The Reasoning View

2.1 Varieties of Reasoning Views

According to a popular view about normative reasons, they can be explained in terms of good reasoning. The idea here is, roughly, that if we are going to give a reductionist account of what normative reasons are, we should look into what reasons do. And what they do is, mainly, to figure as premises in reasoning; not any sort of reasoning, though. If they are normative, and hence speak in favour of some F-ing for us, then they have to figure in reasoning that satisfies some standards – in short, reasoning that can, everything else being equal, steer us towards doing the right thing/having the right attitude. In what follows, I propose to explore this view in some detail, look at its most popular versions, rehearse the most popular arguments in its favour, and focus on a number of worries that this view seems to prompt. This section introduces the view in general and some of its more specific versions.

Imagine that the only way for you to enjoy watching the first episode of the new season of your favourite TV show with your friends tonight is for you to arrive at your friend's place by 7:00 p.m. Imagine that your practical deliberation in this situation contains the following steps: you intend to be at your friend's place by 7:00 p.m., you know that the only way for you to be there by 7:00 p.m. is to catch the 6:30 p.m. bus, you intend to catch the 6:30 p.m. bus. Given the background assumptions in place, it seems perfectly natural to describe the consideration that the only way for you to be at your friend's place by 7:00 p.m. is to catch the 6:30 p.m. bus as your reason for intending to catch the 6:30 p.m. bus – that is, to describe it as your reason in the motivating reasons sense. For it is a consideration on the basis of which you intend to catch the 6:30 p.m. bus. You rely on it in your deliberation. Crucially, however, this consideration is not merely your motivating reason in this situation. It is quite natural to see it equally as a normative reason for you to intend to catch the bus. If anything, it is a

38

consideration that, given the relevant background, can make you immune to a reasonable informed criticism for intending to catch the 6:30 p.m. bus. For instance, it would appear inappropriate or unreasonable for a well-informed colleague of yours to criticise you for intending to catch the 6:30 p.m. bus. Insofar as she knows that you intend to be at your friend's place by 7:00 p.m. and that the only way for you to be there on time is to catch the 6:30 p.m. bus, her criticism of your intention to catch the 6:30 p.m. bus would appear groundless. For instance, she cannot reasonably claim that your intention is baseless. Moreover, one might think that there is something more than mere immunity to reasonable criticism. One might think that your intention to catch the 6:30 p.m. bus is perfectly rational and that it is rational in virtue of it being based on the consideration about the 6:30 p.m. bus being the only way to be on time (plus the relevant background details). Furthermore, perhaps, this property of rationality of your intention makes it the case that intending so is also what you ought to do (at least in the so-called deliberative sense of 'ought'). In sum, this consideration about the 6:30 p.m. bus being the only way to arrive on time at your friend's place is not merely a basis for your intention; it has some normative force and import for you, however exactly we may spell it out. Whatever else we can say about the situation, it seems difficult to deny that there is a normative reason for you to intend to catch the 6:30 p.m. bus and that that reason is exactly the same consideration that figures crucially in your practical deliberation, a consideration on the basis of which you intend; in other terms, it is also your motivating reason.

If one accepts the apparently natural idea that in a number of ordinary cases, like the bus case, a normative reason corresponds to one's motivating reason and, in particular, to a consideration that plays a role in one's deliberation, one may be tempted by a very natural, or so it seems, further claim – namely, the view that normative reasons are just a subset of motivating reasons or, at any rate, a subset of considerations that can play a role in deliberation. It is a naturally tempting view given its incredible simplicity (no need for genuinely different sorts of reasons), its naturalistic flavour, and its straightforward explanation of the link that many think of as a crucial constraint on any theory of reasons, namely the link between motivating and normative reasons (more on this ahead).

The Reasoning view of reasons takes this line of thought seriously and attempts to work it out by filling in the details and exploring its implications. We can capture this idea in its most general form with the following schema:

> A consideration r (on many accounts, a fact) is a normative reason for S to F just in case r is a content of a premise-response (along with other possible premise-responses) in S's possible good/sound reasoning towards F-ing

(i.e. conclusion-response). (cf. Williams 1979, 1989, 2001; Velleman 1999; Hieronymi 2005, 2013 (on some interpretations); Setiya 2007, 2014; McHugh and Way 2016; Silverstein 2016; Asarnow 2017; Way 2017)

We can then introduce the Reasoning view as a proposal that takes the apparently intuitive claim that normative reasons have to be available as motivating reasons, and then provides us with a recipe for how exactly to specify which subset of motivating reasons (or possible motivating reasons/considerations that play a role in one's deliberation) are normative reasons. More specifically, according to the Reasoning view, normative reasons are those potential or actual motivating reasons that correspond to appropriate premises in good patterns of reasoning that S can undertake to F-ing (given the relevant background). In our case, the fact that catching the 6.30 p.m. bus is the only way for you to make it on time to your friend's place corresponds to a premise in a pattern of good/sound reasoning from this premise (and the relevant other premises) to the F-ing (or, intending to F). In sum, there are two elements in the Reasoning view that are appealed to in order to define normative reasons: reasoning and soundness/goodness/norms of reasoning. The former is a descriptive element, the latter a normative element. Thus the general thought underlying the Reasoning approach can be captured as follows: 'If reasons in general are considerations that figure in reasoning, normative reasons are considerations that figure in sound reasoning' (Silverstein 2016: 2).

Now, there are a number of versions of the Reasoning view, some of them more elaborated than others. Typically, the 'first generation' modern versions of the Reasoning view are programmatic and have a number of implicit assumptions and underdeveloped aspects. I suggest classifying Williams (1979, 1989, 2001), Raz (1999), Harman (1986), Velleman (1999), and Grice (2001) as instances of this 'first generation Reasoning view'. Proposals in Hieronymi (2005) and Alvarez (2010: 42) come very close to the general idea of the Reasoning view but may be open to a different interpretation, though. See Section 2.2 for an in-depth overview of the literature and exegetical comments.

The situation is different with respect to more recent variants of the Reasoning approach or what we may call a 'second generation' of Reasoning views. Recently some authors have gone into considerable detail in elaborating the view, considering the problems for some of its versions, and proposing positive arguments in its favour. Among the main recent ('second generation') defences of the Reasoning view are Setiya (2007, 2014), Asarnow (2017), Silverstein (2016), Way (2017), and McHugh and Way (2016). If we assume a relaxed sense of 'reasoning' then,

arguably, Gregory (2016), which claims that normative reasons are good bases, would also count as a version of the Reasoning view. Let me conclude this introductory section by merely putting on the table some of their views without going much into the details of their proposals. We will return to some of the relevant details of their proposals throughout this chapter when we examine the pros and cons of the Reasoning view more systematically.

In his 2007 book, Kieran Setiya proposes the following version of the Reasoning view (which he develops further in Setiya 2014):

> Reasons: The fact that p is a reason for A to φ just in case A has a collection of psychological states, C, such that the disposition to be moved to φ by C-and-the-belief-that-p is a good disposition of practical thought, and C contains no false beliefs. (Setiya 2007: 12)

A central element of his approach is to specify the goodness/soundness aspect of good reasoning in terms of good dispositions of practical thought (the focus is more specifically on reasons for action).

A different way of precisifying the exact nature of the goodness/soundness aspect in the general Reasoning approach has been recently undertaken by Conor McHugh and Jonathan Way (see, in particular, McHugh and Way 2016; see also Way 2017 for relevant details). According to them, the goodness/soundness aspect is explained in terms of fittingness (or correctness, appropriateness, rightness):

> For that p to be a reason for a response is for that p to be a premise of a good pattern of reasoning from fitting responses to that response. (McHugh and Way 2016: 586; compare to Way 2017: 254)

On their view, reasoning is understood quite broadly, to include any passage from some mental attitude (premise-response, in their terminology) to another attitude or action (conclusion-response) where the latter is held on the basis of the former.[1] On their view, good reasoning will be defined, roughly, in terms of fittingness preservation. And the qualification of 'from fitting[/correct] responses' in their account is supposed to appeal to a general requirement covering not only the requirement of having (or potential having) true beliefs but also fitting non-doxastic states (including fitting/appropriate intentions, perceptual states, and others). This

---

[1] 'Here, reasoning is understood broadly, as a certain kind of transition in which a set of responses, which we can call premise responses, leads to some (further) response, which we call the conclusion response. This transition is such that the conclusion response counts as based on, or held in the light of, the premise responses' (McHugh and Way 2016: 586).

constraint ensures that considerations towards clearly and radically immoral acts are not recognised by the Reasoning view as normative reasons. Consider, say, a villain who strongly desires to terminate human life on Earth. Assume that the only way to terminate any human life on Earth is by initiating a global nuclear war. Now, there is a pattern of reasoning that contains the villain's desire and his true belief about the nuclear war being the only way to terminate human life on Earth as premise-responses and initiation of a global nuclear war as the conclusion-response. This pattern of reasoning is in a sense 'good'; it is valid in the intuitive sense introduced earlier: if premises are appropriate/ fitting, so is the conclusion. But that the way to terminate human life on Earth is by initiating a global nuclear war is clearly not a reason for anyone to initiate a global nuclear war. That it is not a normative reason is ensured by the soundness condition (not the validity condition), and this sound- ness condition is explicated in McHugh and Way's account in terms of having fitting premise-responses (all the actual or potential premise- responses have to be fitting). The particularity of their approach is that they take fittingness to be fundamental and a prime normative property that they don't define in further terms (another defence of the fittingness first approach is Chappell 2012). Silverstein (2016) assumes Way's (2017) version of the Reasoning view and provides a further defence of the view.

Asarnow specifies his version of the Reasoning view by appeal to norms of practical reasoning together with a soundness condition (incorporating an anti-defeat condition) on possible premises of the patterns of reasoning. What exactly are the norms of practical reasoning is left undefined in his account. However, he does point to some uncontroversial examples of such norms – for example, the *modus ponens* rule. His recent statement of the view is as follows:

> REASONING VIEW* A normative reason for A to ɸ is a set of facts, F, such that the norms of practical reasoning endorse the transition from a set including beliefs with those facts as their contents and (optionally) one or more elements of A's practical standpoint, to A's intention that A ɸ, and there are no defeaters for that transition. (Asarnow 2017: 626)[2]

It may be useful to note that Asarnow separates the goodness condition of reasoning from what he calls the 'soundness condition'. Roughly, the soundness condition ensures that only true beliefs (or more generally states

---

[2] A version of the Reasoning view that is more specifically attentive to the possibility of non-belief states providing reasons is proposed by Asarnow (2016: 174), in the following terms: 'RV NORMATIVE REASONS. A normative reason for A to ɸ is a set of facts, F, such that the norms of

corresponding to facts) can be reasons. Asarnow's formulation doesn't include an appeal to the fittingness of [potential] premise-responses, contrary to McHugh and Way's formulation. How then does his view block considerations towards immoral acts from counting as reasons in cases where these considerations can play a premise role in valid reasoning? The trick is accomplished here by the anti-defeat clause and an assumption that strict moral requirements can play the role of defeaters: 'While the norms of reasoning endorse the transition from Caligula's desire to have pleasure and his belief that harming innocents will bring him pleasure to the intention that he perform that violent act, the fact that the violent act is morally forbidden is a defeater for that piece of reasoning' (Asarnow 2016: 628). Note that moral requirements are not the only thing that can play the defeater role in Asarnow's theory. He thinks that another category of potential defeaters comes from 'an agent's especially strong or especially deeply held volitional commitments' (Asarnow 2017: 627).

