praiseworthy, it is plausible that nevertheless the benefitting of another, φ, is rightly credited to the benefactor's moral life, and that the beneficiary may be right to feel something akin to gratitude. Corresponding to these are, respectively, the monadic and dyadic forms of whatever kind of moral response has appropriateness-conditions grounded in satisfaction of these positively valenced φ and φ-reasons (irrespective of motivation by those reasons). It strikes me that many of the benefits people receive and bring to others in their roles as con-

---

[12] Similar points apply to cases where S breaches a duty owed to another, acting for a defeated φ-reason, but where S has an adequate excuse for φ, such that S is not blameworthy for φ.

[13] For example, Williams quite clearly says that "We should not entirely assimilate agent-regret and the wish, all things taken together, to have acted otherwise" (1976, 127). Of the dyadic form of agent-regret, i.e., as expressed in reparations, apology and the like, Williams says "it would be a kind of insanity never to experience sentiments of this kind towards anyone" (*op. cit.*, 125). The purely monadic form is less relevant here because I am considering cases involving breach of directed duty (which is blameworthy unless adequately justified or excused), whereas monadic agent-regret presupposes no such wrongdoing. An example of purely monadic agent-regret would be what Williams's Gauguin is supposed to lack with respect specifically to his choice of life, *if*, that is, his life achievements turn out to justify it (*op. cit.*, 134). See Section 3.2 for much more on these issues.

sumers and producers in market economies are like this.[14] Under what moral-emotional rubric we should sort these positively valenced cases is a tricky question. To my mind the relevant moral responses are variants on pride and expectance of gratitude, which are, respectively, the monadic and dyadic forms of whatever becomes *vainglory* in excess. Words like these are, I think, in the right ballpark, though arguably the concepts expressed by 'pride' and 'gratitude' on their most natural readings have motivation-sensitive application-conditions.

Now that the falsity of ELIGIBILITY$_3$'s suggestion has been spelled out and elaborated on, it has to be shown why this does not undermine ELIGIBILITY$_3$'s validity. That the false suggestion is not a logical entailment is relatively easy to demonstrate. It should be clear that from the true premises that all moral responsibility depends on eligibility for 'Why?'-questioning, and that all answers to 'Why?'-questioning (viz. motivating reasons) ground the appropriateness-conditions of *some* kinds of moral response, logically speaking nothing immediately follows about what grounds the appropriateness-conditions of *all* kinds of moral response. *A fortiori* it cannot follow that the appropriateness-conditions of moral responses are always grounded in motivating reasons elicited by 'Why?'-questioning. So whilst all moral responsibility depends on eligibility for the very 'Why?'-questioning answers to which are analytically the grounds of hypological judgment, not all kinds of moral response share those same grounds. ELIGIBILITY$_3$ is not undermined by the correct observation that agent-regret and its inverse are counterexamples to what it seems to suggest, because the suggestion only follows by fallacious reasoning.

1.3. Conclusion

In this Chapter I have offered an interpretation, or perhaps a reconstruction, of Smith's view of moral responsibility as answerability. If what I have said is correct, then what she presents as two separate conditions of moral responsibility are in fact just two appearances of a single con-

---

[14] This is not necessarily a criticism of such an arrangement of social life. Of course, the rampant capitalist who sincerely claims credit for the lifting of all boats effected by the rising tide of their self-interest is a deluded figure. But the force of their threats to take jobs and taxes elsewhere shows that, whilst they are pretty deluded, they are not totally deluded. What is delusional is their hyperinflated sense of their own beneficence, and the corresponding measure of gratitude they expect in the form of (say) derisory corporate tax rates.

dition, and a disarmingly permissive one at that, which principally requires that the thing for which someone's moral responsibility is in question be a thing with respect to which that someone possesses normative reasons for or against. It is important to highlight the qualification 'principally' because of the *ad hoc* amendment made early on in the Chapter to clause (i) of the condition. This amendment, which concerns disability, will be considered in more detail in Section 3.1. Putting this on the backburner, the main takeaway from this Chapter is that Smith's view makes moral responsibility principally a matter of reasons-possession of the specifically *normative* variety. This clarification will hopefully be of interest to Smith's adherents and detractors alike. For adherents like myself, however, the proposed clarification brings with it an obvious challenge. In order for the view to be of any assistance to our practices of responsibility-attribution, it becomes crucial to have a workable account of the conditions of reasons-possession. Developing such an account will be the primary task of the next Chapter.

The Conditions of Reasons-Possession

2. Introduction

In the last Chapter, I gave an interpretation (or perhaps a reconstruction) of Angela Smith's view of moral responsibility as answerability. The point laboured there was that the apparently separate conditions she presents as 'intelligibility' and 'eligibility' in fact amount to a single condition, expressing the core insight that moral responsibility is principally a matter of possession of normative reasons bearing on the thing at hand. Since this condition will be a frequent reference point from here on in, it will be helpful to give it, and the view encapsulated by it, a more perspicuous name purified of the distorting language of 'answerability.' I shall call the Smith-inspired view the *Reasons View* and rename its guiding condition accordingly:[15]

REASONS CONDITION: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S has a φ-reason.

In strictly analytical perspective the REASONS CONDITION should be a welcome result. It reduces moral responsibility to something simpler, in terms of which an answer can be given to the question: '*What is it* for someone to be morally responsible for something?' From a richer theoretical perspective, however, the same result leaves much to be desired. It is generally, and I think rightly, considered a constraint on a satisfactory theory of concepts in practical philoso-

---

[15] Not to be confused with the 'Reason View' proposed by Susan Wolf (see, e.g. 1993, 94). Since the Smithian view I am developing foregrounds the mere possession of (rather than motivation by) normative reasons, and indeed of bare (rather than 'right,' i.e. undefeated) normative reasons, it has, I believe, at least as strong a claim to the moniker.

phy that the theory in question is not silent on the application-conditions of the relevant concept in specific contexts. Failure on this front will be familiar from critiques of consequentialism to the effect that, though it may succeed in analysing morally right action in terms of maximising good consequences, the proposed analysans is too opaque to help us determine whether specific actions are right or wrong. Similarly, whether adopting the Reasons View renders us any more confident in our attributions of moral responsibility in specific contexts will depend on whether we have any better a grasp of when people possess normative reasons than when they are morally responsible *simpliciter*. As it happens, I think we do, but even if that is so, it will not be by virtue of anything the REASONS CONDITION says. In theoretical perspective, then, the REASONS CONDITION would be more successful if it were supplemented by a workable account of the conditions of reasons-possession. Developing such an account is the task of this Chapter.

Before proceeding, let me briefly spell out the argumentative strategy deployed throughout the Chapter. My strategy for inquiring into the conditions of reasons-possession owes its possibility to a philosophical commonplace recently given forthright expression by Bart Streumer. He says,

"Most philosophers think that 'ought' implies 'can' […]. Many philosophers also think that there is a close connection between what we ought to do and what there is reason for us to do. This suggests that 'reason' may also imply 'can' […]." (2018, 233)

Call the philosophical commonplace Streumer draws attention to the *Reasons Implies Can* principle, or RIC for short.[16] At a first pass, the principle might be given the following formulation.

RIC*: S has a φ-reason only if S has the ability to φ (or not-φ as the case may be).

There are a couple of immediate clarifications to make concerning the 'reasons' and the 'implies' parts of this principle. As we saw in Chapter 1, the φ-reasons in the antecedent of RIC* are

---

[16] See John Gardner (2013) for another recent employment of RIC.

not motivating but normative reasons, and indeed normative reasons in the widest sense, i.e., whether defeated or not (and if undefeated, whether uniquely or not).[17] They are, as it were, *bare* reasons. As for the logical connective, it seems fair to assume that RIC* aspires to state a principle about the conceptual relations between reasons and abilities. Without great ceremony, then, I simply stipulate that the 'implies' in RIC* is the 'implies' of conceptual entailment.[18]

The antecedent of RIC* proudly displays its intention to collaborate with the REASONS CONDITION so as to secure the Reasons View's success in theoretical perspective. Moral responsibility is all about the φ-reasons, and the φ-reasons are determined by looking to people's abilities. With this in place, it is important to unpack the parentheses that figure in the consequent. They are there because φ-reasons include not only normative reasons for φ but also against φ. When considering cases where S has normative reasons against φ, the relevant 'can' is the ability to not-φ. In other words, RIC* in fact combines what we could loosely call a positive and a negative condition. Since a normative reason against φ is a normative reason for not-φ, and not-φ (i.e. *omissions*) are being treated as themselves amongst the φ, we can avoid needless verbal complexity by reformulating RIC* in the positive way, in terms of normative reasons for rather than against.[19] We shall therefore use the following formulation instead.

RIC: S has a normative reason for φ only if S has the ability to φ.

With this in place, my strategy for inquiring into the conditions of reasons-possession can be spelled out. It will not have escaped notice that, in presenting RIC, I only fixed the senses of the 'reasons' and 'implies' parts. That the sense of 'can' in the consequent is left indeterminate is no accident, since finding the strongest sense of 'can' that sustains RIC's validity *is* the strategy de-

---

[17] Unique undefeated reasons are usually called *conclusive* reasons, and whatever they support things for which there is 'most' reason.

[18] As the Streumer quote suggests, philosophers' commitment to RIC has been less than explicit. There is therefore not a lot written about ways of interpreting the connective in RIC, or what exactly hangs on the decision. For a good resource on the decision regarding the more familiar *Ought Implies Can* principle, see Peter Vranas (2007, 197-8n1).

[19] Remembering that 'reasons for φ' cover both reasons for and against, and, of course, that the choice is arbitrary (bar considerations of expository convenience). I see no reason to think that the formulation in terms of reasons against could not subserve my strategy, nor that it would lead to interestingly different results. For the decision to include omissions amongst the things, see footnote 3. Here, as before, it should be clear from context whether or not the φ in question is an omission.

ployed in this Chapter. My strategy therefore revolves around varying the sense of 'can' in RIC's consequent. By successively substituting different senses of 'can,' different varieties of ability, it can then be asked whether lacking the relevant ability, whether not-can, intuitively supports the inference to not-reasons, i.e., the absence of reasons-possession. In other words, RIC allows us to get at the conditions of reasons-possession negatively, by contrapositive inference. In what follows, the procedure of varying the sense of 'can' and checking the results of contraposition against intuition will be referred to as *RIC-reasoning*. My strategy is to use RIC-reasoning to arrive at the most exclusive sense in which 'S has the ability to *φ*,' so as to supply the maximally non-trivial conditions of reasons-possession. By doing this, I hope to show the REASONS CONDI-TION fruitful for our moral practices of responsibility-attribution, and thus to show the Reasons View successful from the richer theoretical perspective.[20]

After briefly submitting the weakest sense of 'can' for RIC-reasoning, namely, that of *meta-physical* possibility, the rest of Section 2.1 is spent considering the stronger sense of *physical* possibility. The mainline argument of Section 2.1 develops in three stages. The first stage is geared towards motivating the heuristic assumption of determinism, which in the wider context of my strategy and its aims, translates to a compatibilist form of the Reasons View. In the second stage, I argue that the best compatibilist Reasons View takes a *negative-asymmetrical* form, and that its only workable version does *not* interpret the RIC-relevant sense of 'can' physically. In the third stage I drop the deterministic assumption, and argue that the best incompatibilist Reasons View *does* interpret the RIC-relevant sense of 'can' physically. My central contention in Section 2.1 is that, whether compatibilist or incompatibilist, the best Reasons View appeals to a special sense of 'can' in RIC-reasoning. Does this special sense of 'can' express a physical modality? Well, it depends. And perhaps tendentiously, it does not matter. Section 2.1 shows that the special sense of 'can' can be explored without coming down on this question, answers to which are determined more than anything by extrinsic commitments.

---

[20] Since my focus is on the conditions of reasons-possession, the arguments in this Chapter should be of interest to all theorists of moral responsibility who appeal to normative reasons. Though I mostly suppress these links in the body of the Chapter, this would include *reasons-responsive* compatibilists such as Wolf (1993), R. Jay Wallace (1994), John Martin Fischer and Mark Ravizza (1998) and Dana Nelkin (2011). Though unlike the Reasons View, reasons-responsive views often foreground (at least putative) normative reasons in their motivating role, the reasons are still normative, and so reasons-responsive views have a stake in their possession-conditions.

Section 2.2 explores the special sense of 'can' by taking seriously philosophers' reference to powers, or specifically, *general rational powers*, in similar contexts. As a preliminary, I fix my use of dispositional language in accordance with a minimally committing framework for thinking and talking about powers. The bulk of Section 2.2 explains how the powers-interpretation of the special sense of 'can' satisfies a pair of constraints encountered in Section 2.1, viz., that it has non-trivial utility for RIC-reasoning and that it respects the 'error condition.' The specific powers-reading that satisfies these twofold constraints, I argue, invokes *two-way* powers in an *actualist* manner. What this means is that according to the best Reasons View it is a condition of normative reasons-possession, and thus moral responsibility, that rational power manifests itself (privatively or otherwise). This is the major upshot of Section 2.2, and indeed Chapter 2 as a whole.

2.1. The 'Can' of (Meta)physical Possibility

A natural starting point for getting at the most exclusive sense of 'S can φ' (i.e. 'S has the ability to φ') is with the weakest and most inclusive sense of 'can': metaphysical possibility. This interpretation of 'can' makes the not-can relevant to RIC-reasoning that of metaphysical impossibility. If negating RIC's consequent reads 'It is metaphysically impossible for S to φ,' does the antecedent 'S has a normative reason for φ' come out false? Intuitively it does. Nobody has a reason, say, to draw a square circle, or to give birth to one's own father, or to approach that shadow over there and give it a squeeze, and so on. RIC retains its validity when the sense of 'can' is fixed as metaphysical possibility. The problem is that this validity comes cheap, in this case discounted at a rate proportionate to the weakness of 'can.' Though RIC comes out valid, it is only trivially so. To see this, consider how easily we can think up values for φ where it is metaphysically possible for S to φ, but where S intuitively lacks a reason to φ. For instance, though it may be metaphysically possible for Nina to jump to the moon, it is obviously false that she has a reason to jump to the moon—quite irrespective of whether there would be any value in doing so, could she in fact do so. If we want an interpretation of RIC non-trivial enough to explain why we lack reasons to jump to the moon, we will have to do better than the 'can' of metaphysical possibility.

Another candidate in the running is the 'can' of physical possibility. There is an awful lot more to say about reading the RIC-relevant 'can' physically, and the rest of this Section is dedicated to what the Reasons View should say about doing so. Physical possibility is standardly understood to be stronger than metaphysical possibility in that claims of physical possibility entail corresponding claims of metaphysical possibility, whereas the reverse does not hold. This is because physical possibility is a species of nomological possibility, and in general, a claim of nomological possibility is true if and only if there is a metaphysically possible world in which the state-of-affairs obtains *and* the relevant laws of nature obtain there. To say a state-of-affairs is physically possible is therefore to say that that state-of-affairs is metaphysically compossible with our actual physical laws. In turn, to say a state-of-affairs is physically impossible is to say that there is no metaphysically possible world containing that state-of-affairs, where our actual physical laws obtain.[21] Since we are after a less trivial interpretation of RIC than metaphysical possibility, we should consider states-of-affairs whose physical impossibility is due, not to their metaphysical impossibility *per se*, but to that of the obtaining of our actual physical laws in the possible worlds that contain those states-of-affairs.