Let me stress that a common feature of these views is that they consider a pattern of good reasoning/disposition of [practical] thought as a sort of abstract entity (cf. Asarnow 2017: 616). A subject is not required to have *all* the relevant premise-responses (to use McHugh and Way's terminology), let alone actually undergo a concrete piece of reasoning in order for there to be a reason for her. All that is required is only that there is *a* good pattern of reasoning/disposition of [practical] thought and that the agent has some of the relevant states (that constitute premise-responses). It is possible that *r* is a reason for a subject even if the subject doesn't believe that *r*. It is only required that there is a possible reasoning from a possible belief that *r* and some other premise-states to the relevant conclusion-response.

With all these views on the table, let us examine whether we have good arguments for adopting one or another version of the Reasoning approach. Before that, I propose a brief exegetical historical overview of the most influential variants of the 'first generation' of Reasoning views. The section can be skipped without losing anything of substance from the overall argument.

2.2 A Fuller Exegetical Overview of Reasoning Views

The most prominent proponent of the Reasoning view was probably Bernard Williams (cf. Williams 1979, 1989, 2001). His version of the

practical reasoning endorse the transition from a set of possible mental states, M, the elements of which are appropriately related to the elements of F, to A's intention to ϕ'.

view is somewhat implicit, and it is not entirely clear how exactly his view is supposed to go, but it is closely tied to his famous internalism about reasons, a version of Humeanism according to which reasons are defined partly by appeal to one's motivational set – that is, 'the set of [one's] desires, evaluations, attitudes, projects, and so on' (Williams 1989: 35). He takes reasons to be fundamentally motives. On his account, when '*A* has a reason to *ϕ*' is true '*A* has some motive which will be served or furthered by his *ϕ*-ing' (Williams 1979: 101). He maintains that any conception of reasons has to respect the constraint according to which we should be able to act on the basis of our normative reasons in order for them to have any normative force. The closest we get to a definition of reasons in Williams is:

> (1) A has a reason to *ϕ* only if there is a sound deliberative route from A's subjective motivational set [. . .] to A's *ϕ*-ing. (Williams 2001: 91; compare to Williams 1979, 1989)

He is not clear on how exactly we should understand the soundness element in his view of reasons. Actually, he seems to endorse this vagueness and hints towards the idea that the vagueness of his account is a point in its favour. For example: 'It is sometimes held against the combination of the internalist view with this broad conception of deliberation that it leaves us with a vague concept of what an agent has a reason to do. But this is not a disadvantage of the position. It *is* often vague what one has a reason to do' (Williams 1989: 38, original emphasis). He also doesn't seem to conceive of patterns of reasoning/deliberative routes as always corresponding to well-known patterns of reasoning. He maintains that imagining can constitute a sound route of reasoning (Williams 1989: 38), which raises the question of how even to think about the standards of sound reasoning.

Another prominent proponent of the Reasoning view (at least in its general form) is, arguably, Joseph Raz. For some passages in Joseph Raz's work – in particular, in his earlier work – may be interpreted as expressing a version of the Reasoning view. Consider, for instance, the following passage:

> (2) Statements of facts which are reasons for the performance of a certain action by a certain agent are the premises of an argument the conclusion of which is that there is reason for the agent to perform the action or that he ought to do it. (Raz 1999: 28, second edition of *Practical Reason and Norms*)

Under the assumption that 'argument' here corresponds to patterns of reasoning, this quotation seems to be in the spirit of the Reasoning view. One may, however, be suspicious about this interpretation, given the absence of the 'soundness' element in this quotation. However, it is not

clear that there is no such implicit assumption about the soundness or goodness requirement of the argument in Raz's formulation (see also Raz 1978, *Practical Reasoning*, introduction, where he appears to be more explicit about the truth condition of premises). That is, it is not clear that Raz thinks of normative reasons in mere terms of inferences without any references to good or sound inferences. Raz does mention the validity aspect of a practical inference and holds that it is an important aspect. Once this validity aspect of inferences is recognised, it is only a small step from there to recognition that some inferences are sound. Soundness is understood in the usual way as validity plus truth of premises in the case of theoretical inferences and, perhaps, appropriateness/fittingness plus satisfactoriness in the case of practical inferences (see Kenny 1966). Indeed Raz thinks that the appeal to valid inferences is an advantage of his account. For instance, he writes:

> (3) One welcome result of this approach is that practical inferences are defeasible, that is, the addition of further premises can turn a valid argument into an invalid one. (Raz 1999: 29)

And Raz does maintain that he understands practical inferences as inferences that conform to the logic of satisfactoriness as it appears, for instance, in Kenny (1966) (where, roughly, satisfactoriness is to practical reasoning what validity is to theoretical reasoning; cf. Raz 1999: 207, fn 7). Hence, it is not that unreasonable to classify at least early Raz as a proponent of the Reasoning view.

Something like the Reasoning view also seems to be accepted by Gilbert Harman in his *Change in view* (1986). For instance:

> (4) To say that a consideration C is a reason to do D is, I suggest, to say that C is a consideration that has or ought to have some influence on reasoning, leading to a decision to do D unless this is overruled by other considerations. The consideration C might be an end or a belief one has, or it might be some line of thought which one finds or would find attractive or persuasive on reflection, for example, an argument of some sort. (Harman 1986: 129–130)

An appeal to the idea that reasons have to be connected to reasoning is also present in Grice's work. For instance:

> (5) Reasons (justificatory) are the stuff of which reasoning is made, and reasoning may be required to arrive (in some cases) even at the simplest of reasons; so it seemed proper to proceed from a consideration of reasoning to a consideration of reasons. (Grice 2001: 67)

'Justificatory' reasons seem to correspond to what we refer to as 'normative' reasons.

The Reasoning view or something quite close to it is also endorsed by David Velleman (1999). His version of the Reasoning view is to be understood in accordance with his specific account of practical reasoning and within his wider philosophical project, which we are, unfortunately, in no position to fully rehearse here. According to Velleman:

> (6) The reasons for an action are things represented in premises from which intending or performing the action would follow as a conclusion in accordance with practical reasoning. (Velleman 1999: 198)

This following as a conclusion is understood as '[enhancing] the agent's self-knowledge by satisfying some self-conception' (cf. Velleman 1999: 198). This, then, provides the basis for the official definition of reasons on Velleman's account as follows:

> (7) [R]easons for an action are those things belief in which, on the agent's part, would put him in a position to enhance his self-knowledge, in this distinctively practical way, by intending or performing that action. (Velleman 1999: 198, original emphasis removed)

One particularity of Velleman's account is that, contrary to much of the current orthodoxy, he allows for false beliefs to be reasons (cf. Velleman 1999: 200). However, he qualifies this contention by insisting that in the case of false beliefs, their falsity speaks against them and against acting on them: 'their falsity is a reason for abandoning them and hence also a reason against acting on them in their capacity as reasons' (Velleman 1999: 200, fn 9). Presumably this move and the link between reasons and an agent's enhancing her self-knowledge in practical reasoning is enough to satisfy the 'soundness/goodness' aspect within the general form of the Reasoning view according to which reasons are premises in good/sound reasoning. However, a more complete exegesis of Velleman's view would be necessary to establish this classification of his view within the camp of the Reasoning view with any degree of certainty.

It is sometimes claimed that Pamela Hieronymi is another prominent proponent of the Reasoning view (for instance, in Silverstein 2016; Way 2017; Whiting 2018). However, it is not entirely clear that Hieronymi subscribes to the Reasoning view as it has been introduced here. Certainly, some passages may be reasonably interpreted as committing Hieronymi to something close enough to the Reasoning view. For instance: 'To start reflection, we can note that, most generally, a reason is simply an item in a piece of reasoning' (Hieronymi 2005: 443). However, it is also clear that for Hieronymi a reason is a consideration that bears on a question rather than on F-ing itself. For instance: 'So I suggest, for consideration, the

following account of a reason: a reason is a consideration that bears on a question' (Hieronymi 2005: 443). And: 'This account differs from the original formulation in taking the fundamental relation in which a consideration becomes a reason to be a relation to a question, rather than to an action or attitude' (Hieronymi 2005: 443). Nevertheless, it is not unreasonable to see Hieronymi as a proponent of a version of the Reasoning view, in particular given her comment on her view, where she doesn't seem to make a difference between bearing on a question and bearing on a conclusion: 'One could say, "a consideration that bears on a conclusion." I do not think there would be any relevant difference, though I find the idea of answering a question more intuitive for capturing the activities of rational agents' (Hieronymi 2005: 443). For now, we can tentatively classify Hieronymi's view as a version of the Reasoning view. But we will come back to Hieronymi's view shortly, when elaborating our positive proposal, since, as we will observe, it contains another crucial insight for our view beyond its focus on reasoning.

One might also think that something close to a version of the Reasoning view is presupposed in Maria Alvarez's argument in favour of propositionality and factivity of all reasons (though she doesn't appeal to validity or soundness):

> (8) A better reason for arguing that the most perspicuous way of expressing reasons is propositionally is that reasons must be capable of being premises, i.e. things we reason, or draw conclusions, from, whether in theoretical or in practical reasoning. Otherwise, the connection between reasons and reasoning would be lost. (Alvarez 2010: 42)[3]

However, we lack sufficient grounds for ascribing a full-blooded endorsement of the Reasoning view to Alvarez.

If it can be reasonably held that (most of) the abovementioned views are versions of the Reasoning view (e.g. the 'first generation Reasoning view'), it can nevertheless also be recognised that they are not really elaborated in detail. They appeal to the general idea of explaining normative reasons in terms of some norms or value of reasoning (e.g. sound/good patterns), but they don't spend much time working out the details of how exactly the view works. Also, they are a bit shy on giving positive arguments in favour of the view. Rather, the impression is that they are happy with putting this view on the table and working out its implications for their further

---

[3] See, for the factivity part: 'My view is that all facts are indeed reasons merely by virtue of being potential premises in (theoretical or practical) reasoning' (Alvarez 2010: 42).

theoretical views, which *can* be taken as a case in favour of the Reasoning view, if successful. Moreover, they don't appear to spend much time considering possible shortcomings of the Reasoning view and how to address them.

2.3   Arguments in Favour of the Reasoning View

Given the state of the debate about the Reasoning view, it is not always obvious how to extract positive arguments in its favour. On some occasions (in particular, in the context of what I have called the 'first generation' of the Reasoning view; see Section 2.2), the Reasoning view appears to be merely endorsed, taken as obvious, and considered to be in no need of further theoretical defence. Nevertheless, a number of more elaborated positive lines of thought can also be found (in particular, within the context of the 'second generation' Reasoning view). As far as I can see, there are five general lines of argument in favour of the Reasoning approach within contemporary debates. (Some of these are so closely related that it would not be unreasonable, however, to lump them together.) Let us look at them briefly.

First, as Jonathan Way (2017) observes, it is a valuable feature of any view of normative reasons if it is general enough to be applicable to normative reasons of *all* varieties. Arguably, there are normative reasons not only to act or intend but also for attitudes – for example, beliefs, emotions. The Reasoning view is perfectly adapted to account for this and hence to have the necessary level of generality. As long as there is a good pattern of reasoning (understood broadly enough to count emotional attitudes – for example, fear, anger, pride, as conclusion-responses) towards the relevant conclusion-response (and the other relevant conditions obtain), there can be normative reasons for the conclusion-response in question. Thus, the fact that the Reasoning view can deal with reasons of various sorts speaks in its favour.