When we think of our actual physical laws as nondeterministic, it is often difficult to say when a state-of-affairs is physically impossible, and correspondingly difficult to say whether or not instances of 'It is physically impossible for S to φ' are true. This is easy enough when values of φ straightforwardly contravene our actual physical laws, that is, when no calculations are needed to determine the contravention. Nobody has a reason to cause something to move faster than the speed of light (say), however metaphysically possible this may be. Call cases like these *transparent* cases. Conversely, for *opaque* cases lacking this character, it is difficult to know what to say. For example, if the only conceivable way for A to rescue B is to divert the oncoming train by pulling a lever, but the lever is actually jammed, and B actually doomed, is it physically possible for A to rescue B or not? If the 'can' of physical possibility is to be any less trivial than that of metaphysical possibility, then for a target state-of-affairs like 'A rescues B' our ac-

---

[21] The relationship between metaphysical and nomological modality is contested. A lot turns on the question of whether conceivability suffices for claims of metaphysical possibility. Needless to say I am not in a position to contribute to these difficult discussions. My understanding of physical modal claims is lifted more or less wholesale from the helpful introduction to the topic in Tamar Szabó Gendler and John Hawthorne (2002), which I regard as safe to call 'standard' in contemporary Anglophone philosophy.

tual physical laws must yield a list of n metaphysically possible worlds ($n \geq 0$) that both satisfy the laws and that contain 'A rescues B.' Under ideal conditions this could be done by running a complete physical description of the history of the actual world through the laws. This makes the question whether A has a reason to rescue B turn on empirical calculations, presumably very complex, and under non-ideal conditions likely arbitrarily circumscribed in principle.

Things get easier when we think of our actual physical laws as deterministic. Straightforward cases remain as transparent as under nondeterminism. Where values of φ explicitly contravene our actual physical laws, physical impossibility can be concluded without further ado. But even the less straightforward cases lose their opacity under determinism, because the deterministic assumption effectively obviates the need for complex empirical calculations. For example, returning to A and B, where B is actually doomed regardless of A's conduct, the determinist need not entertain calculations because the answer is already built into the case. Determinism says our actual physical laws are such that for any possible worlds sharing a complete physical description D_t at time t, if the laws obtain, then for each of the worlds at every later time $t+i$ we can calculate exactly one complete physical description D_{t+i}.[22] So if the lever is actually jammed and B actually doomed, then, assuming a deterministic physics, this state-of-affairs figures in the sole complete physical description D_{t+i} calculable at any possible world whose physical history is completely described by D_t. Once again, to avoid triviality, our actual physical laws must yield a list of n metaphysically possible worlds ($n \geq 0$) that both satisfy the laws and that contain the target state-of-affairs 'A rescues B.' In this case the list is maximally short ($n = 0$), so that 'It is physically impossible for A to rescue B' is true, and so 'A has a reason to rescue B' comes out false. But if the lever were not jammed, so that in fact A rescues B, then by the same logic, no metaphysically possible world with the same physical history as our own would satisfy the laws whilst containing the state-of-affairs 'A fails B.' Either way, no calculations are required. What for nondeterminism are opaque cases are transparent for determinism. Unlike the nondeterminist, the determinist is able to make judgments of physical impossibility by simply reading off the actual.

Under nondeterminism, many concrete cases are modal-epistemologically opaque to RIC-reasoning, and so nondeterministic Reasons Views will be *prima facie* unsuccessful from the

---

[22] This is a standard definition of determinism Gordon Belot (1995) calls the Earman-Lewis view.

richer theoretical perspective. Under determinism it is much easier to make judgments of physical impossibility, so these troubles are avoided. In the context of my strategy and its aims, then, it makes sense to work from a heuristic assumption of determinism. It pays to stress its heuristic rather than properly philosophical nature. Later on, after its heuristic value has been mined, the assumption will be dropped. Whilst it is in force, however, it is important to see that the assumption brings a further commitment alongside it. Reflection on the thesis of determinism quickly raises the question of its compatibility with moral responsibility, whether in concept or in practice. Either they are compatible, or they are not. There are traditionally thought to be two major incompatibilist options, broadly speaking: *skepticism* and *libertarianism*. Skeptics and libertarians agree about the incompatibility of determinism and moral responsibility. But whilst skeptics argue from the truth of determinism to the unreality of moral responsibility, libertarians cleave to the reality of moral responsibility, arguing on that basis that determinism is false. The deterministic assumption of course rules out the libertarian option. But it is important to see that, equally, the theoretical perspective we are trying to assuage also rules out the skeptical option. If the theoretical perspective is such as to be assuaged at all, it is presumably left cold by a global skepticism about moral responsibility (in concept or in practice). So with both incompatibilist options ruled out, it appears our deterministic assumption comes with compatibilism built in.

Compatibilism is a view about the relations between determinism and moral responsibility. It is not a view about the interpretation of 'can' in RIC. There is therefore no reason to think a compatibilist, in virtue of their compatibilism, should interpret RIC physically. It is no surprise, then, that there are two forms of compatibilist Reasons View, depending on whether or not the 'can' in RIC is interpreted physically. *Prima facie* it is not obvious how it could be open to the compatibilist to interpret the 'can' in RIC physically. Surely the skeptic is right that if determinism obtains, so that it is physically impossible for states-of-affairs to be otherwise than they actually are, then moral responsibility must be either fictitious or illusory? Recall, however, that on the Reasons View moral responsibility turns principally on reasons-possession. Sure enough, if 'can' means physical possibility, then where it is physically impossible for S to φ, S has no reason and so no moral responsibility for φ. But where S in fact φs, it is physically possible for S to φ. Indeed by determinism this is not just physically contingent but necessary, i.e., it is physi-

cally impossible for S to not-φ. And if it is true that 'S can φ,' in the sense of φ's being physically possible, then by RIC-reasoning 'S has a reason to φ' will not come out false. In that case, if S intuitively has a reason to φ, then S is morally responsible for φ according to the Reasons View —determinism notwithstanding. Of course, it is only by sheer luck that S conforms to, and thus possesses, φ-reasons. But if 'can' means physical possibility, then this unusual form of compatibilism is what the deterministic assumption implies.

Coherent as this compatibilist Reasons View may be, it has some unwelcome implications. It can be called *positive-asymmetrical compatibilism* on account of the fundamental asymmetry it implies about what people can be morally responsible for. All those things that do not actually come about become physically impossible things, and thus things we have no reasons bearing upon. Holding fast to the validity of RIC, this implies, as a conceptual matter, that nobody can act contrary to reasons they possess. Equally, by the REASONS CONDITION, nobody is morally responsible for what we would naturally call their transgressions. In turn, nobody can be blameworthy, since *a fortiori* no one ever acts contrary to their unique undefeated reasons.[23] On the flipside are all those things that are determined to actually come about. By dint of their physical possibility, these are things we can have reasons bearing upon, and by dint of physical necessity we actually conform to those reasons (when we intuitively possess them). This asymmetry is already starkly at odds with our existing practices. Worse still, though we come out morally responsible for these conformities with reasons, we never come out praiseworthy for them. Praiseworthiness requires not only conformity with an undefeated reason, but also conformity to that reason *for* a further undefeated reason.[24] Since under determinism it is only ever by sheer luck that S φs as their φ-reasons prescribe, S never φs for φ-reasons (even when S possesses them), so *a fortiori* S never comes out praiseworthy. Not only does this lump the Reasons View with an unfamiliar asymmetry, it makes our practices of praise and blame groundless. In

---

[23] If the condition of blameworthiness implied here is considered too controversial, then blameworthiness can still be ruled out *a fortiori* by the absence of moral responsibility for deviations from normative reasons possessed.

[24] Sometimes, that is, though many things are justifiable, such that doing any of them would be justified, there are further undefeated reasons that speak in favour of some or other of the justifiable things, and being motivated by these further reasons is what praiseworthiness is all about. In Joseph Raz's terms these further reasons are *second-order* reasons (1979, 17). Here I follow Gardner, who argues persuasively that unless we observe this distinction "one wrongly eliminates the logical gap between justified and creditable [i.e. praiseworthy] action" (2007, 103-4).

spite of its coherence, positive-asymmetrical compatibilism is doubly revisionary and should be rejected as implausible.

This leaves us with arguably the more natural way of proceeding with the strategy in a compatibilist vein. On this route, the compatibilist Reasons View accepts that the truth of determinism makes non-actual states-of-affairs physically impossible, but denies that the 'can' of physical possibility figures in a valid and non-trivial reading of RIC. For reasons that will emerge, call this view *negative-asymmetrical compatibilism*. According to negative-asymmetrical compatibilism, then, the physical impossibility of S's φ does not rule out S's having a reason for φ.[25] Suppose that determinism holds, S does not φ, but S intuitively does have a reason to φ. Then, holding fast to RIC's validity, S has the ability to φ in whatever non-trivial sense is required to sustain RIC's validity. Since we are describing a scenario where S has the ability to φ despite S's φ being physically impossible, it must be that the question of whether S has the ability to φ, in the negative-asymmetrical compatibilist's sense of 'ability,' is totally indifferent to conceivable physical impediments to S's φ.

This observation has a flipside that raises a *prima facie* conceptual problem about the sense of 'can' in question. To see this, consider the following case: again under determinism, S does in fact φ, and intuitively has a reason to φ. By our commitment to RIC, S has the ability to φ in the very same sense, namely, that for which the question of its ascription is totally indifferent to conceivable physical impediments. But this time, S has the ability to φ, in this sense of 'ability,' despite S's φ being physically necessary—i.e. despite S's not-φ being physically impossible. So the indifference of 'can' to conceivable physical impediments in the two cases conceals an important difference in the physical realities underlying them. Physically speaking, whereas in the former case S is necessarily in error relative to S's φ-reason, in the latter case error is impossible. The conceptual problem lies in this. Many philosophers agree that it belongs to the concept of

---

[25] Negative-asymmetrical compatibilism is not available if one takes a necessitarian view of the laws of nature, as does, perhaps most famously, Sydney Shoemaker (1980). For necessitarians, modal claims underwritten by our actual physical laws will be metaphysically necessary truths, so that, in effect, the impossibility of non-actual states-of-affairs under determinism is strengthened to metaphysical impossibility. Since RIC with the 'can' interpreted metaphysically comes out trivially true, the physical (= metaphysical) impossibility of S's φ will, for the necessitarian, rule out S's having a reason for φ. At this stage, the necessitarian can either adopt a variant of the positive-asymmetrical compatibilism we just rejected as implausible, or fall back to an incompatibilist acceptance of a nondeterminism whose modal-epistemological resources render many concrete cases opaque to RIC-reasoning.

a normative reason that those possessing them can be in error relative to them. Call this the *error condition* (EC).[26] But in the relevant sense of 'can,' someone might, consistent with determinism, conform to a reason to φ, despite it being physically impossible not to have so conformed. In other words, *prima facie*, negative-asymmetrical compatibilism tacitly understands a subset of normative reasons (viz. those actually conformed to) in a manner that violates EC. This is where negative-asymmetrical compatibilism gets its name. Removing the EC-violating reasons results in an inverse asymmetry—a meaner one at that—in which people only ever come out transgressive with respect to reasons (perhaps even blameworthy), and never in conformity (let alone praiseworthy).

The way to iron this out is to bring the 'errors' whose possibility is required by EC into the modal space of the relevant sense of 'can.' In other words, ascriptions of ability in this sense must not only be totally indifferent to conceivable physical impediments, they must moreover actively countenance impediments of a *non-physical* variety. Abilities of S's that could be gotten in the way of on non-physical grounds are, it can be agreed, presumably somewhat spooky. Consider the palpable spookiness of empirical possibilities whose actualisation is in principle susceptible to interference by the non-physical. But absent such abilities, RIC will be invalid, and in turn RIC-reasoning will be strategically useless to negative-asymmetrical compatibilism. This is because its failure to meet EC will mean there are no normative reasons to seek the conditions of possession for, at least where they are intuitively—now illusorily—conformed to. Call the EC-respecting 'can' and the corresponding sense of 'ability' in 'S has the ability to φ' the *special senses* of 'can' and 'ability.' The problem for negative-asymmetrical compatibilism is that for it to be workable it must countenance special abilities that are ontologically spooky.

At this point it seems that the compatibilist way of proceeding with the strategy has reached an impasse insofar as the debate with incompatibilism is concerned. In the context of the historical development of the debate, the dialectical point of compatibilism for non-skeptics about moral responsibility was always that it represented allegiance to the scientific worldview, whose

---

[26] This terminology is owed to Doug Lavin (2004). In his paper, Lavin evidences widespread assent to EC, and gives it one of the only sustained explorations in the literature to date. I take it that the positive-asymmetrical compatibilism we rejected falls foul of what Lavin calls the 'logical interpretation' of EC, since given its background commitments (viz. to determinism and the physical 'can'), all normative reasons people could conceivably possess are ones whose prescriptions turn out to be inviolable.

best physical theory to date has far from ruled out fundamental determinism.[27] Historically, the opposing non-skeptical (i.e. libertarian) incompatibilism has been the option easily dismissible as scientifically dubious, ontologically profligate, and in the end probably of quasi-theological motivation. But if the preceding paragraphs are anything to go by, compatibilist Reasons Views cannot maintain the dialectical advantage, since even the best (viz. negative-asymmetrical) compatibilist Reasons Views countenance spooky abilities susceptible to non-physical impediment in principle. The obvious follow-up is why the scientifically-minded are supposed to find this any less suspicious than the libertarian alternative. This ought to give pause to any compatibilism that makes a basic appeal to normative reasons, which includes, for example, contemporary *reasons-responsive* compatibilists. Unless the conditions of reasons-possession can be established by a strategy other than RIC-reasoning, they will need some way of addressing this challenge to their comparative sobriety. To my knowledge there is little to no reasons-responsive literature that gives airtime to EC, so if there are plausible ways of retaining the dialectical advantage, as there may well be, then I am not aware of them.

Let us briefly take stock of the last few moves. On a compatibilist Reasons View, the interpretation of 'can,' if valid and non-trivial, may mean either physical or non-physical possibility. If physically interpreted, the resultant view implies *inter alia* an implausible positive asymmetry in our reasons-possession (thus moral responsibility), and should be rejected accordingly. If non-physically interpreted, a different though equally implausible negative asymmetry awaits, which can only be avoided by countenancing special abilities that are, at least on the non-physical interpretation, ontologically spooky. The takeaway message is that for the best compatibilist Reasons View, the 'can' in RIC assumes a special EC-respecting sense, expressive of a non-physical modality.

Naturally, from this new dialectical vantage non-skeptical incompatibilism can begin to look attractive. Taking this seriously requires that we drop the deterministic assumption. As we have

---

[27] The rise of quantum mechanics is sometimes thought to falsify this claim. Since the inception of quantum mechanics, however, its proper interpretation has been much debated by the scientific community, with very little consensus to show. If we let pragmatic considerations creep into what counts as the 'best' physical theory, this lack of consensus tends to favour fundamental determinism. But even if we reject such pragmatism as alien to the objectivity of scientific theory (*pace* most working physicists), it is still respectable to adopt the de Broglie-Bohm theory of quantum mechanics, which has deterministic dynamics in the classical sense. For a recent and accessible account of these matters, see Jean Bricmont (2016).

seen, however, under nondeterminism the trouble is that many concrete cases become opaque to RIC-reasoning. This is because it is comparatively difficult to say what is physically impossible under nondeterminism. But thanks to the useful heuristic of determinism, we now know of one way forward, namely, whichever way negative-asymmetrical compatibilism was going to proceed before getting mired in the present debate—an exercise deferred to Section 2.2 onwards. Transposed to a nondeterministic assumption, there is no reason to think the relevant special abilities—those whose ascriptions must allow for EC-respecting impediments—must allow for *non-physical* impediments. That is, nothing prevents impediments from occupying the modal space freed up by the nondeterminism of our actual physical laws, so that the 'possibility' of error required to meet EC may just be good old physical possibility. Quite how this is supposed to go is of course pretty obscure, but the point is simply that there is nothing scientifically unrespectable about it in principle.[28] Nothing spooky need be going on with special abilities under nondeterminism. From the new vantage, it is the best incompatibilist Reasons View that has the dialectical upper hand, and its compatibilist counterpart that looks suspiciously theological.

As with any overtly dialectical position, the best incompatibilist Reasons View is not without its flaws. The heuristic assumption of determinism led us to see how the RIC-relevant sense of 'can,' and so of S's ability to φ, might be developed despite S's φ being physically impossible—namely, via appeal to special abilities. This lesson remains salient even if nondeterminism turns out to be true, since the thesis (or theses) of nondeterminism certainly need not imply that nothing is physically impossible. Imagine, once again, that our actual physical laws are known to be nondeterministic, and that we have some less straightforward cases to decide. Surely we should still deploy RIC-reasoning by plugging in the special sense of 'can,' i.e., the one whose ascriptions allow for EC-respecting impediments? Faced with the prospect of a later, more scientifically sophisticated time, in which it is proven physically (im)possible to have done otherwise, the best incompatibilist Reasons View has basically two options. One option is to make peace with the worry that present pieces of RIC-reasoning (with the special 'can'), and thus present responsibility-attributions according to the REASONS CONDITION, may turn out false. This

---

[28] Couching this obscurity in the vocabulary of the debate in its classical 'free will' guise, Jenann Ismael says, "Quantum mechanics adds a bit of microscopic randomness to nature, but nobody knows how to transform that randomness into the sort of control that we think that we have over our own actions" (2016, *x*).

worry is due to the physical (and so in principle scientifically tractable) nature of the relevant abilities, and does not, for all its ills, apply to negative-asymmetrical compatibilism. The alternative is to make peace with the ontological profligacy of placing special abilities into a non-physical realm. Both options cede something of the newly gained dialectical advantage.

In the end, then, whether or not the non-skeptic thinks the Reasons View should interpret the 'can' in RIC physically has everything to do with whether the Reasons View is considered more at home in the compatibilist or incompatibilist camp. In turn, however, answers to this latter question have almost nothing to do with determining the maximally non-trivial reading of RIC, and do not help to vindicate the Reasons View in the eyes of the richer theoretical perspective. They express, more than anything, the force of extrinsic commitments. Beyond what level of opacity do concrete cases evidence the unworkability of a sense of 'can'? How concerned should we be about the charge of anachronism in the face of science? What *is* the current science? Is it a problem that a future science could debunk our present moral practices? Our heuristic determinism helped expose the strange emptiness of coming down on these issues, positions on which are something like a matter of taste. My point in this Section has not been to persuade anyone out of their tastes. My point is that for the best Reasons Views, whether compatibilist or incompatibilist, the RIC principle must assume the special sense of 'can.' Notwithstanding tastes, this sense of 'can' can and must be explored. In Section 2.2, I begin this exploration in an agnostic silence on this Section's debates. It is possible to make whatever is sayable about the special sense of 'can' amenable to compatibilists and incompatibilists alike—and thus whether it expresses a physical modality or not—just so long as possession of normative reasons figures substantively in the account.

2.2. The 'Can' of Power

Even when we bracket the vexed debate between compatibilists and incompatibilists, the special sense of 'can' remains elusive. This is partly due to the difficulty of getting senses of 'can' to satisfy the twofold constraints encountered in Section 2.1. On the one hand, to get any mileage out of RIC with a candidate sense of 'can,' plugging it into the consequent must at least sometimes support, by negation, the contrapositive inference to the absence of reasons-possession.

39

That is, it must have utility for RIC-reasoning. But on the other hand, at least some conceivable impediments to 'ability' in the very same sense are in fact a condition of reasons-possession, so it cannot be that they all exclude reasons-possession by making 'S has the ability to φ' come out false. That is, a candidate sense of 'can' must not be EC-violating. Though the special sense of 'can' was said to meet these twofold constraints, quite how it manages to do so remains mysterious.

One avenue that suggests itself is to take seriously the talk of *general rational powers* commonplace amongst philosophers writing in contexts similar to our own.[29] Indeed when dealing with ability-ascriptions whose subjects are restricted to agents such as to have normative reasons (see Section 1.1), something like this reading of 'can' is arguably the most natural. Though it is true that, if anything, sounding natural tends to favour an interpretation, an interpretation's natural-soundingness can just as well signal its ambiguity as its correctness. The dispositional language of 'powers' is perhaps especially prone to imprecise use in practical philosophy, as it is easily filled by sympathetic readers with whatever meaning best suits its user at the time. I will therefore endeavour to be as clear as possible about my own use of this language.

In what follows, my 'powers'-talk will accord with George Molnar's theory of powers. Given that by 'powers' Molnar means whatever properties exhibit "the deep characteristics, or marks, of the family of dispositional properties" (2003, 57), one basic clarification that can be made straight away is that my use of 'powers' is interchangeable with 'dispositions' as it often appears in the literature. My 'powers'-talk therefore does not encode a distinction between agential and non-agential dispositions. In this regard, as in many others, Molnar's theory of powers is broadly continuous with other non-reductive realist theories of dispositional properties.[30] However, the fragment of Molnar's theory I will present is also meant to be shallow enough to encompass what it shares with reductivist or anti-realist theories. What I present is in this sense a *powers-framework*, that is, a minimally committing framework for thinking and talking about powers in

---

[29] Fundamental reference to general rational powers is prominent in the reasons-responsive wing of the moral responsibility literature, as well as Hartian criminal jurisprudence. For a clear example, consider Wallace: "the general abilities that matter to responsibility are distinctively rational powers" (1994, 188). In a more recent context closer to our own, Raz says that reasons "are reasons because rational creatures can recognize and respond to them with the use of Reason," by which he means a general rational power (2011, 86).

[30] Such as, e.g., Nancy Cartwright's (1999) theory of 'capacities' and Stephen Mumford's (1998) theory of 'dispositions.'

general. Points of contention will be indicated where appropriate, though, needless to say, it is not my intention to wade too deep into debates in analytical metaphysics.

According to Molnar, powers are properties of objects. Their manifestations are sensitive to the circumstances those objects find themselves in, where *which* circumstances matter is given by the nature of the power in question. S's having a power, P, is therefore always S's having the power to φ given circumstances C. Here, substitution instances of φ and C are always determinables with determinate manifestations and determinate circumstances respectively.[31] To spell this out: S's having a power P is always S's having the power to manifest determinate manifestations ψ_n of the determinable φ selected by a function, F, from determinate circumstances falling under the determinable C (*op. cit.*, 165).[32] The function F always at least tacitly includes a *ceteris paribus* clause, so as to exclude any funny business going on in determinate values of circumstances C.[33] Typically, practical philosophy concerns itself with macroscopic powers of macroscopic objects. Macroscopic objects have parts, i.e. they are complex objects. For Molnar, if S and P are macroscopic, instances of P will be what he calls 'derivative' powers (*op. cit.*, 29).[34] If S's power P is derivative, then its determinate manifestations ψ_n (of the determinable φ) are the joint exercise of several immediately grounding powers, belonging either to S or some of S's suitably related parts (*op. cit.*, 145). For expository convenience, call these immediately grounding powers the grounded power's *underlying capacities*. Piecing all this together, and speaking on the level of determinables, manifestations φ of S's power P—a power to φ given C—are themselves joint manifestations of S's underlying capacities Q, R, T, given their respective circumstances D, E, F (determinates of which constitute determinates of C). To say S has the *opportuni-*

---

[31] For a vivid presentation of the idea that the manifestations of (at least some) dispositional properties are "highly generic or determinable," see Gilbert Ryle (1949, 118). I take it that 'general' as it figures in talk of 'general rational powers' expresses just this genericity, which manifestations of rational powers exhibit to perhaps the greatest known extent.

[32] The idea that the determinate manifestations of powers are outputs of functions whose arguments are the determinate circumstances bearers of those powers find themselves in is reflected in Cartwright's view of the relationship between powers (she prefers 'capacities'), laws and what she calls 'nomological machines.' Her discussion of Coulomb's law evidences this confluence pretty clearly (1999, 59).

[33] Funny business is afoot when S finds itself in determinates of determinable circumstances C containing nonstandard features like finks, maskers, etc. Though Molnar does not say much about the place of *ceteris paribus* clauses in his theory, I have filled the gap just as Cartwright does in her own very similar theory (1999, 188). Given that Molnar is not offering a conditional (i.e. reductivist) analysis of powers, I see no reason not to take this step.

[34] Barring the controversial possibility of emergent properties (*op. cit.*, 143).

ty to φ is to say S has a power P (i.e. to φ given C) and that S finds themselves in a determinate of the determinable circumstance C. Which is to say, the capacities underlying S's P have the opportunity in the same sense, not to φ, but to make their respective *contributions* to φ.

How does this machinery work in concrete cases? Consider, for example, the fragility of a particular glass. In the present use of 'power,' fragility is a type of power. It is a power to break (= φ) given the right circumstances (= C). More specifically, it is a power to break like this, that, or the other (= determinates $\psi_1, \psi_2, \ldots, \psi_n$ of the determinable φ) as selected, *ceteris paribus*, by a function F from determinates of the determinable circumstances C. Determinate circumstances mapped by F to the glass's breaking like this, that, or the other might include, for instance, circumstances where the glass is struck like this, dropped like that, twisted like this, etc. As with all macroscopic powers, the glass's fragility is grounded in what I am calling its underlying 'capacities.' In this case, the underlying capacities are presumably themselves further powers of some of the glass's parts (suitably related), rather than of the glass as a whole—perhaps its atoms, molecules, and their arrangement. In my use of 'opportunity,' when the glass finds itself in a determinate circumstance C it has the opportunity to break, that is, the capacities underlying the glass's fragility have the opportunity to make their respective contributions to the glass's breaking. Note that for a range of determinate circumstances C the function F fails to map them onto a determinate way for the glass to break. If the glass is bubble-wrapped, for instance, then even when dropped it will not break. This sort of funny business is what the *ceteris paribus* clause tacitly qualifying the function F rules out. But of course, *ceteris paribus*, given its fragility, the glass will break (i.e. it manifests its fragility) if and only if it has the opportunity to do so, or, what is the same, its underlying capacities have the opportunity to make their respective contributions to doing so.

Now that I have fixed my use of dispositional language, especially viz. talk of 'power,' 'capacity' and 'opportunity,' let us return to the mainline question of how reading the special sense of 'can' in powers-terms satisfies the two constraints mentioned above. It is fairly straightforward to show how it has at least minimal utility for RIC-reasoning, and therefore meets the first constraint. To see this, suppose for a moment (and not implausibly) that there is a deep conceptual connection between something's being a glass and its having the power of fragility, such

that if something is a glass, then by conceptual entailment that same something is fragile. What is being supposed is the truth of a conditional, call it *Glass Implies Fragility*, or GIF for short.

GIF: S is a glass only if S is fragile.

For any S that lacks the power of fragility, if GIF is true, then that licenses the contrapositive inference to the claim that S is not glass. So for instance, were a glass to lose its fragility by exposure to extremely high temperatures, then according to GIF the resultant mass of liquid could not properly be called 'glass.' As long as GIF is true, then the fact that its consequent ascribes a power to whatever figures in the antecedent's subject-position is enough to reason contrapositively from the thing's lacking the power altogether. Similarly, to the extent that the RIC principle is valid, then on a powers-interpretation of its consequent—where 'S has the ability to φ' reads 'S has the power to φ'—that S lacks the power to φ suffices to infer that S lacks a reason for φ, by RIC-reasoning. Since macroscopic powers are grounded in their underlying capacities, something loses a macroscopic power when that power loses its underlying capacities. In the case of a glass, that some of its parts lose their (grounding) powers in ceasing to be suitably related is what amounts to the glass's fragility losing its underlying capacities. The same holds between whatever macroscopic power interprets the special sense of 'can' and whatever capacities underlie it—more on which in Chapter 3. For present purposes what matters is that, albeit fallibly, it is well within our reach to make or withhold all sorts of power-ascriptions like these. Interpreting 'can' in powers-terms thus guarantees at least minimal utility for RIC-reasoning.

It may be thought that the role of opportunity in power-ascriptions actually allows the powers-interpretation of 'can' to avoid EC-violation, so that it satisfies the second constraint as well. The thought would be as follows. If a glass is fragile, then given the opportunity, it breaks—or what is the same, its underlying capacities contribute to its breaking. Now imagine a glass that possesses the relevant underlying capacities but lacks the opportunity in the given circumstances. This glass is of course fragile by hypothesis. But since it lacks the opportunity, fragile though it may be, it is not manifesting its fragility here and now. This suggests that on a powers-interpretation of 'ability' in 'S has the ability to φ,' lack of opportunity could be a conceivable impediment to ability that is EC-respecting. After all, the glass has the 'ability' to break in the

powers-sense, even whilst its ability is impeded by a lack of opportunity. The thought, in effect, is that power-ascriptions in the consequents of true conditionals like GIF always have another guise signalling the presence of opportunity. Call this the *actualist* guise of the power-ascription that figures in its *potentialist* guise in GIF. For example, under the actualist guise GIF would look like this:

GIF (ACTUALIST): S is a glass only if S manifests its fragility.

Imagine again a fragile glass that in the given circumstances lacks the opportunity to break. It does not manifest its fragility, so the consequent of GIF (ACTUALIST) comes out false. Clearly, however, that the glass does not manifest its fragility does not thereby show that it is not a glass. So whilst contrapositive inference is intuitively supported by GIF, it is unsupported by GIF (ACTUALIST). Since the potentialist and actualist guises are guises of one and the same power-ascription, and so one and the same sense of 'ability,' this sense of 'ability' allows impediments that do not rule out whatever having the ability is a conceptual condition of. In the case of GIF it is a condition of being a glass, whilst in the case of RIC it is a condition of reasons-possession. Attending to the role of opportunity in power-ascriptions leads to a distinction between their guises, and this shows how the powers-interpretation of 'can' avoids EC-violation. Or so the thought concludes.