Second, the Reasoning view not only seems to correctly predict and explain what things can have normative reasons but also accurately predicts for what kinds of things there cannot be reasons. In this manner, for instance, the Reasoning view can explain why there are no reasons for values. This point is again observed by Way (2017), who suggests that given the Reasoning view we can easily explain why there are no reasons for, say, having green hair, being tall, being healthy (as distinct from eating healthy food), and perceiving so and so. These are things towards which there cannot be any reasoning. Hence, there cannot be good patterns of

reasoning towards having green hair, being healthy, perceiving red, and so on. The Reasoning view has the right degree of discrimination: it excludes precisely the things we intuitively don't want to count as being reasons-sensitive.

Third, as many proponents of the Reasoning view observe, it is also well placed to explain the pre-theoretical thought that reasons to F have to be somehow connected to reasoning and to reasons for F-ing (motivating reasons). Indeed, many think that it is a platitude that reasons are what reasoning is made of (cf. Grice 2001: 67). The Reasoning view has a straightforward explanation for this. Other views arguably struggle to explain it in simple terms (this observation is strongly connected to the next point). Now, if one takes on board a further somewhat natural assumption that all motivating reasons (reasons *for* which we F) are pre-mises in our reasoning (assuming again that reasoning is broadly construed to include all kinds of relevant transitions towards F-ings), the Reasoning view has again a simple and powerful story about how and why normative reasons are connected to motivating reasons. Reasons, normative or moti-vating, just are premises in patterns of reasoning. And in the normative case, they are appropriate premises in good/sound patterns of reasoning.

Fourth, and probably the most popular line of argument in favour of the Reasoning view, is a comparative argument. Strictly speaking, the observed points can be also understood as boiling down to comparison to other views. Hence, the lines between these five points are not really strict; it is more about dialectical accents and framing. Let me give three examples here that illustrate the comparative argument.

Pamela Hieronymi (2005), for instance, thinks that the Reasoning view (or at least something similar to it; see Section 2.2) is clearly better off than the reasons-first approach with respect to the so-called wrong kind of reasons problem (see Section 1.4). In fact, according to Hieronymi, there is not really a problem for the Reasoning view. Roughly put, proponents of the reasons-first approach cannot satisfactorily explain why, say, a demon's threat is not a normative reason for the demon's victim to admire it, despite clearly counting in favour of one admiring it, given the demon's threat to punish one severely for non-compliance with the order to admire. Proponents of the Reasoning view have tools to explain this sort of case: the threat from the demon is not a reason to admire, since it cannot be a premise in a good piece of reasoning to admiration, but we can still account for the counting in favour intuition in such cases, since the threat can figure in a premise of good reasoning not towards admiration but, say, towards attempting or wanting to admire the demon. According to this

line of thought, the Reasoning view is to be preferred to the reasons-first approach (which is not, of course, to say that it is winning against all of its possible alternatives). Note also that the objection against the reasons-first approach according to which their distinction between the 'wrong' and 'right kind' of reasons to admire is ad hoc or self-contradictory doesn't apply to the Reasoning view (see Section 1.4). The Reasoning view is not presupposing that reasons are prime or fundamental; it does provide a theory of reasons and hence can legitimately propose substantial distinctions among kinds of reasons, contrary to views that consider reasons as undefinable.

For Kieran Setiya (2014), the best arguments in favour of the Reasoning view rely on its comparison to other views. One line of thought here is that if we assume that reasons are somehow connected to rationality, a theory of reasons has to explain that link. According to Setiya, the Reasoning view does better on this account than its competitors (by connecting rationality to good reasoning dispositions and by defining reasons in terms of good patterns of reasoning). However, notice that similar to the abovementioned point, Setiya also gives at best only the beginning of a full positive argument here. He compares the Reasoning view only to some of its competitors on this topic. He shows that views that connect reasons to the way a rational or virtuous (exemplary) person would be moved to act in a situation have shortcomings in cases of non-virtuous agents who still can have normative reasons to act in certain ways in which virtuous agents would not be moved to act. And Setiya does suggest that a possible improvement on that view – namely, the ideal adviser model, where the focus is on an idealised and fully rational version of oneself as an adviser for the current situation in which one finds oneself (cf. Smith 1994, 1995), is still unsatisfactory compared to the Reasoning view. However, one might worry that these views, which seem to lack the appropriate degree of abstraction from actual agents to capture the nature of normative reasons, are not the only competitors with respect to the explanation of the connection between reasons and rationality. For instance, some reasons-first approaches might claim to be able to account for the reasons–rationality connection by appeal to the possession condition of reasons or the perspective dependence of reasons (for a recent version of this strategy, see Lord 2017 and Kiesewetter 2017, 2018). Arguably these other alternatives will not have the same problems as the abovementioned views (e.g. the 'example' and the 'ideal adviser' models). More work is probably needed in order for this line of thought to gain real traction against all the Reasoning view's competitors.

According to Samuel Asarnow (cf. Asarnow 2016, 2017), we have to accept Rational Internalism, the view that connects normative reasons to motivating reasons (e.g. normative reasons have to be able to be motivating reasons) on the basis of a roughly Davidsonian idea of rationalisation, rather than on the basis of Setiya's (2007, 2014) internal dispositionalism (see also Gibbons 2013). According to Asarnow, Rational Internalism provides an argument against Objectivism about normative reasons (cf. Broome 2013). Now, the consideration that speaks in favour of the Reasoning view, according to Asarnow, is that it provides an attractive alternative to those who are sympathetic to Objectivism, since it allows for objective values (where the existence of objective values was, according to Asarnow, a central motivation in favour of Objectivism about reasons). Whatever the merits of this sophisticated argument, it is, again, a comparative argument that relies on some substantive assumptions but also suggests that accepting the Reasoning view brings in some unexpected advantages over Objectivism about reasons (e.g. reasons as facts about oughts) and over hardcore internalism (e.g. reasons as one's psychological states).

The fifth line of argument explores the idea that the Reasoning view enables a simple reductive story where normative is reduced to descriptive. One version of such an argument appears in Silverstein (2016). A central task of that paper is to elaborate in detail an account of soundness or goodness of reasoning. A central assumption there is that reasoning has to have not only a merely formal aim but also a substantive aim and that such a substantive aim can be put in descriptive terms. If he is right, giving grounds for such a reductionist approach is another advantage of the Reasoning view, compared to other alternatives that sometimes struggle to provide a clear and straightforward story on this account. But again, the argument has a limited scope, since it is not absolutely evident that all the possible alternatives will be ruled out here.

In this section, we have looked at five existing lines of defence of the Reasoning view. The rest of this chapter is concerned with its potential shortcomings.

2.4 First Set of Worries I: Outweighing and Weight

As we have seen earlier, the Reasoning view enjoys some initial plausibility. However, on reflection, it has also some puzzling aspects. Let us start with worries that are already well-known (in Sections 2.4 and 2.5 in particular) before turning to some new problems (in Sections 2.6).

There are, as far as I can see, two general lines of prima facie problems for the Reasoning view within the literature. The first one arises from the observation that not all acts or attitudes that are recommended by reasons always correspond to outputs of a good pattern of reasoning. Some reasons are outweighed and yet remain normative. This is the problem of so-called outweighed reasons. The second worry arises from the observation that good patterns of reasoning can contain among their premises statements of mere enabling conditions for F-ing. And yet it doesn't feel always right to consider mere enabling conditions as genuine normative reasons to F. Let us look at the details of these worries a bit more attentively and review some of the most influential existing responses to these worries. This section is devoted to the former worry (outweighed reasons), while the next one focuses on the latter worry (enabling conditions).

To begin with, let us first clarify certain aspects of the Reasoning view a bit more. The general idea of the Reasoning view, as we saw earlier, is that there are good/sound patterns of reasoning (which presumably mirror patterns of good/sound arguments) and normative reasons are premises of such patterns. Let us start with the general version of the view. Recall:

> **The Reasoning view (general, rough)** A consideration *r* (on many accounts, a fact) is a normative reason for *S* to F just in case *r* is a content of a premise-response (along with other possible premise-responses) in S's possible good/sound reasoning towards F-ing (i.e. conclusion-response).

How should we understand what is meant by 'good reasoning' here? Variations of the view exist in this respect. However, on a very general level of abstraction, everyone agrees that reasoning is, roughly, a transition from some mental states to others, where the arrival state (i.e. the state at which one arrives through such a transition) is held on the basis of the initial state/s. That is, it is not a *mere* transition, where one happens to transit, say, randomly or in a purely mechanical way, from one state to another. There has to be some more substantive link between the arrival state and the initial state. And this required link can be, for the time being, described roughly as 'basing' – holding one state on the basis of or in virtue of the other one.

Reasoning so conceived can be evaluated. Some such transitions are appropriate, while others are clearly not. Jumping to a conclusion via purely fearful or wishful reasoning is not appropriate. Say, jumping to the conclusion that you will be able to meet a short deadline for submitting a new project, when you know that there is a massive past record of your failure to meet deadlines that speaks against you having such an ability, is

inappropriate. The fact that you very much want to meet the deadline doesn't make your reasoning any better in such a situation. On the other hand, the reasoning that originates in your intention to have a party at your place together with your belief that the only way to have that party is to invite some friends, and that terminates in your intention to invite/ invitation of some friends, appears to be an appropriate piece of reasoning.

Pieces of reasoning correspond to patterns of reasoning (we can think of them as abstract entity). Some patterns are good/sound, while others are not. Say, a pattern of reasoning that mimics a *modus ponens* argumentative structure and has the appropriate initial states is clearly good. And a pattern of reasoning that mimics the Affirming the Consequent argumentative structure is clearly not good. We will return to this in a moment. For now, let us only focus on the claim to which all proponents of the Reasoning view are committed, namely that normative reasons are premises (or contents of initial mental states/premise-responses) in good patterns of reasoning. So, on the Reasoning view, if r is a normative reason for a subject S to F, then there has to be some good pattern of reasoning for S from some premise-responses (initial mental states), where one of these premise-responses has r as a content towards F-ing/intending to F. S is not required to have all the relevant premise-responses (for instance, S may not even believe that r), but S has to have at least some of the relevant premise-responses. Maybe I only have the intention to have a party but haven't yet realised that the only way to have a party is to invite friends. The Reasoning view predicts, plausibly, that the fact that the only way for me to organise the party is to invite some friends is a reason for me to (intend to) invite some friends. In this case, all that is required is that there is a good pattern of reasoning along the following lines: intention to G, belief that the only way to G is to F, (intention to) F. Even if I do not undergo any concrete piece of reasoning, the mere fact that there is such a pattern of good reasoning from some of my mental states (premise-responses) to the relevant F-ing is enough for r to be a normative reason for me to F in case where r is a content of one of these actual or some merely possible premise-responses of mine.

Now, the problem of outweighed reasons arises from the simple observation that a pattern of reasoning is either good or bad; that is, it either complies or doesn't comply with the standards of reasoning/argument. And a consideration either is or is not a content of the premise-response (possible or actual mental state) of a good pattern of reasoning. However, given standard assumptions, some patterns of reasoning will be classified as bad even though we might have a strong inclination to see one or more of

the considerations in the premise-responses as normative reasons. To make the objection a bit more concrete, consider the following version of a classical example (the original example goes back to Ross 1930: 18). You have promised to meet your friend for a coffee today. On your way to the coffee shop, you witness a traffic accident. As it happens, you are the only witness and some of the people involved in the accident are severely injured. Suppose you undergo reasoning (indeed, it may be almost instantaneous) that concludes in (an intention to) call the ambulance and help the injured. Now, you did the right thing, no doubt about that. Everyone agrees that there is no all-things-considered ought for you to leave the injured and go to see your friend instead. There is no *sufficient* reason for you to go on to your meeting with your friend. However, it is nevertheless natural to think that there is still *a* reason for you to go to see your friend. That you promised to meet your friend over a coffee is still something that counts in an intuitive sense in favour of going to meet your friend. This promise still exercises some normative force upon you. Of course, everyone agrees that it is a massively outweighed reason in the circumstances of the accident, but it seems to be a reason nonetheless. One popular way to further motivate this observation is to appeal to the fact that if it were not a reason for you to go to meet your friend, it would be difficult to explain the fact that it is appropriate for your friend to be somewhat annoyed with you for not showing up. It seems that the sensible thing for you to do would be to excuse yourself later or at least to explain him why you didn't show up (we are assuming here that helping the injured in the accident necessarily entails that your will miss your coffee break). Moreover, if your promise (and additional considerations, such as the desire to meet your friend) is not a reason to go on, then we have a hard time explaining why it appears OK for you to feel slight regret at missing the coffee break.