Though this is an interesting attempt to understand how the powers-interpretation of 'can' avoids EC-violation, it sadly misses the mark. There are two reasons for this. First, the reference to guises of power-ascriptions is not coherent. It perhaps contains a kernel of truth, namely, that power-ascriptions say not only how things actually are with the object in question (i.e. it is powerful), but also how things could potentially be with it (i.e. its power could manifest in various ways). But, ultimately, that something has a power and that it manifests that power are two very different realities. Indeed the whole point of keeping powers in our metaphysical toolkits is that they can be had without being manifested. To say that these different realities are in fact one reality under different guises, or that judgments about these different realities are but different guises of a single power-ascription, is incoherent. The second reason is closely related to the first. If GIF is true, and GIF (ACTUALIST) is simply GIF but with the consequent appearing under a different guise, then it is not clear how contraposition could fail in one case but not the

other. Assuming the logical connective stays the same, there are two options. If the two guises are of the same sense, then there is simply nothing left to explain differences in inferential behaviour. If the two guises are of the same reality (i.e. they are like different senses), then only an intensional context could explain the difference, but there is none.

The correct explanation is that whilst GIF is true, GIF (ACTUALIST) is false. Reference to guises, and the related notion that different conclusions could be reached by the same premises and inference rule, is spurious. If the consequent of GIF (ACTUALIST) figures in a true conditional, it will look less like GIF and more like this:

GIF2: S, a glass, breaks like this, that, or the other only if S manifests its fragility.

What GIF2 says is that for any given glass, its manifesting its fragility is a conceptual condition of its exhibiting a certain range of properties, e.g. breaking like this, that, or the other. As we have seen, powers manifest themselves in determinate ways only when their bearers find themselves in determinates of whatever determinable circumstances are written into the nature of the power in question. The effect of rendering the consequent of GIF2 in an actualist manner is to say that having an opportunity to break is a condition of a fragile glass's breaking. That is, a glass's having the 'ability' to break in the sense of having both the power of fragility (its underlying capacities) *and* the opportunity to break is a condition of a glass's breaking. In other words, GIF2 states a *power-* or *capacity-plus-opportunity* condition of a glass's breaking. Moreover, unlike GIF (ACTUALIST) but like GIF, the power- or capacity-plus-opportunity condition stated by GIF2 is also a true conditional.

Powers figure in the consequents of both GIF and GIF2, and they are both true conditionals. They are two different interpretations of the 'ability' an instance of S must have, as a conceptual matter, if it is to figure truthfully in the antecedent. GIF says it needs the 'ability' in the sense of being fragile, i.e., having the power to break. GIF2 says it needs the 'ability' in the sense not only of having but manifesting its fragility, which is to say, of having both the power and the opportunity to break. These different interpretations bring with them a formal distinction between the conditionals they enter into. This naturally raises the question whether RIC is more like GIF or GIF2 formally speaking. To be clear, the two conditional forms are as follows.

FORM I: S is a K only if S has the power to φ.

FORM II: S, a K, has $\psi_1, \psi_2, \ldots, \psi_n$ only if S manifests the power to φ.

So, is the RIC principle a FORM I or a FORM II conditional? To my mind it is pretty clear that RIC is a FORM II conditional, for two main reasons. For starters, much like the predicate '__has ψ_n' in the antecedent of FORM II, the predicate '__has a reason for φ' in RIC's antecedent picks out the properties of an object, rather than bringing it under a sortal like '__is a K' as in the antecedent of FORM I.[35] Moreover, substitution instances of S in RIC are restricted to those of some sort K, namely, objects such as to have normative reasons (see Section 1.1). This feature of RIC also strongly suggests that it bears FORM II rather than FORM I. Those left unpersuaded by these formal similarities should note that in the context of my wider strategy there are also good theoretical reasons for opting for FORM II. Comparatively speaking, opting for FORM II loses nothing and gains something in terms of utility for RIC-reasoning. This is because where substituting one set of values into FORM I and FORM II yields two true conditionals, the FORM II conditional will be stronger than the FORM I conditional in the following way. Whatever leaves the consequent of the FORM I conditional unsatisfied leaves the consequent of the FORM II conditional unsatisfied, but not vice versa. As mentioned, FORM II conditionals express power- or capacity-plus-opportunity conditions, meaning that objects fail to meet them either when they lack the power or when they have the power but lack the opportunity. Whatever has the power but lacks the opportunity fails to meet the FORM II conditional but satisfies the FORM I conditional. So taking RIC in an actualist manner, that is, as a FORM II conditional, is sensible on both formal and theoretical grounds.

---

[35] There is a discrepancy here as to why the properties picked out by RIC are reasons rather than determinate ways to φ. To anticipate the discussion a little, this can be resolved by noting the constitutive rationality of the RIC-relevant power. Minimally, this means that each of S's determinate ways of φ-ing ($= \psi_n$) stands under a normative reason, so that they each correspond to a normative reason to φ in a determinate way ($=$ a reason to ψ_n). Here I follow Randolph Clarke and Thomas Reed, who say that "A manifestation of a rational power is something done for a reason, something responsive to the normativity of reasons" (2015, 20n36). I have slightly adapted their dictum to reflect the fact that manifestations of rational powers admit of privation. That manifestations *stand under* normative reasons (and so being *in principle* responsive to them) seems to me to be what is common to all manifestations of rational powers, whether privative or non-privative.

Let me briefly recap the preceding points. The task of this Section is to explore the powers-interpretation of 'S has the ability to φ' as it figures in RIC's consequent. The present suggestion is to do this in what I called an *actualist* manner, i.e., by reading 'S has the ability to φ' not as 'S has the power to φ' but as 'S manifests the power to φ.' In other words, RIC should be taken as a conditional of FORM II as opposed to FORM I. To do so is to read RIC as a power- or capacity-plus-opportunity condition on reasons-possession.

Recall the twofold constraints on the special sense of 'can' outlined at the start of this Section, viz. utility for RIC-reasoning and avoiding EC-violation. What we have said about powers represents progress as far as the first constraint goes. There are, as we have seen, not one but two ways for 'S has the ability to φ' read as 'S manifests the power to φ' to come out false: where S lacks the power (the underlying capacities) altogether, and where S has the power but lacks the opportunity for it to manifest. So the actualist powers-interpretation of 'can' has plenty of utility for RIC-reasoning. That lacking either the power (the underlying capacities) or the opportunity to do some thing rules out having normative reasons bearing on that thing follows simply from reading the 'can' in RIC in powers-terms.[36] Disappointingly, however, no progress has been made in terms of the second constraint. Since both ways of falsifying 'S has the ability to φ' read as 'S manifests the power to φ' support contraposition, we have yet to discover any conceivable impediments to 'ability' in this sense that do not rule out reasons-possession. Quite how reading 'can' in powers-terms allows it to be EC-respecting remains as mysterious as ever.

With both lack of power and lack of opportunity ruled out, the only place that remains to look for conceivable impediments in virtue of which the powers-sense of 'can' is EC-respecting is, as it were, *in* S's power (its underlying capacities) itself. This may sound *ad hoc* given what I said, following Molnar, about powers more generally. Recall that if a macroscopic object (or its parts) possesses whatever capacities underlie its power to φ, then *ceteris paribus* given the opportunity it manifests its power to φ, and if it finds itself in circumstances that do not present an opportunity, then *ceteris paribus* it does not manifest its power to φ. On our powers-framework,

---

[36] In Section 2.1 we saw that the best compatibilist Reasons View thinks ability-ascriptions, in the special sense of 'ability,' are totally indifferent to conceivable physical impediments. If, as I ultimately suggest, the 'can' in RIC assumes the actualist powers-sense, then this has the following notable implication. Suppose S has the power to φ but cannot φ (in this sense of 'can'), and that S's φ is known to be physically impossible. Here, S lacks the opportunity to φ but the best compatibilist Reasons View may not infer that S lacks the opportunity to φ because φ is physically impossible (though it may well be).

then, there is no room for anything other than lack of opportunity to get between a macroscopic object's having a power and manifesting it. *A fortiori*, nothing about a macroscopic object's power (its underlying capacities) can do so. Does this not show that, to the extent that a macroscopic object has a power, the power (and so its underlying capacities) simply cannot harbour conceivable impediments within it, let alone EC-respecting ones?

Here I appeal to a quotidian observation often mentioned by moral and legal philosophers in similar contexts. People are rather prone to messing things up even when they possess the power and an opportunity presents itself. Premier league strikers have the power to score, despite being known on occasion to miss open goals. Albeit increasingly rarely, cabinet ministers have the power of oratory, despite being known on occasion to garble the message in set-piece press conferences. In these cases the failure to properly manifest powers of goal-scoring and oratory, respectively, is most plausibly attributed not to a lack of power or opportunity, but rather to an impediment to the proper manifestation of the power (and so the underlying capacities) itself. One way to sustain these appearances is to introduce a distinction between those powers (and underlying capacities) that are susceptible to impediments to proper manifestation even in the presence of opportunity, and those that are not. A premier league striker's power to score goals and a cabinet minister's power of oratory belong in the first group, whereas a glass's fragility belongs in the second. Not uncontroversially, but for reasons that will emerge, I call the former *two-way* powers and the latter *one-way* powers.

Although the distinction between one-way and two-way powers is genuine and fruitful, it is crucial not to overstate the difference. What we are now calling 'one-way' is nothing other than what we had been treating, following Molnar, as the generic concept of a power. Independently of any tacit contrast with two-way powers, to say a power is 'one-way' says nothing more than that it manifests itself given the opportunity—a feature in any case uncontroversially built into thought and talk about powers *per se*. 'One-way' was introduced to distinguish it from a *species* of power with additional differentiating features, which we called 'two-way.' In other words, strictly speaking, two-way powers are a species of one-way powers. For a property to be a two-way power, it must satisfy, minimally, the powers-framework's criteria for being a (one-way) power. This includes the fact that S manifests the power to φ if and only if S has the opportunity to φ. As such if S is in the right circumstances, so that S has an opportunity, but S's two-way power (its underlying capacities) are impeded in their manifestation, then it does not follow that S's

two-way power does not manifest itself. As a species of one-way powers, two-way powers *always* manifest themselves in the presence of opportunity.[37]

What differentiates two-way powers is not that, given the opportunity, they may or may not manifest themselves. Rather, as indicated, what differentiates them is their susceptibility to impediment, that is: they manifest themselves better or worse *qua* whatever two-way powers they are. This explains why sometimes in the above, 'manifestations' came qualified with the word 'proper.' Impediments to proper manifestation of a power make its manifestations less proper—i.e., worse *qua* manifestations of the power manifested—not no manifestation at all. Since impediments to two-way powers (their underlying capacities) explain privative manifestations, but privative manifestations count as manifestations of those two-way powers nonetheless, impediments to 'ability' interpreted in *actualist two-way powers*-terms do not falsify 'S has the ability to φ' as it figures in the consequent of the RIC principle. Clearly, if 'S has the ability to φ' reads 'S manifests the two-way power to φ,' and impediments to something's two-way power do not render suitable instances of this sentence schema false, then RIC-reasoning will not in these cases yield false instances of the antecedent 'S has a reason to φ.' So on the actualist two-way powers interpretation of 'can' there are conceivable impediments that do not rule out reasons-possession. This dispels the mystery about how the 'can' in RIC meets EC.

2.3. Conclusion

In order for the Reasons View to satisfy the richer theoretical perspective, the REASONS CONDITION requires a workable account of the conditions of reasons-possession. The overarching strategy of Chapter 2 was to employ RIC-reasoning to get at whatever sense of 'can' supplies the maximally non-trivial conditions of reasons-possession. The upshot of Section 2.1 was that the best Reasons View appeals to a special sense of 'can'—whether compatibilist or incompatibilist,

---

[37] When Aristotle, to whom we owe the notion of two-way powers, says "where it is in our power to act it is also in our power not to act" (*NE* 1113b7), I read him as saying *both* action and inaction count as manifestations of one and the same power. Similarly, Aquinas, who picks up on the Aristotelian theme, distinguishes objects with two-way powers as those "capable of determination in several ways and to various things" (*Summa Theologica*, Ia-IIae, 49, 4), by which he means that any of the several ways count as determinations (= manifestations) of the same capability (= power).

and so whether the 'can' expresses physical or non-physical modality. The special sense of 'can' is valuable to the Reasons View because it has non-trivial utility for RIC-reasoning whilst avoiding EC-violation. In Section 2.2 I explained just how the special sense of 'can' meets these two-fold constraints. It achieves this by bearing the powers-sense, more specifically the *actualist two-way powers*-sense. Objects such as to have normative reasons 'can' do a thing in this sense when they have the two-way rational power (and thus the underlying capacities) to do that thing plus the opportunity to do it. Like all powers, two-way powers *always* manifest given the opportunity, but those manifestations can be better or worse—they admit of privation. In other words, the major upshot of Chapter 2 is as follows. According to the best Reasons View it is a condition of normative reasons-possession, and so of moral responsibility, that rational power manifests itself (whether privatively or otherwise).

CHAPTER THREE

A Razian Reasons View

3. Introduction

In Chapter 2, I argued that the best Reasons View operates with an actualist two-way powers account of the conditions of reasons-possession. Whatever is such as to have normative reasons 'can' do a thing in this sense when they have the two-way rational power to do that thing plus the opportunity to do so. Though the idea of a power- or capacity-plus-opportunity condition is not in itself unfamiliar, the 'actualist' gloss I gave it may understandably lead one to think that the best Reasons View must be a rather outlandish view.

The goal of this Chapter is to exhibit Joseph Raz's work as bearing structural features of the best Reasons View, and to build on materials he provides in light of this interpretation of what he is up to. Although the Reasons View that emerges will be a construction out of recognisably Razian materials, it will, of course, be my own Razian view as opposed to Raz's own.[38] Raz's work, though dense and difficult, undoubtedly enjoys wide influence in contemporary Anglophone philosophy. Hopefully my proposal's Razian spirit, together with its theoretical benefits, go some way to establishing that the best Reasons View is not as outlandish as it might first appear. Before jumping in, however, let me start off by supplying some motivation for interpreting Raz in this way at all.

My suggestion is that the basic machinery Raz invokes when explaining the relationship between reasons and rational power looks a lot like an account of the conditions of reasons-possession operating an actualist two-way powers-interpretation of 'can.' If this can be substan-

---

[38] In my usage, the 'Razian' stands to Raz in whatever relation what is commonly labelled 'Marxian' stands to Marx. So 'Razian' means something like 'Raz-inspired,' and does not suggest that whatever it qualifies is Raz's own view, or even necessarily that he would endorse it.

tiated, it would evidence the stirrings of the best Reasons View in Raz, in turn supplying motivation for reading Raz's work along these lines. This is, admittedly, quite a bold suggestion, given the relative obscurity of the passages in which Raz lays out the machinery I have in mind. My hope is that closer examination of the relevant passages will make the suggestion more palatable, or perhaps, better still, even *prima facie* obvious. Here are the relevant passages, which I have divided into five chunks **[a]** to **[e]** for exegetical purposes.