The problem of outweighed reasons is that the Reasoning view, in its simple form (given some further standard background assumptions), entails that there is no normative reason for you to go to meet your friend in the circumstances of the accident. That is, there is not the *slightest* reason, according to the general Reasoning view, not even a tiny one. This is so because there is no good pattern of reasoning for you in this case (given your beliefs, desires, intentions, and other states) towards going to meet your friend. More precisely, there is no good pattern of reasoning given some standard assumptions about good patterns of reasoning. The pattern that corresponds to our accident case is, roughly, the following: intention to respect a promise to meet a friend for a coffee, belief that the only way of respecting the promise is to go to the coffee shop (the meeting

place), intention to help the heavily injured persons in an accident, belief that the only way to help them is to stay and call the ambulance (which entails not going to the coffee shop), intention to go to the coffee shop. The presence of the intention to help and the belief that the only way to help is to stay and call the ambulance is what 'makes' the pattern of reasoning bad, according to this line of objection. These elements undermine, so to speak, the reasoning from the initial premise-responses of intention to respect the promise and belief that the only way to respect the promise is to go to the coffee shop, to the conclusion-response of going to the coffee shop. These additional elements (intention to help the injured and the belief that the only way to help is to stay and call the ambulance) function as defeaters of the goodness of reasoning. We can represent the pattern of reasoning in this case more schematically as a transition from initial states (premise-responses) intention to F, belief that P-ing is the only way to F, intention to G, belief that Z-ing (entailing not-P-ing) is the only way to G, to the conclusion-response of P-ing. The fact that the agent has the intention to G and the belief that Z-ing is the only way to G *in addition* to the intention to F and belief that P-ing is the only way to F is what makes arriving at the conclusion-response of P-ing because of the given premise-responses inappropriate (given the assumption that good patterns of reasoning cannot be undermined by additional defeating elements). Hence, the Reasoning view in its general form (plus standard assumptions) appears to predict that the consideration that the only way for you to respect your promise is to go to the coffee shop cannot be a normative reason for you to go to the coffee shop. There seems to be no good pattern of reasoning for you that has this consideration as a content of one of your premise-responses to the conclusion-response of going to the coffee shop. Being a premise of a good pattern of reasoning is categorical. There are no degrees of being a premise of a good pattern of reasoning. Thus, the Reasoning view appears to predict a counterintuitive conclusion. For, as we observed above, that the only way for you to respect your promise is to go to the coffee shop is a normative reason for you to go to the coffee shop. It is a reason, even if it is clearly not a sufficient reason. And this presents a problem that cannot be easily dismissed, since our ordinary lives are over-packed with all sorts of outweighed reasons. Arguably, every situation where one faces a non-trivial choice (and that is not a situation of a genuine dilemma) is a situation with at least one outweighed reason. A view that predicts that there cannot be outweighed reasons is problematic.

I follow Way (2017) in classifying the possible replies from the proponents of the Reasoning view to this problem in two camps. On one side,

there are those who revise the view (the definition in the Reasoning View) in ways that are supposed to allow for outweighed reasons. On the other side, there are those who revise some of the background assumptions concerning patterns of good reasoning. Let us look briefly at both of these lines of reply.

The first line of reply proposes to specify the general version of the Reasoning view by qualifying its domain of application. Several specific ways of doing this exist. According to one influential approach (cf. Williams 2001; see also Williams 1979, 1989), in order for one to have a normative reason to F, one has to have some motivational set S (e.g. desires, emotions, and so on) such that there is for one a sound deliberative route from S together with a true belief that p to F-ing. According to one interpretation of Williams's view (see Way 2017), what Williams really meant is that the sound deliberative route ends in one being motivated to F, rather than F-ing *tout court*, where F stands for some action. Such an interpretation is indeed supported by textual evidence. Consider the following:

> This does not mean that when an agent has a thought of the form 'that is a reason for me to ϕ', he really has, or should really have, the thought 'that is a reason for me to ϕ *in virtue of my S*'. The disposition that forms part of his *S* just is the disposition to have thoughts of the form 'that is a reason for me to ϕ', and to act on them. (Williams 2001: 93)

Assuming that disposition to act on a consideration amounts to being motivated to act by it, Way's interpretation of Williams may indeed bear out. An even clearer case for attributing this interpretation of the Reasoning view to Williams can be found in the following passage:

> There is indeed a vagueness about '*A* has reason to ϕ', in the internal sense, insofar as the deliberative processes which could lead from *A*'s present *S* to his being motivated to ϕ may be more or less ambitiously conceived. But this is no embarrassment to those who take as basic the internal conception of reasons for action. (Williams 1981: 110)

Now, even if Williams is not explicitly committed to it, one might think that this would be a natural way for a proponent of Williams's account to develop such a view in further detail. It makes sense to go in this direction in particular given the context of the debate in which Williams's account takes part. His opponents are, of course, externalists about reasons, like Scanlon (1998), who do not want to postulate a necessary or essential link between one's motivation set and normative reasons. But Williams also opposes some other views within the broad family of Humean approaches (of which Williams's proposal is a part). For instance, Williams rejects the

ideal reasoner models. He doesn't endorse Smith's account, which is committed to classifying desires to F as outputs of the relevant pieces of practical reasoning by one's idealised counterpart. Thus, we may think that on a charitable interpretation of Williams's view reasons are considerations that play a role in a sound deliberative route – good reasoning, broadly understood from one's desires and other states to being motivated to F rather than to F-ing. If so, the objection from outweighed reasons might be blocked as long as we think of being motivated as coming in degrees.

A more detailed version of this line of reply to the objection has been recently developed by Kieran Setiya (2007, 2014). Recall Setiya's account of reasons for action:

> *Reasons*: The fact that *p* is a reason for A to ɸ just in case A has a collection of psychological states, C, such that the disposition to be moved to ɸ by C-and-the-belief-that-*p* is a good disposition of practical thought, and C contains no false beliefs. (Setiya 2007: 12, 2014: 222)

On Setiya's view, the relevant element is not a desire (of, say, an idealised and fully informed self), but being moved to F. Arguably, being moved to F and being motivated to F is the same kind of thing. And according to Setiya, this mere fact is enough for his preferred version of the Reasoning view to account for outweighed reasons:

> [T]his principle [that is, 'Reasons' above] is concerned with *pro tanto* reasons, reasons that can be outweighed; accordingly, it speaks of being moved, not of acting or intending. (Setiya 2014: 222)

On the face of it, revising the Reasoning view in such a way allows us to avoid the objection from outweighed reasons. Insofar as motivation, desire, or similar items to which a proponent of the Reasoning view can appeal, come in degrees, there is a possible story to tell about the outweighed reasons. One can claim, following Setiya, for instance, that you have (or there is) both a disposition to be moved to stay and help the injured in the accident (by some of the relevant premise-responses/psychological states) and a disposition to be moved to go and meet your friend (by the relevant premise-responses/states). Both dispositions are good dispositions of practical thought. Hence, the view predicts that the relevant considerations are both normative reasons. The crucial move is to claim that the relative weights of reasons correspond to 'the relative strength of motivation' (Setiya 2014: 229). More specifically:

> Reasons correspond to collections of psychological states that fuel good practical thought. One reason is *stronger* than another just in case it is a

good disposition of practical thought to be *more strongly moved* by the collection of states that corresponds to it, than by the collection that corresponds to the other. (Setiya 2007: 13)

A similar move is also available to theorists who appeal to the desires of an ideal and well-informed self (cf. Smith 1994). On such a view, one (e.g. the ideal self) has a stronger desire to help the injured than to go to the coffee shop. Hence, the relevant considerations (about the accident) outweigh the considerations about respecting the promise to the friend and correspond to a stronger or 'weightier' reason. Yet both remain reasons. On the motivation view: insofar as there is a good disposition to be somewhat moved by the promise consideration, the promise consideration still remains a reason for you to go to meet your friend, even if it is outweighed by the accident consideration, which is stronger since you are more strongly moved to help the injured.

This line of reply, attractive as it may appear, actually raises more problems than it promises to solve. Here are two of the most urgent ones. For one thing, it has trouble in explaining reasons for responses other than action. Take, for instance, reasons to believe. Motivation and belief don't seem to go well together. Normally, we don't have motivation for believing that such and such is the case. Typically, we are not moved to believe, while we are moved to act in some ways. Belief is a state at which we may arrive through reasoning or other belief-producing processes. Normally we don't arrive at a belief that p by being moved to believe that p. Action and belief appear to be quite different in this respect. Note also that when we do have a motivation to believe something, such a motivation is of a pragmatic sort and hence arguably can be linked only to the 'wrong kind' of reasons. Blaise may well be motivated to believe in God, since he thinks that such a belief will make him better off. But this sort of practical consideration is traditionally not seen as a normative reason for Blaise to believe in God (but see Chapter 6 for more on this). Without varying degrees of motivation to believe, we are thus back to square one. It is not clear how proponents of the Reasoning view who appeal to varying degrees of motivation (or of desire) to explain outweighed reasons to act could account for outweighed reasons to believe, as well as outweighed reasons to have other attitudes (i.e. it doesn't seem clear that we can have motivation to fear, to be angry, and so on, but we have reasons to fear, to be angry, and so on).

Now, one may try to provide even more sophisticated manoeuvres to account for reasons for beliefs and other attitudes. Elaborating on Setiya's account (cf. Setiya 2014), one might think that as we replaced acting by

being moved we may also replace believing by having an increased degree of confidence within a more sophisticated Reasoning view. And, according to this line of thought, we might reconcile the Reasoning view with there being outweighed reasons for belief. Alternatively, one might focus on inclinations to believe instead of being moved to believe.

However, critics of such a move remain unconvinced. Consider the case where one is already certain that p (and believes that p). In such cases, no increase in the degree of confidence is possible (and there is no more inclination to believe, for one already believes that p). And yet we might still discover previously unnoticed/new reasons for one to believe that p. That DNA analysis confirms the suspect's presence at the crime scene, that we have witnesses confirming her presence and we know that the suspect had a motive for the crime may convince us, indeed make us certain, that the suspect did it. We believe that she did it. And yet that we find out later that her fingerprints were on the temporarily lost murder weapon can nevertheless constitute another reason for us to believe that she did it, even if it does not incline us to believe that she did it (for we already believe that), nor does it increase our confidence (for we are already certain). It seems that the Reasoning view that appeals to inclinations to believe or degrees of confidence predicts that that there are fingerprints on the weapon cannot be reason for us to believe that the suspect did it.

One could try to get out of trouble by appeal to counterfactual considerations and claim that the relevant consideration only need to incline one to believe or increase one's confidence in some other possible circumstances. This move, however, is a tricky one. As Way notes (cf. Way 2016: 261), it requires, for one, that these other possible circumstances can be specified without appeal to reasons (otherwise, a vicious circularity looms), and it is not clear whether it can be done. For another thing, appeal to counterfactual inclinations to believe and counterfactual increases in degrees of confidence in this dialectical situation presents a risk of committing the conditional fallacy: the relevant changes in circumstances might be such that all other things are not equal and thus the appeal to possible other circumstances in which one is inclined to believe cannot constitute theoretically satisfactory grounds for drawing lessons about what reasons there are for one to believe in the actual circumstances (see Way 2016 for further details and references on these worries). In sum, it is unclear whether proponents of this version of the Reasoning view have any easy and fully satisfactory way of accounting for outweighed reasons to believe and to have other attitudes.

The second problem with this general line of reply is that it is not straightforward that we may be moved to act (or desire to act) in ways

recommended by outweighed reasons in situations where the outweighing reason massively outweighs the outweighed reason. Crucially it is not clear that in such cases there is a good disposition of practical thought to be moved to act in ways recommended by the outweighed reason. An objection similar to this one appears in Silverstein (2016).[4] Consider again our example of a promise for a coffee break versus a street accident. Arguably, when you witness the accident and have the relevant premise-responses (e.g. the intention to help, and so on), you are strongly moved to stay and help and not at all moved to go to the coffee shop. Crucially, it is not straightforward that there is a good disposition (or pattern, for that matter) of practical thought in this case from your intention to keep your promise to your friend and the belief that the only way to keep your promise is to walk away to being even the slightest bit moved to walk away. Given this and the abovementioned problem with outweighed reasons for belief, we can tentatively conclude that the move to revise the Reasoning view by focusing on patterns of good reasoning/good dispositions of practical thought that have as their conclusion being moved (or having the desire) to F rather than F-ing (intending to F) are unsuccessful in the light of outweighed reasons.

Given the problems of the revisionary versions of the Reasoning approach, some theorists have turned to an alternative line of reply to the problem of outweighed reasons. Namely, instead of revising the Reasoning view, they propose to rethink some of the background assumptions about patterns of good reasoning (cf. McHugh and Way 2016; Silverstein 2016; Way 2016; Asarnow 2017). The central move here is to reject the idea that good reasoning cannot be defeated; or, put more positively, they suggest that there can be patterns of good reasoning that may be defeasible. According to this line of thought, that reasoning from some premise-responses to a conclusion-response is good doesn't mean that its goodness cannot be undone if other elements were 'added' – namely, if other premise-responses were present. On this view, good

---

[4] I worry there is a more basic problem here, though. Once I realize I can save lives by breaking my promise and leaving you to find your own way home from the airport, I am not moved or motivated to pick you up at all. I recognize that my promise counts in favor of picking you up, but this just does not motivate me in the face of the countervailing considerations. Of course, I may regret that I will not be there to meet you, but I need not feel any motivational tug (or nudge) in the direction of the airport. If that is correct, then it is a mistake to identify the normative force of a reason with the motivation produced by sound deliberation from that reason. (Silverstein 2016: 12)

reasoning is always good reasoning *ceteris paribus*. If other things are not equal, then the reasoning is not good. The core idea is not new. It is a well-known topic in logic and theory of argumentation/reasoning (see Pollock 1986 and Horty 2012 for classical statements). To take a standard example from defeasible logic: one may draw the conclusion (c) that Tweety can fly from the premise (p1) that Tweety is a bird (and some background assumptions, perhaps, that most birds fly). It is an instance of a good inference (even if not classically valid). But it can certainly be defeated; for instance, when another premise is added, say, (p2) that Tweety is a penguin. The idea is often captured by reference to the formal property of consequence relation, known as *monotonicity*. An inference satisfies monotonicity where, roughly, in the case of a valid inference (e.g. where a conclusion follows from the set of premises), adding any other premise to the existing set of premises will not alter the validity of the inference. *Modus ponens*, for instance, is often presented as an inference that satisfies monotonicity: whatever you can add to the set of p, and if p then q, will not undermine the validity of inferring the conclusion q. Proponents of non-monotonicity insist that there can be good or cogent inferences even without such a strong constraint (the term of *cogency* is sometimes used as the equivalent of deductive *validity* in the context of non-deductive inferences).[5] For instance, the Tweety inference (i.e. the inference from p1 to c) is good/cogent within the defeasible logic that gives up on the monotonicity requirement for good/cogent inferences. Applying this to the case of outweighed reasons, proponents of the Reasoning view insist that an outweighed consideration, such as the promise consideration in our promise versus accident case, may still count as a reason to leave for the coffee shop. For there is a pattern of good reasoning from your intention to keep your promise to your friend to meet at the coffee shop and your belief that the only way to keep your promise is to leave, to leaving. It is a good pattern, albeit a defeasible one. And in this specific scenario, the reasoning is indeed defeated, since your other premise-responses – namely, your intention to help the injured and your true belief that the only way to help them is to stay – makes the situation such that all other things are not equal. These additional premise-responses undermine the *ceteris paribus* condition. If you were to reason from the promise considerations to an intention to go to the coffee shop, your reasoning in this accident situation would be a bad one – but not because there is no corresponding good

---

[5] Thanks to Aleks Knoks for a helpful discussion on defeasible inferences, and for drawing my attention to the notion of *cogency*.

pattern of reasoning. There is. It's just that moving to the intention to go to the coffee shop, given your promise *and* given further premise-responses in this situation (e.g. your true belief that there is an accident and so on), is not good reasoning.

Now, a question remains: what makes it the case that it is the promise-reasoning that is defeated, rather than the helping-the-injured reasoning? On this view, it has to do with the comparison of the relevant patterns of reasoning. See Way on this:

> [S]uppose R_1 is a good pattern of reasoning which concludes in φ-ing and R_2 is a good pattern of reasoning which concludes in a conflicting conclusion ψ. For R_1 to defeat R_2 is for it to be good reasoning to move from the premise-responses of R_1 to the conclusion-response of R_1, given the premise-responses of R_2, but not good reasoning to move from the premise-responses of R_2 to the conclusion-response of R_2, given the premise-responses of R_1. (Way 2017: 264)

So, the idea here is that two patterns of reasoning can be compared with respect to resilience: whether and which of the two defeats the other one. This measure then is what explains why one of the two reasonings in the cases of outweighed reasons is defeated and the other one is not. The promise-reasoning in our case is the defeated one, because it is less resilient. It is still good reasoning to move from the premise-responses of the accident-reasoning to the conclusion-response of the accident-reasoning, given the premise-responses of the promise-reasoning, while it is not good reasoning to move from the premise-responses of the promise-reasoning to the conclusion-response of the promise-reasoning, given the premise-responses of the accident-reasoning. The promise-reasoning is defeated in this case. Thus, the promise is the outweighed reason in our case. It is still a reason, though, for it still corresponds to a good pattern of defeasible reasoning (that happens to be defeated in the present instance). (For more details of this move, see Asarnow 2016; McHugh and Way 2016; Silverstein 2016; Way 2016.)

This second line of reply appears to deal better with some of the problematic points that the revisionary reply couldn't deal with satisfactorily. For one thing, it applies equally well to the case of reasons to act as well to reasons to believe and other attitudes. For the defeat condition seems to apply equally well to practical as well as to theoretical reasoning. Moving from a belief that Tweety is a bird to the belief that Tweety can fly is *ceteris paribus* good. Also, this line of reply doesn't commit the proponents of the Reasoning view to the questionable claim that one has to be

motivated (or desire) to some degree to leave the injured in order to have a reason to go to see one's friend. Given that this second line of reply can account for the existence of the outweighed reasons without being committed to some of the problematic claims to which the first line of reply was committed, it does seem to enjoy better argumentative support.

We can, thus, conclude that the proponents of the Reasoning view may be able to account for the problem of outweighed reasons by appeal to the defeasibility of patterns of good reasoning. However, before moving on, let us note quickly two further worries that one might have about this response and how a proponent of the Reasoning view may reply to these. First, one might worry that this view doesn't really explain why one reasoning is defeated by another by resilience and how exactly this defeat gives rise to the weight of reasons. One might think that there is something mysterious in this element of resilience. Are we supposed to assume, without argument, that some reasonings just are defeated and that this constitutes an ultimate element of explanation of the weight of reasons? Being able to explain the gradability aspect of reasons is a constraint on any viable theory of reasons. Many think of gradability here in terms of the weight of reasons. Where does the relative weight of reasons come from exactly in this picture? This may be understood as a question about what makes it the case that one reasoning defeats another (rather than, say, the other way round).

One may respond to this worry by pointing out that the talk of the 'weight' of reasons shouldn't be taken literally. The analogy has its limits (the next chapter expands on this a bit more). Perhaps an expectation of some deeper explanation of the weight of reasons comes merely from taking the 'weight' of reasons too literally. But there is no mystery here after all. Reasons surely appear to have a gradable aspect. And there are perfectly sensible comparisons of reasons. But there is nothing more substantial to it than that. Appeal to good reasoning and the possibility of one reasoning defeating another good reasoning is all there is as far as the gradability aspect of good reasoning and hence of reasons goes. Asking for a more substantial explanation of the weight of reasons is merely entertaining the illusion that the weight analogy should be taken to be somewhat more substantial. Good reasoning and defeat is where the explanation stops on this view.

A second, related worry here is about how the appeal to defeated reasonings can explain the fact that there can be different reasons for one and the same act, F, that can be compared among them. Both that the injured urgently need my help and the fact that helping the injured will

make me feel good about myself might be reasons for me to help the injured. But the former is certainly 'weightier' than the latter. How can the appeal to defeated reasoning account for this difference?[6]

One might reply to this worry that the explanation here is similar to the case of outweighed reasons. Roughly, in situations where r is a reason for S to F and q is a reason for S to F, but r is a more important ('weightier', 'stronger') reason to F than q, the r-reasoning is more resilient than the q-reasoning in the face of further considerations. That is, you can add more considerations to S's premise states without defeating the reasoning from r-premise-response to F than you can add to S's premise states without defeating the reasoning from q-premise-response to F. Concretely, there are further potential considerations that you could add as my premise-responses without defeating the goodness of reasoning from the premise-response of my belief that the injured urgently need my help to the conclusion-response of me helping the injured than there are potential considerations that you could add as my premise-responses without defeating the goodness of reasoning from the premise-response of my belief that helping the injured will make me feel better about myself. Typically, the mere fact that I gave a promise to be at the coffee shop will already defeat the goodness of that latter reasoning (note that to keep comparisons valid, we have to assume that in the case where my reason to help is that it will make me feel better about myself is not a case where the injured urgently need specifically my help to survive). Thus, the Reasoning view also seems to have the resources to account for the apparently different 'weight' of non-competing reasons. Ultimately, the explanation is provided by appeal to reasoning and degrees of resilience of a good pattern of reasoning in the face of further considerations. We can conclude that by revising the standard assumptions about good reasoning, one can save a version of the Reasoning view in the face of the prima facie worry of outweighed reasons (and related considerations about the comparative weight of reasons).

2.5   First Set of Worries II: Enablers

We are now at the following dialectical juncture: we have introduced the Reasoning view and some considerations in its favour, and then observed that there are two main prima facie worries in the current literature about

---

[6] For a somewhat similar line of objection see Star (2018: 255–256), who labels the problem as the problem of 'capturing ranges of weight'.

the Reasoning view. The previous section explored what is arguably the most discussed one of these, the objection from outweighed reasons. The focus of this section is on the remaining one.