"**[a]** Reason can malfunction. Therefore, reasons cannot be defined as what Reason recognizes and responds to. As Reason may fail there are criteria by which success and failure are determined, and they determine what reasons there are to be recognized. Put another way: **[b]** our rational powers are a general capacity to recognize and respond to facts that make certain responses appropriate, and **[c]** such facts are reasons because they can be recognized and responded to by our rational powers. […] **[d]** Reason does not make reasons into reasons (Reason is not a source of reasons). But **[e]** they are reasons because rational creatures like us can recognize and respond to them with the use of Reason." (2011, 86)

It is clear from **[a]**-**[e]** that Raz uses 'Reason' (capital 'R') interchangeably with 'our rational powers,' and that Reason is a power to φ given C, where determinates of φ are ways of recognising and responding to determinates of C (from **[b]** and **[c]**), and where determinates of C are facts that make certain responses appropriate (from **[b]**).[39] In specifying determinates of C not just as facts but facts that make certain responses *appropriate*, Raz indicates that Reason is such as to manifest itself privatively, that is, to respond inappropriately to the same circumstantial facts (from **[a]** and **[b]**). He calls these privative manifestations 'malfunctions' (see **[a]**), and it is clear from the surrounding discussion that he takes 'functionings' correspondingly broadly to include manifestations, privative or otherwise, so that Razian functionings encompass malfunctionings too.[40] Hence it is because Reason is a two-way power that only those facts that

---

[39] For further evidence that Raz's conception of Reason accords with something like our powers-framework, see, e.g. 1999, 76-7.

[40] For the surrounding discussion, see Section 3.1.

make certain responses appropriate present determinate circumstances for its determinate manifestation (from **[a]** and **[b]**).

It may well be that any fact could in principle make certain responses appropriate for some possible organism. But for a concrete organism of a specific kind, not all facts can make certain of its responses appropriate, even in principle. So for concrete organisms of a specific kind, those facts constitute a subset of the set of all facts. As concrete organisms of a specific kind, our simply being endowed with the two-way power of Reason is therefore enough to deselect a range of facts as even *potentially* reasons, namely, those facts that do not make certain of our responses appropriate (this unpacks **[b]**). By the same token, each of us, depending *inter alia* on our spatio-temporal location, find ourselves faced with facts that do make certain of our responses appropriate—we find ourselves in determinates of determinable circumstances C. Faced with circumstantial facts that make certain of our responses appropriate, we manifest our Reason by recognising and responding to those very facts, privatively or otherwise. So chunk **[c]** is felicitously unpacked as follows. Concrete organisms faced with those circumstantial facts possess corresponding reasons because their Reason manifests itself in recognising and responding to those very facts, privatively or otherwise. In other words, the 'can' in **[c]** bears the actualist two-way powers-sense, which, in context, expresses a power- or capacity-plus-opportunity condition of reasons-possession.

This makes sense of the difficult chunks **[d]** and **[e]**.[41] When Raz says that Reason does not make reasons into reasons, he should be read as saying that though endowment with Reason suffices to select certain facts as possible reasons—i.e. those that make certain of our responses appropriate—it does not *create* them in that character. Their existence as a subset of the set of all facts is independent of and presupposed by Reason's selecting them. This unpacks **[d]** in parallel with **[b]**. When Raz says they are reasons because we 'can' recognise and respond to them with the use of Reason, as he does in **[e]**, the 'can' here is that of power- or capacity-plus-opportunity, in parallel with **[c]**. Admittedly, Raz opts to speak in terms of circumstantial facts *being* reasons in virtue of Reason, whereas I read him as making claims about when we *possess*

---

[41] Others have found these chunks difficult. Ulrike Heuer observes that they are in at least *prima facie* tension with one another (2013, 186). Doug Lavin highlights them as expressive of Raz's middle way between realism and idealism about reasons (2011, 368), and goes on, albeit sympathetically, to problematise this middle way.

reasons. But then again, not all facts that make certain of our responses appropriate are themselves reasons, and those that are receive that status by falling within the circumstances of concrete Reason-endowed organisms. Facts are only ever reasons in relation to such organisms, or to collections of them perhaps. That is, they are only ever reasons in relation to some S that has them. It may be ungrammatical to speak of the 'having' of circumstantial facts. But we certainly do 'have' circumstances.

This concludes my attempt to motivate reading Raz's work as a best Reasons View in the first place. As mentioned, this Chapter's goal is to further exhibit Raz's work as one such view, and to develop my own Reasons View out of Razian materials in light of explicit acknowledgement of this character of his work. Recall that for the best Reasons View, there are two ways for S not to have a normative reason to *φ*: by lacking Reason altogether, and by lacking the opportunity to *φ*. Section 3.1 investigates these aspects of Raz's work. After briefly considering what Raz says about *total* lack of Reason, I suggest that he overlooks the distinction between total and *partial* lack. The rest of Section 3.1 is dedicated to building the notion of partial lack into a Razian Reasons View. Doing so involves sustained reflection on the practical-rational profile of what often gets labelled 'disability,' revealing it to be a specific form of lack of *opportunity*. This yields an attractive way of circumscribing those passages of life that count as Reason's manifestations from those that do not, in turn deepening our understanding of the best Reasons View itself.

The best Reasons View says that in the absence of neither Reason nor an opportunity to manifest, Reason manifests itself (privatively or otherwise). Section 3.2 considers what goes on when Reason manifests itself privatively, or in Raz's terms, Reason's *malfunctionings*. The Section begins by presenting and rejecting the distinction between core and ancillary capacities lying at the heart of what Raz says about malfunctioning. I go on to reconstruct a functionally parallel distinction out of Razian materials, and walk through how it may be used to introduce familiar practical-philosophical terms of art: non-accidents (intentional or reckless), carelessness (including negligence), pure accidents, fault, and excuses. In closing, I offer some programmatic remarks about how my Razian Reasons View can accommodate resultant moral luck and strict liability in law.

3.1. 'Can'-Not: Lack of Power or Opportunity

According to Raz, Reason—or 'capacity-rationality,' as he sometimes calls it—is immediately grounded in a constellation of subcapacities, such that lacking any one of them leaves an organism without Reason altogether (1999, 68-9).[42] There are, of course, palpable barriers to conceptualising the relevant subcapacities' diverse functions very clearly, so listing them will inevitably be an impressionistic exercise to be taken without comprehensive or eternal ambitions. With this caveat in place, and loosely organising Raz's own list, I assume in what follows that the subcapacities include *inter alia* perceptual-attentional, motor, cognitive-emotional and volitional capacities.

Recall that on our powers-framework (see Section 2.2), a macroscopic object's powers are 'derivative' in that they are immediately grounded in underlying capacities, where underlying capacities are further powers either of that macroscopic object itself or of some of its suitably related parts. For instance, the capacities underlying a glass's fragility will be located in its suitably related parts—its atoms and molecules, say, in some specific arrangement. By contrast, the subcapacities underlying a concrete organism's Reason are probably better located in the whole rather than its parts. No parts of a concrete organism are themselves the bearers of perceptual-attentional or cognitive-emotional capacities (say), even if these grounding powers are grounded in yet further powers of the organism's parts (perhaps anatomically circumscribed). In any event, the fact that Raz grounds Reason in various subcapacities further buttresses the claim that his work exemplifies the best Reasons View.

For Raz, then, total lack of any of the subcapacities suffices to remove Reason from its bearer. It is important to be clear about what exactly this means. *Total* lack is not *partial* lack. This is

---

[42] This makes Razian Reason *concrete* rather than *abstract*, avoiding fallacies associated with the notion of abstract rational power. Only if it is concrete does it make any sense to say that rational power lacks the opportunity to manifest itself on a given occasion. The very abstractness of abstract rational power presumably lets it float free of constraint by anything so worldly as opportunity. Compare Maria Alvarez's discussion of the rational power of agency. Manifestations of agency, she says, require *inter alia* "the opportunity to φ and the opportunity not to φ" (2013, 108), so that whether it manifests or not "will be up to me" (*op. cit.*, 110). But, as we saw in Section 2.2, the same determinate circumstances cannot present powers with both the opportunity to manifest *and* not to manifest. So perhaps Alvarez means that manifestations of agency require the opportunity to $<\varphi$ or not-$\varphi>$. But then all conceivable circumstances present an opportunity for manifestation, that is: agency floats free of opportunity altogether.

a crucial distinction I believe Raz overlooks, but which can be constructed out of materials he provides. The absence of the notion of partial lack can be brought out by reflection on the place within Raz's picture of what contemporary Anglophone political culture calls 'disability'. Raz gives the following as a case of total lack of volitional capacity: "Someone who lacks a minimal capacity to make up his mind about anything […] lacks capacity-rationality" (1999, 69). This description evokes those living with certain forms of profound cognitive impairment or of mental disorder. Indeed, this reading makes natural Raz's subsequent contrast with people who merely "constantly vacillate, finding it always difficult to make up their minds," and who are on that account criticisable (ibid.). But cognitively impaired and mentally disordered people certainly do not lack Reason altogether, nor are they typically thought criticisable for those φ affected by their respective conditions.[43] Similarly, of perceptual-attentional and motor capacities, Raz says that so long as they are not totally lacking, then "lacking the ancillary [i.e. perceptual-attentional and motor] capacities may impede the exercise of rationality" (*op. cit.*, 68n2).[44] This suggests that persons living with visual impairment or motor disorder (say) manifest their rational power privatively if they fall short of some φ due to their condition. Once again, this does not seem to get things quite right.

What the examples show is that, for Raz, disability sits awkwardly between total lack and privative manifestation. Without a further distinction between total and partial lack, Raz's work does not accommodate disability very well at all. Raz and I agree that total lack of any of the subcapacities, if such there can be, suffices to remove Reason from concrete organisms. To my imagination there could be no condition observable of any person, nor behaviour they could exhibit, which would satisfy this description with anything approaching certainty—other than the condition of death. Conversely, as I understand it, partial lack of any or all of the subcapacities leaves Reason perfectly intact. To get a better grip on partial lack and its possible location within a Razian Reasons View, I will first give a tour of the relevant areas of Raz's work, and

---

[43] If members of these groups are said to lack Reason altogether, then, for Raz, they will no longer count as persons, since "Rationality [i.e. capacity-rationality] makes us into persons" (*op. cit.*, 68). Needless to say, though neither option satisfactorily accommodates members of these groups, this particular option comes with especially troubling moral consequences.

[44] By 'ancillary' subcapacities, Raz means those of Reason's underlying capacities "which do not directly contribute to our rationality," in contrast with those that do, e.g. cognitive-emotional and volitional capacities (*op. cit.*, 68-9). Here, 'rationality' carries the sense in which our failing to exhibit it makes us criticisable. Section 3.2 problematises this distinction between core and ancillary subcapacities.

then elaborate upon the opportunity-conditions of functionings of Reason and its underlying subcapacities.

Since, as I contend, Raz's work exhibits structural features of the best Reasons View, it is no surprise that he believes responsibility is a matter of rational power's manifestation—or as he would say, Reason's functioning—as expressed in the following principle:

"**Rational Functioning Principle**: Conduct for which we are (non-derivatively) responsible$_2$ [i.e. the sense of 'responsibility' we are after] is conduct that is the result of the functioning, successful or failed, of our powers of rational agency, provided those powers were not suspended in a way affecting the action." (2011, 231)

The point of Raz's principle is to explain responsibility not only for conduct guided by Reason, but also for unintentional omissions and actions—in short: *accidents* (*op. cit.*, 243). This is why functionings come qualified as "successful or failed." They are manifestations of rational power such as to admit of privation, as per the best Reasons View's account of the conditions of reasons-possession. In turn, for Raz, conduct counts as Reason's functioning if it falls within what he calls the *domain of secure competence*. All such conduct is our responsibility, again in keeping with the best Reasons View. Hence Raz says, "we are responsible$_2$ because we are rational agents, but only for those actions […] regarding which our capacities of rational agency were available to us," where those actions are available to our capacities "so long as they belong within our domain of secure competence" (*op. cit.*, 246). It is crucial to note Raz's repeated reference to different persons' *respective* domains of secure competence, which indicates that what falls within the domain of secure competence can differ for different persons occupying identical circumstances (*op. cit.*, 245).[45]

Given that functionings of Raz's Reason are such as to include malfunctionings, a concrete functioning might look like anything from solidly φ-ing, through φ-ing so badly that it cannot in good faith be called φ-ing, all the way through to doing nothing whatsoever. This invites the

---

[45] He does however complicate this by suggesting that there is usually a community-relative "basic domain of common competence" beyond which "individuals develop further more advanced competences" (*op. cit.*, 247), and that it is justified to hold people to reasonable social expectations about what falls within this community-relative domain (ibid.).

question how to distinguish those passages of concrete Reason-endowed organisms' lives that count as its functionings—and therefore its responsibility—from those that do not (if any do not). At least part of the point of Raz's domain of secure competence is to answer just this question. The domain of secure competence, he says,

"[M]ark[s] the most basic domain in which agents' rational powers are available and functioning (as they automatically are when available), even though not always perfectly." (*op. cit.*, 250)

Clearly, then, Raz's point with the domain of secure competence is at least partly to circumscribe Reason's functioning, and with it, the matter of responsibility. Recall that for the best Reasons View, it is the presence or absence of opportunity that circumscribes manifestations of rational power (privative or otherwise), and those manifestations in turn conceptually condition reasons-possession, and with it, responsibility. So interpreting Raz's work as a best Reasons View recommends understanding the domain of secure competence in terms of opportunity. But does what Raz says sustain this way of seeing things?

According to our powers-framework, whatever a macroscopic object S lacks the opportunity to do cannot count as manifestations of S's power P (to φ given circumstances C), whereas everything else must so count. S has the opportunity to φ when, and only when, *ceteris paribus*, S's determinate circumstances present *all* the capacities underlying P with the opportunity to make their respective contributions to φ. The *ceteris paribus* clause indicates that determinate values of C are *normally* sufficient to present an opportunity for the underlying capacities to do their respective bit. Suppose a concrete organism S possesses Reason (a two-way rational power to φ given C) and has the opportunity to φ, so that S manifests its two-way rational power to φ, whether privatively or otherwise. In that case, S is in such circumstances that, normally, none of S's underlying subcapacities lacks the opportunity to do its bit. To see how this is supposed to work, consider the following unfortunate scene.

EARPHONES: After a long day of Zoom calls, Bob is about to enjoy a quiet vape out of his second-floor window. He is just in time to see a sentimental urn topple and fall from

the opposite window of the flat across the road. Down on the pavement below, Passerby happens to be passing by. She will catch the urn if alerted of its falling. Bob shouts to alert Passerby, but she has earphones in. She fails to hear Bob's shout, and the urn smashes on the pavement, kicking up a cloud of ashes.