The second prima facie problem for the Reasoning view is the problem of enabling conditions. In a nutshell, the problem is that the Reasoning view in its simple form is unable to distinguish normative reasons to F from mere enabling conditions for F-ing. For, according to this line of objection, both of these can figure as premises in patterns of good reasoning. The original point goes back (at least) to Dancy and his defence of holism about reasons according to which, roughly, what is a reason to F varies from case to case (see Dancy 2004). Dancy merely observes that not all premises in practical reasoning are [favouring] reasons for the F-ing that figures in the conclusion of the reasoning. Dancy doesn't focus specifically on the Reasoning view of reasons as we have presented it. However, his point does constitute a serious prima facie worry for anyone who wants to define or explain normative reasons in terms of patterns of good reasoning. Another way to put the worry is that the Reasoning view seems to predict that considerations that we are naturally inclined to treat as mere background conditions are themselves genuine reasons to F. Setiya (2014: 226) extrapolates:

> Dancy's more contentious claim is that even some of these considerations, which are premises of sound reasoning, are mere enabling conditions, not reasons to act. In his example, the fact that I have promised to do something is a reason to do it, while the absence of duress, possession of ability, and lack of competing reasons, though relevant to practical reasoning, are said to be mere conditions (Dancy 2004, pp. 38–41). Here *Reasons* [i.e. 'Reasons', from above] disagrees. It counts every premise of sound reasoning as a reason to act.

It does seem reasonable to take Dancy's observation at face value. Indeed, it is pre-theoretically plausible to see some considerations as mere conditions that make it possible that some other conditions are normative reasons to F without themselves being normative reasons to F. Dancy proposes to:

> [c]onsider the following piece of practical 'reasoning':
>
> 1. I promised to do it.
> 2. My promise was not given under duress.
> 3. I am able to do it.
> 4. There is no greater reason not to do it.
> 5. So: I do it.
>
> (Dancy 2004: 38)

The suggestion here is that not all elements in 1–5 are normative reasons. Indeed, Dancy claims that only 1 corresponds to a normative reason (in his terminology, to a 'favouring' reason). Premises 2, 3, and 4 correspond, according to Dancy, to enabling conditions. Very roughly, not all enabling conditions are alike on Dancy's view. On his account, 2 and 3 enable 1 to favour 5 (we are skipping the details about the differences between these two); whereas 4 enables 'the move from 1 to 5' (cf. Dancy 2004: 40).

Now, it would be unpromising for a proponent of the Reasoning view to dispute the fact that it does seem pre-theoretically very plausible to assume that at least considerations 2 and 3 are not reasons to do the thing referred to in 5. Premise 4 is arguably a slightly less clear case (cf. Setiya 2014).

But then what can a proponent of the Reasoning view propose that could block the argument from these intuitions against the Reasoning view? Unfortunately, there haven't been a lot of replies to this worry in the literature. However, there are still some proposals. Notably, Kieran Setiya has considered the issue and provided what one might think is a promising reply to the worry. Basically, he proposes to defend the thesis that 1–4 are, contrary to appearances, all normative reasons. He gives some theoretical explanation why it is so. And crucially, he sketches elements of an error theory of why it may appear to us that 2–4 are not reasons.

Here are some more details of Setiya's proposal (cf. Setiya 2014: 226–228). His strategy is twofold. First, he proposes an explanation that appeals to pragmatic aspects of communication. He claims that typically it is not sensible to cite 2, 3, and the like as reasons to F. But this, according to Setiya, is not because these are not really reasons, but rather that citing considerations about the absence of duress, one's ability, and so on are not informative enough, given the rarity of duress and so on. The idea here is that 'we assume' the absence of duress and that '[a]bility is arguably a condition of every practical reason' (cf. Setiya 2014: 226). Thus, presumably because something like the Gricean maxim of relation (be relevant or contribute all and only information that is relevant) is in place, we tend to focus on 1 (i.e. the promise) rather than 2, 3, or 4 when citing reasons for F-ing, which is represented in 5. According to Setiya (2014: 226): 'Citing the promise is thus a more informative and more natural way to bring out the pattern of practical reasoning under which the relevant motivation falls'. Compare this line of thought to a recent proposal by Fogal (2016), according to which, roughly, the fact that 'S likes dancing' and 'there is dancing tonight' cannot both express at the same time a reason for S to go to the dancing place, and hence, we should not rely on our language use of 'reasons' too heavily. According to Fogal, reasons might just be proxies or

representatives of 'normative clusters' or reason (the referent of mass noun 'reason'). According to Setiya, given the appropriate context, it would be perfectly acceptable to cite 2 and 3 as reasons.[7]

The second element of Setiya's explanation is to deny the additivity principle for reasons according to which, roughly, if $r1$ is a reason of 'weight' x for S to F and $r2$ is a reason of 'weight' y for S to F, then there is a conjunctive reason $r1$&$r2$ (alternatively: two reasons) for S to F of a total 'weight' of $x + y$. This enables him to respond, in particular, to the line of thought according to which 4 ('There is no greater reason not to do it') cannot be a reason to F. According to Setiya, we arrive at some absurd consequences by allowing 4 to be a reason to F only if we also accept additivity (note, however, that Setiya doesn't use the label 'additivity', nor 'weight', but talks of 'strength' of reasons instead). If additivity doesn't hold, we don't arrive at the problematic conclusion that 4 adds some more 'weight' to the total 'weight' of reasons that one has to do the thing one promised to do (i.e. 5) and hence there is nothing really absurd about counting 4 and the like among one's reasons.

It may well be that the additivity of 'weight' of reasons fails. That the thing on the menu is a pizza is a reason for me to order it; that the thing on the menu has chocolate cream on it is a reason for me to order it; however, that the thing on the menu is a chocolate cream pizza is in no way a reason for me to order it and hence has no 'weight'. The failure of additivity, sometimes also labelled 'accrual of reasons', is a lively topic in recent debates on reasons (Horty 2012: 61; Brown 2014; Bader 2016; Nair

---

[7] 'In the right context, however, any fact that is a premise of sound reasoning can be given as a reason. If I have been pressured into making various promises, so that duress is salient, the fact that a particular promise was made without duress will be a sensible thing to cite as a reason for keeping it' (Setiya 2014: 226). Note that this latter thought, that there are situations where it is perfectly fine to appeal to abilities or absence of duress as reasons, constitutes a positive argument in favour of the view that there is no significant distinction among enablers (and modifiers, and other conditions) and reasons, given the assumption that this fact about language and common sense is to be taken on board. A similar argument can also be found in Fogal (2016: 101), who endorses a different view on reasons (he is not a proponent of the Reasoning view, but in this aspect his view is similar to the Reasoning view):

> For example, a lot of philosophers want to deny that the ability to φ is itself a reason – or part of a reason – to φ but there are many contexts in which it seems perfectly acceptable to cite one's ability to φ as a reason to φ. [. . .] One option is to bite the bullet and say that they're strictly speaking false, offering some pragmatic story to explain (or explain away) their acceptability. Another – which I favor – is to accept them as perfectly good reasons claims, since the facts cited are good representatives in the imagined contexts and that's all that being – i.e., counting as – a reason really amounts to. Nothing would then follow about the specific metaphysical role the facts play within the cluster they represent – that would be a further, substantive issue, to be settled on broadly theoretical grounds Fogal (2016: 101).

2016; Hawthorne and Magidor 2018; and Sher 2019 contain considerations against additivity), and we will return to it in later chapters (see Chapter 3, for instance).

However, even if additivity indeed fails and Setiya's second point is well taken, one may still have some reasonable doubts about the first element of Setiya's proposal. Presumably, when Setiya appeals to pragmatic explanations in dealing with 2 and 3, the idea is that the fact that in a given situation it is not felicitous to assert that that one didn't make the promise to F under duress is a reason to F can be explained by an appeal to implicatures and conversational maxims. But can the oddity of citing 2 and 3 as reasons to F really be explained by an appeal to implicatures and conversational maxims? Consider what such an explanation would amount to exactly. We referred to the appeal to the Gricean maxim of relation (or relevance) and an implicature that would be generated by flouting the maxim of relation in the relevant context (cf. Grice 1975) as a possible explanation of the oddity. The idea seems to be that we expect people to communicate the most significant and specific bits of their information. However, according to this line of thought, when one offers 2 or 3 as a reply to the question, say, 'what reasons are there to F?' (i.e. do the promised thing), one is flouting the maxim of relation. That is, one is breaking it, by not communicating the most significant or specific information in the context (namely, 1, the promise itself) and one is breaking the maxim (i.e. not complying with it) in a flagrant manner. This flouting, then, should itself generate an implicature (given the assumption that people are cooperative): others should be entitled to imply something further from such flouting, from the fact that one is not providing the most significant bit of information. A natural thought here, then, would be that according to this line of reply, one should infer from the fact that a subject offers 2 or 3 rather than 1 that the subject doesn't actually have any further information other than 2 or 3 (as reasons to F), which then, presumably, according to this thought, generates the oddity. For we know that there is also 1 and perhaps find it puzzling that one could have 2 or 3 as reasons to F without also having 1. Thus, one might develop Setiya's idea here more specifically, as a case where offering 2 or 3 as reasons to do the promised thing is odd because one is offering some information that is known not to be the most specific information but the resulting implicature (there is no reason other than 2 and 3 to F) clashes with that fact (i.e. that there is a more specific bit of information – namely, that 1 is a reason to F). And, presumably, that 1 is a reason to F is more specific, because ability to F is always present where one has a reason to F, and duress is normally absent in typical cases of promises.

This line of reply appears problematic for the following reason.[8] This pragmatic story doesn't survive the cancelability test. It is assumed that implicatures, typically, can be cancelled. So, for instance, if you ask me 'Where are my keys?' and I reply 'Somewhere in the apartment', you are entitled to infer that, assuming that I am in the business of cooperative communication, I am implying that I don't have any more specific information about your keys. For the maxim of relation requires that I give you the most informative reply, and if I know something more specific than that your keys are in the apartment, then I am not giving you the most specific bit of information that I have. Now, it is observed that implicatures can be cancelled. To see how, consider the following modified version of our communication where, upon you asking me where your keys are, I reply not merely that they are somewhere in the apartment, but that 'they are somewhere in the apartment, but I don't mean to imply that I don't have any more specific information about the location of your keys'. If I say this, you are not entitled to infer that I don't have any more specific information. You will probably conclude that I am in a playful mood or that I am trying to give you a paternalistic lesson about the importance of remembering where you put your stuff. Crucially, there is nothing odd from the linguistic point of view in my communication (even though it might not be a very friendly piece of communication on my part). Cancellability is a common, typical feature of implicatures. But does it hold in the abovementioned case of promise and 1–5 reasoning? If Setiya (or rather our reconstruction of his argument) is right, then by offering 2 or 3, one is implying that one doesn't have anything more specific as reason to F (say, 1). Hence, it should be the case that one could cancel the implicature, by asserting something like 'that I am able to do the promised thing is a reason for me to do it, but I don't mean to imply that I don't have any more specific reason to do the promised thing'. However, this assertion and, in particular, its first conjunct still appear odd. Consider a variation on this: 'That I promised to meet you at 2:00 p.m. is a reason for

---

[8] My strategy here is inspired by a parallel argument from Jessica Brown (see Brown 2013), against the view according to which a piece of evidence *e* can be evidence for itself. Brown observes the oddity of such a proposal and considers in detail putative explanations that have been proposed to explain (away) the pre-theoretical oddity of that claim. The strategies that Brown examines (from Williamson 2000) strongly resemble Setiya's proposal here to appeal to pragmatics of conversation in order to explain why it is odd to cite ability and absence of duress as reasons to do the promised thing. Brown rejects in particular the proposal that the claim that evidence can be evidence for itself is explained by the fact that it generates an implicature given the maxim of relation. Brown rejects that proposal on the grounds that, roughly, the oddity remains even if one attempts to cancel the putative implicature.

me to meet you at 2:00 p.m. but so is the fact that I am able to meet you at 2 p.m.'. This seems odd. If Setiya's line of reasoning was correct and the only problem with offering 2 or 3 as reasons to F was that they generate implicatures along the lines that we drew above, then it shouldn't be odd at all to assert this and thereby cancel the alleged source of oddity, the implicature. But it is still odd. And this raises a serious suspicion about Setiya's pragmatic proposal. It doesn't seem that the problems for the Reasoning view from reasoning like 1–5 can be solved by appeal to some pragmatic aspects of communication. This is not to say that there cannot be another more theoretical motivation to reject the idea of enabling conditions not being reasons (see Fogal 2016 for one line of thought; for another, see Kearns and Star 2013: 84–86). However, it doesn't look like there is a forthcoming solution from the Reasoning approach.