It is uncontroversial that Bob lacks the opportunity to catch the sentimental urn. This is because, normally, the motor and volitional capacities of someone in his circumstances lack the opportunity to contribute to their catching it. By the same token, he does have the opportunity to alert Passerby down below—an opportunity he takes. Unlike Bob, but again, by the same token, Passerby does have the opportunity to catch the urn, since the perceptual-attentional capacities of someone in such circumstances do not normally lack the opportunity to contribute to catching the urn. Unlike Bob, then, Passerby's failure to catch the urn manifests their rational power to do so, because Passerby but not Bob has the opportunity to do so. Of course, in failing to catch the urn, Passerby manifests their rational power privatively.[46]

Reference to the 'normalcy' of circumstances built into our powers-framework makes the best Reasons View's notion of opportunity *universal* in the following sense: different persons occupying determinates of one and the same determinable circumstances do not differ in the opportunities presented to rational power. As we have seen, Raz's domain of secure competence is not like this. What falls within it is *individualised* in the sense that this could differ for different persons in determinates of the same determinable circumstances. This is why Raz's domain of secure competence cannot be straightforwardly understood in terms of opportunity. Here I will argue that Raz has gone astray, not only *qua* Reasons View, but also by his own explicit lights. That *the* domain of secure competence is really *a* domain of secure competence strikes me as not only wrong, but in fact at odds with other things Raz says about it. Let us return to EARPHONES above. For Raz, it is possible that catching the urn falls beyond Passerby's domain even as it falls within the domain of another passerby with earphones in, regardless of

---

[46] This is to say Passerby does indeed have a normative reason to catch the urn, and therefore, absent further qualification, that Passerby is responsible for failing to catch the urn according to the Reasons View. It may be thought offputtingly harsh that Passerby comes out responsible. But here, as in Chapter 1, this is not (yet) to say Passerby is *blameworthy*.

which of them—call them Passerby One and Two—is more representative of what the circumstances normally present. This possibility, I submit, is at odds with the following remark.

"Failure to control conduct within our domain of secure competence threatens to undermine our self-esteem and our sense of who we are, what we are capable of, etc. […] We have grown frail, our competence is diminishing." (*op. cit.*, 245)

Presumably if Passerby One's domain of secure competence does not contain catching the urn, then it is 'diminished' relative to that of Passerby Two (which contains catching the urn), making Passerby One more 'frail' than Passerby Two. For Passerby One to find their relative frailty agent-regrettable, they must have been responsible for failing to do the thing lying outside their domain, viz. catching the urn. For Raz, we have seen, what falls beyond the domain of secure competence cannot be a functioning of Reason, and therefore cannot be something that that person is responsible for. But if Passerby One and Two have their own respective domains, then Passerby One is not responsible for the very failure (to catch the urn) that is presupposed by the agent-regret Raz invokes above. In other words, how could something Passerby One is not responsible for precipitate agent-regret about the very frailty that pushes such things beyond their responsibility? The answer is that it cannot. Perhaps Raz would say Passerby One coherently agent-regrets their relative frailty because they are responsible for the frailty *itself*, i.e. their having a less expansive domain of secure competence than Passerby Two. For Raz, this would presuppose the presence of an act-type like 'expand one's domain' within Passerby One's domain. But since, by hypothesis, domains of secure competence are individualised, the presence of this act-type within one's domain— and therefore the grounds of coherence of the associated agent-regret—will have to be individually variable in principle. Such individual variability is quite counterintuitive.[47]

---

[47] The point generalises from the inter- to the intrapersonal case. If domains of secure competence are individualised to *time-slices* of persons, then persons can be frail relative to themselves at other times (whether past or future) at which their domains were more expansive. For any time-slice of anybody to coherently agent-regret such frailty, they would have to be responsible for it, which presupposes the presence of expanding one's domain within their very domain. Whether or not it is present can in principle vary amongst time-slices of persons, which is, once again, a counterintuitive result.

For a given passage of conduct of a (time-slice of a) person, what determines whether or not it counts as Reason's functioning—and so a matter of responsibility—must therefore be universal as opposed to individualistic. In what follows, then, I build my own Reasons View on broadly Razian foundations whilst rejecting his claim that what circumscribes Reason's functionings is the domain of secure competence. Instead, I work from the assumption that it is the presence of opportunity that fulfils this theoretical role, where this means that the circumstances Reason's bearer finds themselves under *normally* present the underlying subcapacities with the opportunity to contribute to the relevant functioning.

How does this help with the notion of partial lack? The connection can be brought out by reflection on what is commonly labelled 'disability.' This label is misleading in a number of ways. To start with, those with disability do not totally lack any of Reason's subcapacities, so neither this nor the attendant lack of Reason can be what 'disability' expresses. Less objectionably, it may be thought that 'disability' expresses a diminishment of the opportunities presented to affected persons. Without further qualification, however, this is also misleading. To see why, consider the following variation on the EARPHONES case.

> ZENA: After a long day of Zoom calls, Bob is about to enjoy a quiet vape out of his second-floor window. He is just in time to see a sentimental urn topple and fall from the opposite window of the flat across the road. Down on the pavement below, Zena happens to be passing by. Zena is deaf, unable to lipread, and an excellent user of British Sign Language (BSL). She will catch the urn if alerted of its falling. Bob is ignorant of BSL and assumes Zena is not deaf. He shouts to alert Zena, who sees him, but cannot make out the content of his speech. The urn smashes on the pavement, kicking up a cloud of ashes.

If the 'normalcy' of the opportunity presented to people's subcapacities by circumstances like these is not to be tacitly ableist, BSL-ignorant Bob must be said to have the opportunity, and thus a reason, to alert Zena. Just as with Passerby One's agent-regret, without saying this it becomes hard to explain why, were BSL-ignorant Bob to learn the true nature of the situation *ex post*, he might become rationally motivated to learn BSL. By the same token, Zena must be said to have the opportunity, thus a reason, to catch the urn. This explains why, *ex post*, Zena

61

might become rationally motivated to learn how to lipread. Act-types specifying the opportunities circumstances normally present to Reason are highly generic, and unaffected by variations in the underlying subcapacities. This is why it is misleading to say 'disability' expresses a simple contraction of the range of opportunity, at least absent further qualification.

The relevant further qualification can be seen if we consider resistance to this claim along the following lines. Can we not say that Zena lacks the opportunity, via a lack of opportunity for her perceptual-attentional capacity, not to catch the urn *simpliciter* but rather to *try* to catch the urn? If this is correct, then the range of opportunity in this way does suffer a contraction, so that the language of 'disability' is not misleading after all. This sounds plausible as far as it goes, but unless it goes further it remains a tacitly ableist argument. What is true of Zena is equally true of BSL-ignorant Bob, whose Reason, via lack of opportunity for his cognitive and motor capacities, lacks the opportunity to *try* to alert Zena in the way his circumstances demand, despite having the opportunity to alert her *simpliciter*.

It is important to be precise about the sense of 'trying' at play here.[48] To have the opportunity to try to φ in the relevant sense, our underlying subcapacities must each have the opportunity to make their respective contributions to trying to φ. In this sense, BSL-ignorant Bob and Zena lack the opportunity to try. But there is a sense of 'trying' they do not lack the opportunity to do, that is, that their respective underlying subcapacities do not lack the opportunity to contribute to. Consider a parent, who knows they cannot swim, party to their child drowning at sea. If they dive into the sea, they do not do so as an effective means to rescuing their child, since they are, as they know, a non-swimmer. Doing so need not be without reason. Diving in despite knowing they cannot swim might be an expression of the depth of their love. Their subcapacities have the opportunity to contribute to 'trying' in this *expressive* sense, and they have, correspondingly, a reason to expressively try. But their reason to rescue their child is not derivatively a reason to expressively try, since expressively trying is not derivable in deliberation as a means to rescuing their child.[49] By contrast, were they a swimmer, then trying would be a means to rescuing their child, so that their reason to rescue their child would derivatively be a reason to

---

[48] The following ideas about 'trying' are borrowed from John Gardner (2004) and have been significantly influenced by Heuer (2010).

[49] For the idea that some reasons to try are (derivatively or otherwise) *identical* with reasons to succeed, see Gardner (2004, 54-5).

try—a reason to *effectively* try, rather than to expressively try. That their subcapacities have the opportunity to contribute to effective (not expressive) trying is what makes their reason to succeed derivatively a reason to effectively try. BSL-ignorant Bob and Zena may well have the opportunity to expressively try, but not to effectively try (respectively: to alert Zena, to catch the urn).

The subcapacities underlying the Reason of BSL-ignorant Bob and Zena lack the opportunity to contribute to effectively trying to do things that others, in the same determinable circumstances, might have, and *vice versa*. In other words the contraction of opportunity cuts both ways if it cuts at all. If those whom dominant political culture labels 'disabled' are disabled relative to the 'nondisabled,' then the 'nondisabled' are always in other respects disabled relative to the 'disabled.'[50] Since the ascription-rules of the popular label 'disability' do not distribute ascriptions in a way that reflects this fact, the popular label remains misleading. A more perspicuous terminology might say people are *differently abled*, though with the caveat that there are other terminological virtues plausibly achieved by the label 'disability,' such as political utility for actually existing emancipatory projects.

Let me briefly recap. Raz's work exhibits structural features of a Reasons View operating an actualist two-way powers-interpretation of 'can.' This explains why total lack of any of Reason's underlying subcapacities precludes its functioning, and with it, responsibility. The same goes for what falls outside Raz's domain of secure competence, which, in circumscribing Reason's functionings (when Reason is not totally lacking), fulfils the same theoretical role as opportunity on the best Reasons View. I rejected Raz's domain of secure competence on the basis of its indi-

---

[50] This is to accept what Timothy Macklem and John Gardner call the 'reciprocity thesis' (2014, 64). Macklem and Gardner reject the reciprocity thesis on the grounds that the 'nondisabled' have the luxury, in principle, of augmenting their subcapacities such as to have the opportunity to do things how the 'disabled' do, whereas the 'disabled' do not have the reciprocal luxury. This makes their view of disability what Elizabeth Barnes (2014) calls a *bad*-difference view rather than a *mere*-difference view. It is disappointing that Macklem and Gardner adopt a bad-difference view, given that their paper contains the resources to sustain the reciprocity thesis. Though they entertain participation in the Paralympics as a potential example of a luxury the 'nondisabled' do not in principle have, they reject it on the grounds that this lack is merely rule-based (2014, 64-5). Without some additional premises, it is hard to see how this rejection follows. We ordinarily allow what is possible 'in principle' to be circumscribed by rules just as often as by individual biology. Surely an in-principle reciprocal lack of luxury can just as well be generated by social rules as by individual biology? Since the additional premises are not forthcoming, I see no good reason to drop the reciprocity thesis. This makes my Razian Reasons View a mere-difference view politically sensitive to claims made by disability-rights activists.

vidualism, and proposed my own Razian Reasons View with the notion of opportunity in its place. I went on to consider the special case of lack of opportunity to effectively try, which sets itself apart from standard cases by ruling out possession not of reasons to φ *simpliciter*, but of those very reasons in their status as derivatively reasons to effectively try to φ. Finally, I argued that this specific form of lack of opportunity encapsulates the practical-rational profile of differently-abledness.

How does my Razian Reasons View improve on Raz's work as regards responsibility and the differently abled? Now at long last is the time to discharge a promissory obligation incurred back in Section 1.1. There, an *ad hoc* amendment was made to clause (i) of the condition on moral responsibility, which resurfaced in Chapter 2 as the REASONS CONDITION underpinning the Reasons View. The *ad hoc* amendment concerned disability:

REASONS CONDITION: S is morally responsible for φ iff (i) S φs (not from a disability), and (ii) S has a φ-reason.

Given our discussion of the label 'disability,' the REASONS CONDITION can be immediately given an initial and transitional reformulation as follows.

REASONS CONDITION$_1$: S is morally responsible for φ iff (i) S φs (not from differently-abledness), and (ii) S has a φ-reason.

As we have seen, on my proposal, the practical-rational profile of differently-abledness is constructible as follows. Due to variations in Reason's underlying subcapacities, *everybody* is differently abled with respect to a range of instances of φ given C. What this means is: for appropriate pairs of substitutions into φ and C, though everybody has the opportunity to φ given C, some people lack the opportunity to effectively try to φ given C. Act-types figuring in specifications of what opportunities circumstances normally present to Reason are highly generic, and when (and only when) they are effective tryings, their figuring in such specifications can be affected by variations in underlying subcapacities. This allows my Razian Reasons View to help itself to the notion of partial lack. As I understand it, to experience partial lack of one or more

underlying subcapacities *just is* to be differently abled in precisely the sense at hand. Let me give the REASONS CONDITION another transitional rendering that makes explicit the place of partial lack.

REASONS CONDITION$_2$: S is morally responsible for φ iff (i) S φs (not from partial lack), and (ii) S has a φ-reason.

Since what Anglophone political culture popularly labels 'disability' is in fact simply a tacitly ableist selection from the wider phenomenon of differently-abledness, the identification of partial lack with differently-abledness lets my Razian Reasons View accommodate disability as follows. On the one hand, partial lack is not total lack, and so those who experience it not only do not lack Reason—they moreover do not lack the opportunity (thus the reason) to do things affected by their variety of subcapacities. In terms of REASONS CONDITION$_2$, they satisfy clause (ii) but not (i). Whilst BSL-ignorant Bob and Zena have a φ-reason (to alert Zena; to catch the urn), due to their experiencing partial lack these are not derivatively reasons to effectively try to φ. On the other hand, things done from partial lack are not criticisably privative manifestations of Reason. By REASONS CONDITION$_2$, if S φs from partial lack, then S does not come out morally responsible for φ, and so *a fortiori* S cannot come out blameworthy for φ. This is not to say S has no φ-reason *simpliciter*; only that S has no reason to effectively try to φ. Again, S may well satisfy (ii) but not (i).

A final reformulation of the REASONS CONDITION unfolds the *ad hoc* amendment to clause (i) in such a way as to completely rid it of reference even to the notion of partial lack, and with it any remaining worries about its *ad hoc* character. If φ falls within the range of S's partial lack, then S's Reason (via one or more of its underlying subcapacities) lacks the opportunity, not necessarily to φ given C *simpliciter*, but to effectively try to φ given C. This is why, given C, if S has a reason to φ, it will not derivatively be a reason to effectively try to φ. Thus, were S to φ, it is obvious why this should make 'Why?'-questioning unintelligible with respect to S's φ, in the strong sense of INTELLIGIBILITY$_3$ (see Section 1.1). It is intelligible in this sense to interrogate S's φ only if it is intelligible (in the same sense) to interrogate S's effectively trying to φ, that is, S must have the opportunity, and thus a reason, not only to φ but to effectively try to φ. Since the

REASONS CONDITION does nothing more than rename INTELLIGIBILITY₃, the preceding observations indicate how to unfold the *ad hoc* elements of clause (i) into the more substantive clause (ii) without loss. Clause (ii) concerns reasons-possession, and the *ad hoc* partial lack element can be stated in terms of reasons-possession. Here, then, is the final formulation of the best Reason View's core analysis of moral responsibility.

REASONS CONDITION₃: S is morally responsible for φ iff (i) S φs, and (ii) S has a φ-reason that is derivatively a reason to effectively try to (omit to) φ.