2.6 Still More Worries: No Good Reasoning Available

In the two preceding sections, we have been exploring a couple of well-known worries for the Reasoning view. We saw the challenge that the existence of outweighed reasons (and 'weight' of reasons in general) raises for the view, and we also discussed the problem from enabling conditions as distinct from reasons. We also saw that proponents of the Reasoning view have provided some possible replies to these objections and that at least one of these might be on the right track – for example, the appeal to defeasible reasoning as a reply to the objection from outweighed reasons.

This section introduces a new challenge for the Reasoning view. It is a challenge from observations about cases where, roughly, no pattern of good reasoning is available from a given consideration, and yet we have an inclination to count the relevant consideration as a genuine normative reason.

The challenge discussed in the present section, I believe, is more radical and more worrisome than the worries discussed earlier. If the considerations ahead are on the right track, then there is a robust category of normative reasons that cannot even possibly comply with the Reasoning view's definition of reasons.

Let us focus on McHugh and Way's version of the Reasoning view, (RV), according to which 'for that p to be a reason for a response is for that p to be a premise of a good pattern of reasoning from fitting responses to that response' (McHugh and Way 2016: 586). This focus is principally motivated by considerations of simplicity and brevity. I believe that the

same sort of worry, *modulo* adapting it to individual specificities, can be advanced against other versions of the Reasoning view.

The worry I put forward here relies on two as-yet-unnoticed but fundamental counterexamples. These two arise given two of McHugh and Way's (independently plausible) commitments in particular. The first commitment is the claim that reasoning is a transition from premise-responses to conclusion-responses of a certain sort. According to them, 'This transition is such that the conclusion response counts as based on, or held in the light of, the premise responses' (McHugh and Way 2016: 586). The second commitment is the claim that all normative reasons are contents of possible true beliefs or of other fitting premise-responses. According to McHugh and Way, a fitting belief is a true belief. They don't require that all the relevant premise-responses are actually held by the subject for the content of the relevant premise-responses to be normative reasons for the subject in question. (RV) focuses on patterns of reasoning. (RV) requires that all the relevant premise-responses are fitting and are part of a good pattern of reasoning – that is, that there is a possible reasoning (for a given subject) from the relevant (not necessarily actually held) fitting premise-responses to the relevant conclusion-responses that respects the criteria for a good pattern of reasoning. How exactly the good patterns of reasoning are defined is not crucial for our purposes here.

2.6.1   Moore-Paradoxical Beliefs

The first of our counterexamples appeals to possible considerations of the Moore-paradoxical form. Consider the following example. The fact that (r) 'the building is on fire, but John doesn't believe that the building is on fire' is, intuitively, a reason for John to check/consider/reconsider/investigate the hypothesis (h) 'the building is on fire.' Yet there is no possible good pattern of reasoning for John from a fitting belief in (r) to reconsidering/investigation/and so forth of (h). This is so simply because it is not possible for John to have a fitting belief that the building is on fire and that he doesn't believe that the building is on fire. John cannot have a fitting belief that (r) because it is impossible for such a belief to be true. One cannot truly believe that the building is on fire and that one doesn't believe that the building is on fire. The belief in the first conjunct contradicts the belief in the second. On the (RV) account, given that John cannot (in any sense) have a true belief in (r), (r) cannot possibly be a reason for any response from him. However, it appears pre-theoretically plausible to think that (r) speaks in favour of some response for John. Given that the fittingness

requirement is central for (RV), Moore-paradoxical beliefs constitute a non-negligible challenge for the Reasoning view.

2.6.2   Self-Undermining Beliefs

The second sort of counterexample appeals to the possibility of other self-undermining beliefs. The fact that (p) 'I just took a drug that erased all of my memories about the past five minutes' is, intuitively, a reason for me to suspend judgment about what I did in these past five minutes (assuming that I haven't yet learned anything new after that). However, to suspend judgment about what I did in these past five minutes entails suspending judgment about whether I took the memory-erasing drug. But there is no good pattern of reasoning from a belief that p towards a suspension of judgment about p. Such a transition (if possible) doesn't satisfy the basic criteria for reasoning. In fact, McHugh and Way define reasoning as a transition where the conclusion-response is held in the light, or on the basis, of the premise-response. In this case, however, I cannot base my suspension about p on my belief that p. To suspend judgment about whether I took the memory-erasing drug cannot be based on the belief that I took the memory-erasing drug. At the very moment when I would suspend judgment, I would lose the basis for the suspension, and it would not count as being held in the light of the relevant premise-response. It is not possible to believe that p and, at the same time, to suspend judgment about p (on the basis of the belief that p). If the fact that I just took a drug that erased all my memories about the past five minutes is a reason for me to suspend judgment about what I did during these past five minutes, then (RV) must be false, since it entails that it cannot be a reason for me to suspend judgment.

Now, a proponent of the Reasoning view might reply to this line of argument from the cases of Moore-paradoxical and self-undermining beliefs by suggesting that in all such cases there are actually facts in the vicinity of the problematic considerations that are both normative reasons for the same F-ing and, at the same time, considerations that can play the relevant premise role in good patterns of reasoning towards F-ing.[9] According to this suggestion – for example, in the Moore-paradoxical belief case – there is a non-paradoxical fact in the vicinity of the Moore-paradoxical consideration, that is a reason for one to check/investigate the

[9] Thanks to an anonymous reader for Cambridge University Press for drawing my attention to the need to address this potential reply to my argument.

hypothesis that the building is on fire. One might, for example, think that the fact that the building is on fire is one such fact, and that the building may well be on fire might be another such fact. A similar line of response has been recently put forward by Hille Paakkunainen in a discussion concerning a similar worry for what she calls the Deliberative Constraint on normative reasons (see Paakkunainen 2017, 2018). Roughly, according to the Deliberative Constraint, a consideration r is a reason for S to F only if r can be a premise in a good pattern of reasoning/deliberation to F (see Paakkunainen 2018: 156 and Paakkunainen 2017: 65 for the more precise and detailed formulation of the constraint, which may be considered as the left-to-right conditional of the Reasoning view, if we spell it out in a biconditional form). Paakkunainen responds in these publications to objections from cases where the relevant considerations seem to be unable to play a premise role in [good] deliberation. We will return to these cases more attentively in the next section within our more general discussion of what other authors have called the Response constraint (cf. Way and Whiting 2016), which is also connected to the more general and theory-driven discussion about the supposed guidance role of normative reasons. But with respect to our present discussion, a proponent of the Reasoning view might well appeal to the same point that Paakkunainen makes in response to these further cases and claim that there are facts in the vicinity of the Moore-paradoxical and self-undermining considerations that can be reasons to check and reasons to suspend and can also still be premises of good patterns of reasoning.[10]

To this worry I would like to reply two things. First, such a response from the proponents of the Reasoning view would amount either to the claim that we don't have the intuition that – Moore-paradoxical considerations are reasons for checking or to the claim that our intuitions are massively mistaken. Neither of these options is promising, though. We may well accept that there are other reasons in the vicinity of Moore-

---

[10] Another point that Paakkunainen makes against these cases where it doesn't seem that the relevant considerations can play a role in (good) deliberation is that there might be alternative evaluative and normative phenomena that constitute the real focus of these cases. In particular, she suggests that the relevant considerations might be explanatory reasons that explain why the relevant F-ing 'would be a favorable outcome from the perspective of [one's] preference-satisfaction; or why [one's F-ing] would be good, or good for [one]' (Paakkunainen 2017: 68). Such a suggestion comes very close to our own positive proposal below (see Chapter 5). Yet, on our proposal, the considerations that don't play a role in a good deliberation can still be reasons. But let us not precipitate that discussion yet. See also Rossi (2021) for a response to Paakkunainen in which a new kind of elusive reasons are introduced. Rossi also provides further considerations for the claim that such reasons are genuinely normative and authoritative.

paradoxical and self-undermining considerations. Yet the argument from Moore-paradoxical and self-undermining considerations still goes through even if there are these additional reasons. Now, one might, of course, claim that it is somewhat indeterminate or not obvious what exactly our judgments track in the relevant cases (cf. Paakkunainen 2017: 67–68). But why should we doubt the reliability of our self-reported judgments specifically in these cases? Why should we think that we are massively in error when we consider Moore-paradoxical and self-undermining belief cases? In order for this response to be successful, we need an independent, theoretically well-motivated error theory that would explain either why exactly our judgments specifically in Moore-paradoxical and self-undermining cases are not about normative reasons to check and alike but are indeterminate or not obvious, or why we are so massively mistaken in having these judgments. As far as I can tell, no independent error theory that could explain this is forthcoming. But without such an error theory, we can take our judgments about these cases at face value.

The second point that I would like to make in reply to this line of response is that it is not clear that the appeal to the facts in the vicinity of Moore-paradoxical and self-undermining considerations would help the proponents of the Reasoning view at all. Our suggestion is that, for example, the Moore-paradoxical consideration is a reason specifically to check or investigate (or similar). But it is not clear why the invoked alternatives in the vicinity of Moore-paradoxical considerations – for example, that the building is/may be on fire – would count as reasons *specifically* to check or investigate, rather than to run away or to call the emergency number. The proposed alternative considerations don't seem to speak specifically in favour of checking. It is specifically the conjunction of the fact that there is a fire, and that one doesn't know about it, that speaks in favour of checking. The ignorance factor that is part of the Moore-paradoxical consideration has normative importance, it would seem. Thus, it is not clear that appealing to these further facts in the vicinity of Moore-paradoxical and self-undermining considerations could help proponents of the Reasoning view to alleviate the problems that our challenge raises. I suggest that the challenge from Moore-paradoxical and self-undermining considerations for the Reasoning view is still germane and cannot be easily dealt with.

2.7   Are Moore-Paradoxical and Self-Undermining Beliefs Really Worse Than Ice Creams and Surprise Parties?

One thing that proponents of the Reasoning view could do in the light of the objection from the preceding section is to appeal to existing moves that

have been made within the closely related debate on so-called guidance by normative reasons. A central assumption there is that reasons should be able to guide. And this can be understood in terms of there being a response constraint on reasons. That is, the constraint according to which only if you are able to act/have an attitude on the basis of r, can r be a reason for you to act in the relevant way/have the relevant attitude (cf. Way and Whiting 2016; see below for more details).