To see the value of this less *ad hoc* condition, consider what it lets the best Reasons View say about our friends from ZENA and EARPHONES. By RIC, Zena has a reason to catch the urn, but it is not derivatively a reason to effectively try to catch the urn. This explains why she might be rationally motivated to learn lipreading, but without making her responsible for failing to catch the urn. Similarly BSL-ignorant Bob has a reason to alert Zena, but it is not derivatively a reason to effectively try to alert Zena. This explains why he might be rationally motivated to learn BSL, but without making him responsible for failing to alert Zena. The cases of Zena and BSL-ignorant Bob should be contrasted with the original Passerby.[51] Unlike Zena and BSL-ignorant Bob, Passerby experiences no partial lack. Having earphones in makes no difference to the opportunity Passerby's circumstances *normally* present to the perceptual-attentional subcapacities of someone so placed. Passerby not only has the opportunity to catch the urn, but the opportunity to effectively try to do so. Her reason to catch the urn is therefore derivatively a reason to effectively try to do so, and so she is responsible for failing to do so. By contrast, original

---

[51] A remark on technological change. BSL and lipreading are technologies. The subcapacities of neither BSL-ignorant Bob nor Zena have the opportunity to contribute to effectively trying to do what their respective circumstances demand because neither can meaningfully deploy these technologies. This shows why *digital illiteracy* is a form of partial lack. It is increasingly normalised in the UK that effectively trying to access state support requires digital literacy, which the subcapacities of the digitally illiterate lack the opportunity to contribute to. The digitally illiterate are therefore not responsible for failure to access state support. Given that such failure routinely incurs negative consequences including punitive sanctions (e.g. suspension of benefits payments, etc.), the rapid normalisation of digital access to state support without recourse for the digitally illiterate is as gravely unjust as it would be to hold people who φ from disability responsible (indeed blameworthy) for φ-ing. This is a sentiment broadly echoed by the UN Special Rapporteur on extreme poverty and human rights in his 2019 report on the UK, available at https://undocs.org/A/HRC/41/39/Add.1 (accessed 8th June 2020), especially section VI, C.

Bob, who is upstairs across the road, lacks the opportunity to catch the urn, and so is not only not responsible (by REASONS CONDITION₃) but has no grounds for agent-regret (by RIC).

Before moving on, let me summarise the developments of Chapter 3 so far. Raz's work has been exhibited as bearing structural features of the best Reasons View. After presenting some problems in Raz's work, I proposed a variation on Raz's theme that better accommodates disability by understanding partial lack as a specific form of lack of opportunity. The Razian Reasons View that emerges offers an attractive way of differentiating what does and does not count as Reason's functioning, privative or otherwise. In turn, it unfolds the *ad hoc* REASONS CONDITION into the more straightforward REASONS CONDITION₃. My hope is that the effect of all this is to make the best Reasons View look more palatable.

3.2. Malfunctionings

According to the best Reasons View, it is necessary and sufficient for someone's moral responsibility for some thing that they possess a normative reason bearing on that thing, which is derivatively a normative reason to effectively try to do that thing. On the Razian Reasons View proposed in Section 3.1, a person satisfies this condition so long as: they have Reason, their Reason has the opportunity to do the thing, and they do not experience partial lack with respect of such things. The subsequent manifestation of Reason stands under a normative reason the person thereby possesses. This accounts for the person's responsibility for the manifestation, whether privative or not. In this Section I consider what happens when people lack *neither* Reason (whether totally or partially) *nor* the opportunity to manifest it, and so thereby manifest it. My focus will be on privative manifestations of Reason, or in my preferred Razian idiom, those functionings that are *malfunctionings* (2011, 243).

Whether a manifestation is privative depends on how the proper manifestation of the power whose manifestation it is should look, according to the nature of the power that it is. Another way of putting this is that standards internal to two-way powers are constitutive of their determinate manifestations. In Reason's case the internal standards are *rational* standards, which is

why functionings of Reason that do not meet these standards count as malfunctionings.[52] There is a live question regarding whether or not *each* of Reason's underlying subcapacities is constitutively rational. If they are not, then if it is an arational subcapacity that fails, then Reason's malfunctioning cannot evince, as it were, a direct disregard for rational standards. But if they are, then regardless of which subcapacity fails, it is possible that Reason's malfunctioning evinces direct disregard. In drawing the distinction between *core* and *ancillary* subcapacities, Raz gives the first answer to this question (1999, 68-9). This is a distinction I wish to problematise.

Raz deploys what he calls the 'irrationality test' to demarcate core and ancillary subcapacities (2011, 91-3). The irrationality test says a subcapacity is constitutively rational if and only if its failures can be *non-derivatively* irrational. A failure is non-derivatively irrational when our calling it 'irrational' is not owed to its being explained by a further irrational failure, i.e., it is not *derivatively* irrational. Uncontroversial examples of subcapacities whose failures can be non-derivatively irrational are the cognitive and volitional capacities. The non-derivative irrationality of bad reasoning, at least when undertaken seriously, shows that the cognitive capacities are constitutively rational. The non-derivative irrationality of akratic or anomic tendencies in decision-making shows the volitional capacities to be constitutively rational. For Raz, this shows these underlying subcapacities to be Reason's *core* subcapacities. By contrast, he thinks that when our perceptual-attentional (*op. cit.*, 91) and motor (1999, 69) subcapacities fail us, their 'irrationality,' if this is apt at all, will always be derivative. Hence although these subcapacities underlie Reason, they are demoted to an *ancillary* status.

So if a core subcapacity's failure is what amounts to Reason's malfunctioning, its irrationality may be non-derivative, meaning that the malfunctioning evinces direct disregard for its constitutive rational standards. Still, it may be that the failure is derivatively irrational. If S, given the opportunity, falls short of effectively trying to φ due to akrasia, then it is a non-derivative failure of S's volitional capacity that amounts to this malfunctioning, which evinces direct disregard for

---

[52] There is significant controversy over how to understand 'rationality' and its normativity. Here, talk of a functioning's having internal 'rational standards' simply means that one or more normative reasons bear upon it that it may conform to more or less well, i.e. that it stands under a normative reason (see also footnote 35). In Niko Kolodny's helpful terms, this is more an *objective* rather than a *subjective* use of 'rationality' (2005, 510). Though Raz often employs the subjective use of 'rationality,' it strikes me that his understanding of the constitutive 'rationality' of powers (and their manifestations) is, ultimately, an objective notion.

rational standards. But if S falls short owing to a decision to do otherwise, made on dodgy perceptual grounds acquired in the heat of the moment, then it is a derivative failure of S's volitional capacity that amounts to malfunctioning, which at most evinces *indirect* disregard. This is because the volitional failure is only aptly called 'irrational' if it is explained by a yet further failure aptly called 'irrational.' For Raz, the present case need not be like this, because the upstream failure of an ancillary subcapacity (viz. perception) cannot be non-derivatively irrational. Of course, it may be that the perceptual failure is in turn explained by a non-derivatively irrational failure. But it need not be. So for Raz, derivative failures are only aptly called 'irrational' when the explanatory chain—possibly one-membered—terminates in a non-derivatively irrational failure. Where this is so, I will say that the derivative failures are *traceable* to the non-derivatively irrational failure.

For Raz, all irrationality, and so all disregard for rational standards (whether direct or indirect), is traceable to a core subcapacity's failure. This is because failures of ancillary subcapacities cannot be non-derivatively irrational. Here, I present two arguments to the contrary: the so-called 'ancillary' subcapacities *can* in fact exhibit non-derivatively irrational failure. This effectively makes each of Reason's underlying subcapacities 'core,' rendering Raz's core-ancillary distinction defunct.

My first argument generalises Raz's own strategy for showing our emotional capacities to be constitutively rational. Raz distinguishes the derivative irrationality of fear when traceable to irrational belief from the non-derivative irrationality of fear when it disproportionately responds to perfectly rational belief (2011, 92). By Raz's test, it is the latter that reveals the constitutive rationality of emotion. Now, notice that the perceptual-attentional and motor capacities are also derivatively irrational when traceable to irrational belief.[53] My thought here is to complete the analogy with emotion with examples of non-derivatively irrational perceptual-attentional and motor failures. There is a limited range of cases of *implicit cognitive bias* suited to this job. Like emotion, our capacities for perception, attention, memory and movement are all rightly heightened in quality by rational belief calling for such heightening, e.g., non-accidentally true beliefs about the delicacy of the practical context. These capacities are likewise diminished in quality by the presence of rational beliefs that call for no such diminishment, e.g., true but contextually

---

[53] As in a subset of cases of what Susanna Siegel (2017) calls 'wishful seeing.'

irrelevant statistical generalisations about (say) incidence of crime amongst a group to which *this* person belongs. If any of these capacities go awry like this, explanation of the failure need not look beyond the capacity itself. The irrationality is non-derivative. Many cases of implicit cognitive bias are of course derivatives of full-blown (if deeply buried) false beliefs. But if failures within this limited range of cases are aptly called 'irrational,' which plausibly they are, then even the ancillary capacities must be constitutively rational.

My second argument concerns the irrationality test. The irrationality test partly trades on linguistic intuitions about which failures are properly called 'irrational.'[54] But linguistic intuitions are often not very principled. To my mind, any plausible organising principle would discern that the things we call 'irrational' are unified not only by their being privative manifestations of rational power, but their being *criticisably* so. What ordinary moral thought describes as forms of *clumsiness* satisfy this principle. However, though clumsiness is usually (in suitable contexts) considered criticisable, it is not usually called 'irrational.' So do failures from clumsiness pass the irrationality test, or not? It is right, I believe, to give clumsiness a pass, by being led less by linguistic intuitions about what counts as 'irrational' and more by what we convey by its use. This, in turn, puts pressure on the core-ancillary distinction. Though 'clumsiness' is typically used to qualify motor failures, it is, I think, used in recognisably the same way when describing failures of other subcapacities. Under ideal conditions, obliviousness to salient environmental detail is a clumsiness of perception, inattentiveness to known or easily inferable facts is a clumsiness of memory or cognition, being insensitive is a clumsiness of cognitive-emotional capacities, and so on. It pays to emphasise the stipulation of ideal conditions here, since this is what excludes variations on these where the failure is neither non-derivative nor traceable to one non-derivative. But under the stipulated conditions, it seems 'clumsiness' does indeed describe non-derivative failure of ancillary subcapacities.

If these arguments go through, Raz's core-ancillary distinction should be treated as defunct. What remains is the bare idea of a distinction between those of Reason's malfunctionings that are criticisable and those that are not. Such a distinction, I suggest, can still be built with Razian materials, even once we drop his core-ancillary distinction. Call this the distinction between *rational* and *arational* malfunctionings. When the failure of a subcapacity amounting to Reason's

---

[54] Raz himself speaks like the irrationality test, at least as he espouses it, is underwritten by linguistic intuitions about 'irrational,' and may to that extent be optional (2011, 92).

malfunctioning is non-derivative, or is derivative and traceable to a non-derivative failure to another subcapacity, the malfunctioning itself is a rational one. When the failure is derivative but not traceable to a non-derivative failure to another subcapacity, the malfunctioning is an arational one.

To see the appeal of the proposed way of drawing the lines of criticisability, let me say a little more about derivative and non-derivative failures. Whenever the failure amounting to Reason's malfunctioning is non-derivative, the malfunctioning itself is rational *par excellence*. Whenever the failure is derivative, the malfunctioning counts as an *accident*. If the derivative failure traces to a non-derivative failure of another subcapacity, the malfunctioning itself will also be rational, and so counts as *carelessness*—a careless accident. But if the derivative failure does not trace to a non-derivative failure of another subcapacity, the malfunctioning will be arational, and so counts as a *pure accident*. Whenever the failure amounting to malfunctioning is non-derivative, the malfunctioning itself therefore counts as *non-accidental*.[55]

Hopefully with these more familiar terms of practical-philosophical art in the picture, the appeal of my rational-arational distinction becomes clearer. Whether the failure of Reason's subcapacities is non-derivative (as in non-accidents) or derivative but traceable to a non-derivative one (as in carelessness), it amounts to a rational malfunctioning. When Reason's malfunctioning is non-accidental or careless, as is always the case with rational malfunctionings, it makes sense to say the bearer of Reason is at *fault*. A functioning of Reason is a malfunctioning because it falls short of its constitutive rational standard, which is to say, it is unjustifiable. And to non-accidentally or carelessly do something unjustifiable is what it means to be at fault for doing it. The point can be restated from the perspective of fault. People are at fault when they have no adequate justification or excuse for their conduct.[56] If someone does something unjustified, then, necessarily, they have no justification for it. If the same thing was also not a pure accident, which is to say, it was either careless or no accident at all, then they necessarily have no *adequate* excuse for it. All rational malfunctionings satisfy these conditionals, and so rational malfunctioning is sufficient for fault.

---

[55] Non-accidental malfunctionings include paradigmatic cases of intentional wrongdoing and recklessness, although see Douglas Husak (2011) for tricky cases of recklessness at the borderline of non-accidentality.

[56] For this understanding of fault, see Gardner (2005, 112). It also tracks Antony Duff's use of *liability* (2007, 21-2), though as I will suggest, this is a slightly narrow usage of the term.

Non-accidental malfunctionings evince direct disregard for their constitutive rational standards because the underlying subcapacities fall short of them non-derivatively. We have seen, *pace* Raz, how this possibility generalises across all the subcapacities—whether 'core' or 'ancillary.' By contrast, showing malfunctionings to be accidental presumptively indicates a less direct disregard for rational standards, since in accidents the underlying subcapacities fall short derivatively. Minimally, excuses operate by showing malfunctionings to have been accidents. The *ideal* excuse tells a story of how the derivative failure is ultimately not traceable to any non-derivative failure, thereby showing that none of Reason's subcapacities fell short in such a way as to evince even indirect disregard for rational standards. This invites the question as to what lies at the terminus of tracing in the ideal excuse, if not a non-derivative failure. It cannot be a derivative failure, for if it were, then the tracing exercise would not amount to any explanation whatsoever. Rather, in the ideal excuse, tracing terminates in a manifestation of a subcapacity that amounts to a functioning that satisfies its rational standards.[57] It terminates in a notionally separable functioning of Reason that, being non-privative, is justifiable by the very reason it conceptually conditions. In this way, ideally excused malfunctionings evince no disregard (direct or indirect) for their constitutive rational standards. Such malfunctionings are therefore arational, which is to say, on my proposal, that they are pure accidents without fault.

Sadly, not all accidents are as pure as the ideal excuse would desire. Often the derivative failure traces to a non-derivative failure of another subcapacity. In that case, though the subcapacity whose failure amounts to Reason's malfunctioning does not fall short of rational standards, the failure traces to that of another subcapacity which does, so that the malfunctioning evinces indirect disregard. As far as the rational standards are concerned, indirect disregard is disregard nonetheless. A story showing an accident to have been careless at best is therefore inadequate as an excuse. That being said, it is often enough for an excuse's adequacy that it tends towards

---

[57] Some examples of the ideal at work. Say someone who fumbles and smashes a glass is excused because the lights cut out at a crucial moment. This would be because the derivative motor failure traces to a fine manifestation of perception, which rightly responded to the lights cutting out by ceasing to deliver environmental information. Or say someone who goes to the wrong airport is excused because the ticket vendor misinformed them of their flight details. The derivative volitional failure traces to a fine manifestation of cognition, which rightly delivers information there is no reason to distrust into the background of deliberation. Again, say someone who intentionally kills their husband is excused because their husband was, once again, being highly abusive. The derivative volitional failure traces to a fine manifestation of emotion, which rightly responds with explosive grievance to years of provocation (e.g. the appellate decision in *R v Ahluwalia* [1992] 4 All ER 889).

ideality without being so beyond a shadow of a reasonable doubt. In practice, excuses may absolve fault despite an atmosphere of carelessness surrounding the events. But availability of an excuse for some φ (adequate or otherwise) notwithstanding, malfunctionings are *qua* functionings always a matter of responsibility, so long as Reason's bearer possessed a normative reason bearing upon φ that was derivatively a normative reason to try to (omit to) φ.

Our familiar practical-philosophical terms—viz. non-accidents (intentional or reckless), accidents (careless or pure), fault, excuses—should not necessarily be heard as normatively charged. Whether they should depends on the φ-reason with respect to which a given functioning counts as malfunctioning. Suppose someone spills their tea alone in bed. Perhaps it was recklessly done in full knowledge of the risks associated with momentarily balancing their mug on their belly. Or perhaps it was carelessly done without awareness of the risks. Perhaps they mutter excuses to the effect that it was purely accidental. These are perfectly sensible uses of these terms, as I intend them. They are not normatively charged because the reason not to spill tea is, in the given case, not a *categorical* and *mandatory* reason. A reason is categorical if its possession is not sensitive to its possessor's personal goals. It is mandatory if it excludes certain countervailing reasons from entering into the metaphorical battle of strength.[58] *Duties*, or more precisely the fact of their possession, are categorical mandatory reasons.[59] It is when functionings count as malfunctionings relative to duties that our practical-philosophical terms get normatively charged. When it comes to breach of duty, or *wrongdoing* as it is often called, the grounds of fault are the grounds of blameworthiness.[60] Non-accidents (intentional or reckless) are normatively charged by wrongfulness, and carelessness even registers this by being spoken of as *negligence*.

In closing let me offer some programmatic remarks about resultant moral luck and strict liability.[61] My Razian Reasons View is hospitable to moral luck. In moral luck cases what we could

---

[58] In other words, a reason is mandatory because it is 'protected' in Raz's sense (1979, 17-19).

[59] This conception of duty is owed to Gardner and Macklem (2004, 465).

[60] Here 'wrongdoing' does not mean 'doing the wrong thing,' if this means 'acting unjustifiably.' For this distinction, see Gardner (2005, 99-102). Note the following qualification to the claim that blameworthiness is faulty wrongdoing. Some self-regarding reasons are plausibly categorical and mandatory but less plausibly duties 'to oneself,' and even less plausibly 'wrongdoings' when violated. Williams's Gauguin's reason to pursue his art may be like this. If Gauguin faultily fails to conform to this reason he may blame himself, but whether he wrongs himself by violating a duty to himself is less clear.

[61] 'Resultant' moral luck is one of four varieties suggested by Thomas Nagel (1976). Reference to 'moral luck' should hereafter be read as of the resultant variety.

call the quality of a person's will comes apart from the quality of its results, in ways the person cannot control. The unlucky person is often presented in contrast with a lucky other, a foil, for whom the results are qualitatively aligned with their identical will. The unlucky result is typically a malfunctioning relative to a φ-reason that is a duty. Moral luck cases come in two shapes, broadly speaking. In what I call *bad will* cases, both persons have a duty to φ that is derivatively a duty to effectively try to φ, relative to which their Reason is functioning in identically faulty ways (either non-accidentally or negligently). But whilst they both fail to try assiduously enough to constitute effectively trying to φ, only the unlucky one fails to φ. Both wrongfully fail to effectively try to φ, but only one wrongfully fails to φ. Since the lucky one does not fail to φ but merely fails to effectively try to φ, they are, unlike the unlucky one, not responsible for φ (by clause (i) of REASONS CONDITION₃). The unlucky one is responsible and at fault for one more thing, which, being a wrongdoing, makes them blameworthy for one more thing.[62]

In what I call *good will* cases, both persons have a duty to φ that is derivatively a duty to effectively try to φ, relative to which their Reason is functioning in identically faultless ways (neither non-accidentally nor negligently). Both try assiduously enough to constitute effectively trying to φ, but only the unlucky one's effectively trying falls short of φ itself. That is, neither wrongfully fails to effectively try to φ, but one wrongfully fails to φ.[63] One of the unlucky one's functionings, unlike the lucky counterpart, is a malfunctioning. But given absence of fault, the malfunctioning (viz. wrongful failure to φ) is a pure accident. That is, the unlucky one comes out faultlessly responsible for wrongdoing. They are responsible (but not blameworthy) for one more thing, which, being a wrongdoing—i.e. non-conformity with a categorical mandatory reason—supplies rational grounds for an especially recalcitrant kind of agent-regret.[64] Where the wrongdoing is breach of a *directed* duty, agent-regret is in part directed to the wronged, and takes on a dyadic and reparative character.

---

[62] This reflects Raz's explicit take on bad will cases (2011, 252-3). Canonical bad will cases include Nagel's lorry drivers (whether negligent or reckless) and his (would-be) murderers (1979, 141).

[63] Many legal duties to φ anticipate fault, meaning that faultiness is built into the elements of their breach. Where this is so, (apparent) failure to φ will not count as legally wrongful if this comes about faultlessly, as in good will cases. See Gardner for more on fault-anticipating wrongs (2005, 115-6).

[64] Canonical good will cases include Williams's non-negligent lorry driver (1976, 124) and, more contentiously, Neil MacCormick's university teacher who breaks a promise to his kids to respond to a student's urgent mental health needs (1982, 212-3).

Strict liability in law can be grasped on the model of good will cases. If someone incurs 'liability' over some φ, then they are subject to another's normative power over φ.[65] Simplistically speaking, what the normative power is a power to do can be sorted under two headings: either to levy punishment or to compel reparations. Correlatively, liability is either liability to punishment or repair. To call liability 'strict' is to say it can be incurred without fault. This is why strict liability in criminal law is usually considered morally beyond the pale. Commission of wrongdoing only underwrites blameworthiness when faulty, and so to mete out institutionalised blame in the form of legal punishment to faultless 'criminals' is at least *prima facie* artificial. Strict liability in torts is another matter. Claimants in tort seek to obtain from the tortfeasor a civil remedy of a typically reparative character, as perhaps best exemplified by payment of general damages. Good will cases, we have seen, reveal the possibility of faultless responsibility for wrongdoing upon which dyadic agent-regret of reparative character is rationally grounded. So if, not implausibly, tort law institutionalises this reparative moral sentiment, then strict liability in torts—far from being morally incoherent or artificial—reflects ordinary aspects of responsibility in human life.

3.3. Conclusion

The best Reasons View operates an actualist two-way powers account of the conditions of reasons-possession, such that someone is responsible for some thing only if they 'can' do that thing in the sense of having both the two-way rational power and the opportunity to do it. In this Chapter I proposed a properly 'Razian' Reasons View of just this kind. Section 3.1 considered Raz on lack of Reason and lack of opportunity. By sustained reflection on 'disability,' my Razian Reasons View accommodated the notion partial lack as a specific form of lack of opportunity—not to φ *simpliciter*, but to effectively try to φ. Section 3.2 considered Raz on Reason's malfunctionings. First, I offered two arguments against Raz's core-ancillary distinction: from (a) implicit cognitive bias and (b) clumsiness. Next, I proposed to reconstruct a parallel distinction out of Razian materials, and used the reconstructed distinction to introduce some familiar prac-

---

[65] On the correlativity of liability and normative power, see W. N. Hohfeld (1923, 36).

tical-philosophical terms of art. In closing, I indicated how my Razian Reasons View could accommodate resultant moral luck and strict liability. My hope is that the net effect of all this is to shift the best Reasons View's *prima facie* appearance from outlandish to appealing.

CONCLUSION

In this dissertation I defended a highly permissive theory of moral responsibility, which I called the Reasons View, according to which our moral responsibility for things is principally determined by our possessing normative reasons bearing on those things. The view, inspired by Angela Smith (Chapter 1), finds itself at its most theoretically satisfying when supplemented by a powers- or capacity-plus-opportunity condition on reasons-possession (Chapter 2), as exemplified by my development of Joseph Raz's work (Chapter 3). The final formulation of the biconditional central to the Reasons View is this:

> REASONS CONDITION$_3$: S is morally responsible for φ iff (i) S φs, and (ii) S has a φ-reason that is derivatively a reason to effectively try to (omit to) φ.

There are of course many stones left unturned. Let me finish by listing some of what I think are the chunkiest ones. (1) What, explicitly, are the connections between the Reasons View and other existing theories, especially in the reasons-responsive literature? How do they compare? (2) Does accepting the Reasons View imply metanormative commitments? If so, which ones and how? (3) RIC is a necessary condition on reasons-possession. Can it be strengthened to also be sufficient, or, by further conditions, to joint sufficiency? Consider, for example, those reasons that are most obviously *socially* instituted. (4) Much more needs saying about my Razian Reasons View. In particular: (a) How is the normalcy of what opportunities circumstances present to Reason determined? (b) How, if at all, does opportunity (both to φ and to effectively try to φ) relate to value? (c) In more familiar terms from philosophy of mind, how do manifestations of subcapacities relate to their constitutive rational standards? (d) Certain good will cases

of resultant moral luck, aptly called *tragedy*, involve justified breach of duty (including conflict of duty). Does my Razian Reasons View really cover these?

Needless to say, this is just a choice selection of unturned stones. Hopefully in future work I will get a chance to turn at least some of them over. For now, however, I would be happy if what I have said about the Reasons View is enough for it to be taken seriously as a contender in the theory of moral responsibility.

REFERENCES

Alvarez, M. 2013. 'Agency and two-way powers'. *Proceedings of the Aristotelian Society*. 113(1). 101-121.

Anscombe, G.E.M. 1957. *Intention*. Oxford, UK: Basil Blackwell.

Barnes, E. 2014. 'Valuing disability, causing disability'. *Ethics*. 125(1). 88-113.

Belot, G. 1995. 'Determinism and ontology'. *International Studies in the Philosophy of Science*. 9(1). 85-101.

Bricmont, J. 2016. *Making Sense of Quantum Mechanics*. Cham, Switzerland: Springer.

Cartwright, N. 1999. *The Dappled World*. Cambridge, UK: CUP.

Clarke, R. & Reed, T. 'Free will and agential powers'. In D. Shoemaker (ed.). *Oxford Studies in Agency and Responsibility: Volume 3*. Oxford, UK: OUP. 6-33.

Duff, R. A. 2007. *Answering for Crime*. Oxford, UK: Hart.

Fischer, J.M. & Ravizza, M. 1998. *Responsibility and Control*. Cambridge, UK: CUP.

Gardner, J. 2004. 'The wrongdoing that gets results'. *Philosophical Perspectives*. 18(1). 53-88.

——————2005. 'Wrongs and faults'. *Review of Metaphysics*. 59(1). 95-132.

—————— 2007. *Offences and Defences*. Oxford, UK: OUP.

——————2013. 'Reasons and abilities: some preliminaries'. *The American Journal of Jurisprudence*. 58(1). 63-74.

——————2017. 'The negligence standard: political not metaphysical'. *The Modern Law Review*. 80(1). 1-21.

Gardner, J. & Macklem, T. 'Reasons'. In J.L. Coleman, K.E. Himma, S.J. Shapiro (eds.). *The Oxford Handbook of Jurisprudence and Philosophy of Law*. Oxford, UK: OUP. 440-475.

Gendler, T.S. & Hawthorne, J. 2002. 'Introduction: conceivability and possibility'. In T.S. Gendler & J. Hawthorne (eds.). *Conceivability and Possibility*. Oxford, UK: OUP. 1-70.

Heuer, U. 2010. 'Reasons and impossibility'. *Philosophical Studies*. 147(2). 235-246.

——— 2013. 'The reasons that can't be followed: comment on Joseph Raz's *From Normativity to Responsibility*'. *Jerusalem Review of Legal Studies*. 8(1). 185-198.

Hohfeld, W.N. 1923. *Fundamental Legal Conceptions*. New Haven, CT: YUP.

Husak, D. 'Negligence, belief, blame and criminal liability: the special case of forgetting'. *Criminal Law and Philosophy*. 5(2). 199-218.

Ismael, J. 2016. *How Physics Makes Us Free*. Oxford, UK: OUP.

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

Lavin, D. 2004. 'Practical reason and the possibility of error'. *Ethics*. 114(3). 424-457.

————2011. 'Problems of intellectualism: Raz on reason and its objects'. *Jurisprudence*. 2(2). 367-378.

MacCormick, N. 1982. *Legal Right and Social Democracy*. Oxford, UK: OUP.

Macklem, T. & Gardner, J. 2014. 'Human disability'. *King's Law Journal*. 25(1). 60-78.

Molnar, G. 2003. *Powers: A Study in Metaphysics*. Oxford, UK: OUP.

Mumford, S. 1998. *Dispositions*. Oxford, UK: OUP.

Nagel, T. 1976. 'Moral luck'. *Proceedings of the Aristotelian Society, Supplementary Volumes*. Vol. 50. 137-151.

Nelkin, D. *Making Sense of Freedom and Responsibility*. Oxford, UK: OUP.

Raz. J. 1979. *The Authority of Law*. Oxford, UK: OUP.

——1999. *Engaging Reason*. Oxford, UK: OUP.

——2011. *From Normativity to Responsibility*. Oxford, UK: OUP.

Rosen, G. 2015. 'The alethic conception of moral responsibility'. In R. Clarke, M. McKenna & A.M. Smith (eds.). *The Nature of Moral Responsibility*. Oxford, UK: OUP. 65-87.

Ryle, G. 1949. *The Concept of Mind*. London, UK: Barnes and Noble.

Shoemaker, S. 1980. 'Causality and properties'. In P. van Inwagen (ed.). *Time and Cause*. Dordrecht, Holland: D. Riedel. 109-135.

Siegel, S. 2017. 'How is wishful seeing like wishful thinking?'. *Philosophy and Phenomenological Research*. 95(2). 408-435.

Smith, A.M. 2005. 'Responsibility for attitudes: activity and passivity in mental life'. *Ethics*. 115(2). 236-271.

————2008. 'Control, responsibility, and moral assessment'. *Philosophical Studies*. 138(3). 367-392.

————2012. 'Attributability, answerability, and accountability'. *Ethics*. 122(3). 575-589.

————2015. 'Responsibility as answerability. *Inquiry*. 58(2). 99-126.

Streumer, B. 2018. 'Reasons and ability'. In D. Star (ed.). *The Oxford Handbook of Reasons and Normativity*. Oxford, UK: OUP. 233-252.

Teichmann, R. 2008. *The Philosophy of Elizabeth Anscombe*. Oxford, UK: OUP.

Vranas, P.B.M. 2007. 'I ought, therefore I can'. *Philosophical Studies*. 136(2). 167-216.

Wallace, R.J. 1994. *Responsibility and the Moral Sentiments*. London, UK: HUP.

Williams, B.A.O. 1993. *Shame and Necessity*. Berkeley, CA: University of California Press.

————————1976. 'Moral luck'. *Proceedings of the Aristotelian Society, Supplementary Volumes*. Vol. 50. 115-135.

Wolf, S. 1993. *Freedom Within Reason*. Oxford, UK: OUP.

Zimmerman, M. 2002. 'Taking luck seriously'. *The Journal of Philosophy*. 99(11). 553-576.