Thus, a proponent of the Reasoning view may claim that an independent theoretical motivation for not counting our aforementioned examples of Moore-paradoxical and self-undermining beliefs as genuine cases of normative reasons is that it leads to the rejection not only of the Reasoning view but also of the very plausible response constraint on reasons and consequently to the difficulty of explaining how normative reasons are supposed to guide us. Yet the guidance aspect of normative reasons, they may maintain, is a non-negotiable aspect of normative reasons.[11] That reasons are centrally supposed to guide us is indeed a widely held claim. For some, it is even a platitude that any viable theory of reasons has to integrate. Insofar as the aforementioned examples not only undermine the Reasoning view, but they also entail that not all reasons are such that we can act upon them, they make it difficult to understand how normative reasons could guide us. If they don't guide us, it is not clear in what sense normative reasons are still normative, according to this line of thought. If they cannot guide us, they cannot exert any force on us; hence, they lose their normativity, or so it seems.

However, as soon as someone declares a thesis beyond any doubt, philosophers at once come up with endless counterexamples and try to dispute it at considerable length. This is also what happens with respect to the response constraint on normative reasons. Let us explore this debate briefly and see what can it add to our discussion about the Reasoning view. (In what follows, I borrow some elements of Way and Whiting's (2016) terminology and their way of introducing the debate.)

According to Way and Whiting (2016: 214), it is a common assumption that reasons are supposed to guide us and '[c]onsiderations that cannot guide cannot do what reasons are supposed to do'. In other terms, 'it is the job of a reason to recommend that a person perform a certain act or hold some attitude. If it is [to] do that job, the relevant person must be able to heed and respond to its recommendation' (Way and Whiting 2016:

---

[11] Thanks to Daniel Whiting for making me aware of the importance of this possible line of reply on behalf of proponents of the Reasoning view.

214). Here is their formulation of the response constraint, which they seem to present as a natural way to develop the thought that reasons must be able to recommend actions and attitudes and hence guide us:

> **Response Constraint**: That p is a reason for you to φ only if you are able to φ for the reason that p. (Way and Whiting 2016: 214)

Thus, a central argument in favour of the Response Constraint is that it provides a clear and simple way to account for the guidance aspect of normative reasons. Combined with Humeanism, it also avoids a kind of mysticism that other views on reasons and guidance may have. According to Humeanism (e.g. Williams 1979), reasons are necessarily connected to one's motivational set that contain pro-attitudes (note also that strictly speaking Humeanism is not entailed by the Reasoning view of reasons which doesn't specify on its own the sort of premise-responses that one may have – for example, the Reasoning view is compatible with an interpretation on which only beliefs can constitute premise-responses).[12] Think, for a comparison, of the reasons-first approach that only tells us that reasons exist and presumably that at least some people might respond to them. In the case of the Reasoning view that endorses the response constraint, or a Humeanism with a response constraint, we have a straight-forward account of guidance. There is no mystery of how exactly the guidance goes with such a combination. Moreover, the account has a naturalistic flavour, for it seems in principle to be open to a further reductionist account that explores an appeal to some evolutionary story.

Another line of argument in favour of the Response Constraint is that it enables us to explain a number of distinctions within the normative domain. For instance, it allows us to explain why there are no normative reasons for instantiations of value. For example, while there can be a reason for one to want to be a bit taller, there is no reason for one to *be* a bit taller (cf. Parfit 2011: 51). Having certain evaluative properties is not the kind of thing that we are able to *do*, let alone to do for a reason. Hence, it is not something that can have reasons in its favour. More generally, as Way and Whiting observe, the Response Constraint enables one to have the first steps towards an explanation of differences between deontic and axiological properties. Another intuitive distinction that one may be able to explain by

---

[12] Contrary to what can often be seen in the literature, I prefer to avoid using the term 'internalism' for the sort of Humeanism about reasons to act that people associate with Bernard Williams in order to avoid confusions when it comes to the discussion of reasons for belief, where 'internalism' refers to a different sort of view. I am not alone in this. Way and Whiting also avoid using the 'internalism' label.

an appeal to the Response Constraint is the distinction between genuine reasons to believe and pragmatic reasons to believe. One might think that nothing but truth related, evidential considerations can play the role of reasons to believe. Yet sometimes pragmatic considerations are cited as if they were reasons to believe (e.g. that being cheerful is good for one's health is a reason to believe that everything is fine). This gets explained away by appeal to the Response Constraint insofar as we are unable to believe that *p* for reasons that are not related to *p* (cf. Kelly 2002; Shah 2006; among others). Again, this is not to say that pragmatic reasons cannot be reasons to want to believe that *p* (see, again, Chapter 6 for a discussion of this point).

Despite these considerations in favour of the Response Constraint, it is not universally accepted. In fact, there is a battery of counterexamples against the Response Constraint (or at any rate to constraints close enough to it). Julia Markovits (2011a), for instance, elaborates a number of well-known and new cases. Note that at the end of the day, Markovits endorses a version of the Humean (her label is 'internalist') view that still links reasons to one's cares and other states but refuses to accept that reasons have to be able to motivate one. She provides three kinds of cases. First, there are situations where one's being not perfectly rational prevents one from being able to act upon one's reason in a situation (e.g. cases where a sore loser has a reason to leave without greeting their opponent, cases where a delusional subject has a reason to consult a doctor). Second, there are situations where one is not particularly irrational, but the structure of the case is such that one cannot act upon one's reason to F (e.g. cases of deterrent actions, like intending to respond to a nuclear attack, Kavka's toxin case, and soldiers fighting in a just war). Third, there are cases where one ought not to be moved by the reason one has (e.g. an emergency plane landing, complex medical operations, or automatic responses in cases of road accidents, where acting on the basis of the reason one has, say, to save an innocent life, presents a very high risk of screwing the relevant act up). Way and Whiting (2016) centre their discussion specifically on two further cases, and we follow suit in what follows.

First is the case of 'massively outweighed reasons' (cf. Way and Whiting 2016: 215), which comes from Mark Schroeder:

> **[Ice Cream]** Joel's career, his wife and her career, his friends, his Lakers' season tickets, his family, and his loves of surfing and of mountain climbing all tie him to Los Angeles. But Joel also loves chocolate-cayenne-cinnamon flavored ice cream, which he can only get in Madison, Wisconsin. (Schroeder 2007: 166)

The common diagnosis here is that it appears that the fancy ice cream is a (tiny) reason to move to Madison. See Way and Whiting (2016: 216) on this: '[t]he fact that he can only get chocolate-cayenne-cinnamon flavored ice cream (hereafter, ice cream) in Madison is a reason for Joel to move there'. The problem here is that 'Joel is not capable of moving to Madison for that reason, given how much he cares about all of the things tying him to LA' (Way and Whiting 2016: 216).

The second example that Way and Whiting discuss is the case of 'self-effacing reasons' (cf. Way and Whiting 2016: 216):

> **[Surprise Party]** There is a surprise party for Beth at her house that starts at 5pm. Beth loves surprise parties and it would make her very happy to arrive home at 5pm. However, were Beth to find out that there is a surprise party for her at 5pm, the surprise would be ruined, and the party would be a disaster. (Way and Whiting 2016: 216)

This second example is attributed to Markovits (2011b) and Schroeder (2007: 165–166) (Way and Whiting also observe that similar examples can be found in Millgram 1996; Sobel 2001; Shafer-Landau 2003: ch. 7; Markovits 2011a; Smith 2009). The problem here is that, first, it appears that that there is a surprise party is a reason for Beth to go home, but, second, it is not possible for Beth to act upon that reason, since it would require, at minimum, that Beth is aware or believes that there is a surprise party, which would entail, paradoxically, that the party is not a surprise party for Beth anymore.

Now, one can always deny that the aforementioned cases, in particular cases of 'self-effacing reasons', are cases of genuine reasons, because one might think that reasons must satisfy the Response Constraint and must guide us. Way and Whiting suggest that it is the most popular line of reply to these cases and identify, for instance, Setiya (2009: 538), Sinclair (2012), and Kiesewetter (2016) as versions of this line of reply.

However, such a denial is problematic. For denying that contrary to our robust judgments about all these various and numerous situations, the relevant considerations are not really normative reasons is a substantial cost for a theory of normative reasons. We have to realise how massive and pervasive cases like these are (see Markovits 2011a for more on this point). One might think that an outright refusal to count these considerations as reasons has an ad hocness flavour. Moreover, one cannot merely deny that the relevant considerations genuinely count in favour of the relevant F-ings in a sense in these cases since our pre-theoretical judgments indicate otherwise. But then if these considerations are not reasons, but still count

in favour of the relevant F-ing, one is forced to introduce a new normative category to account for them. Recognising them as reasons appears to be a more parsimonious and theoretically fruitful strategy.

Another option that one can take in the face of these counterexamples to the Response Constraint is to modify the Response Constraint in order to try to avoid them. Way and Whiting themselves take this other option (they also discuss some alternative options of this sort; for matters of simplicity of exposition, I will not review them here; my suspicion is that they will face the same problems as Way and Whiting's proposal). Way and Whiting acknowledge that there are two possible interpretations of the Response Constraint depending on how one understands ability. They follow Mele (2002) in distinguishing two sorts of abilities: general abilities (understood 'as a kind of *power* or *competence*'; Way and Whiting 2016: 219) and specific abilities. To take their example, Andy Murray has general ability to serve (in tennis). He has it even when he is on a plane or drunk or when something is interfering with these abilities, or otherwise he happens to be in unfavourable circumstances (say, there is no tennis court around). Murray has the specific ability to serve when he is playing a concrete tennis game.

According to Way and Whiting, interpreting the Response Constraint in terms of specific ability is unpromising. But when interpreted in terms of general ability, the Response Constraint leaves enough room to respond to the problematic cases. The treatment of cases that they provide is twofold. With respect to ice-cream cases, where reasons are massively outweighed, an interference blocks a subject from exercising her general ability to F for the relevant reasons. For instance, one is, in general, able to move to Madison, but one's actual psychology interferes with this ability – one's other cares and desires block one from exercising the ability to move to Madison. In the case of self-effacing reasons, their explanation is that the subject has some sort of general ability that is relevant for the Response Constraint. So, for instance: 'Beth might well have the general ability to reason "there's a surprise F at location L, so I'll go to L", when surprise F is something that she enjoys or is interested in' (Way and Whiting 2016: 224). This response raises some tricky issues with respect to the level of grain of content of reasons. Yet, without going into details of these, one might wonder whether the suggested strategy can succeed. In particular, one may wonder why the suggested general ability of reasoning from some abstract considerations 'there's a surprise F at location L' should be taken as capturing the specifics of the case, the case where the reason is 'there is a surprise party for her at 5:00 p.m.' and not 'there is some surprise,

somewhere'. In what sense is the general, abstract ability of acting upon the consideration that there is some surprise, an ability to act upon the specific consideration that there is a surprise party for me? But even if this line of reply can be worked out for cases like Surprise Party, I think there is another problem that proponents of the Response Constraint face. Let's see this now.

Going back to our initial cases against the Reasoning view, one can claim that there is not even a general coarse-grained ability there. In no sense, neither in a general nor in a specific sense, can one act on the basis of Moore-paradoxical considerations ('p and I don't know that p'). In a way, our case is even more radical than the self-effacing reasons cases. The claim that one might have a general coarse-grained ability to check the building/ investigate the situation on the basis of the fact that there is a fire *and* one doesn't know that there is a fire appears even more problematic than the parallel claim about self-effacing reasons cases – for example, the ice-cream case. One doesn't have the general coarse-grained ability to reason 'there is a fire in the building but I don't know about it, so let's check whether there is a fire in the building'. I don't see how one could have an ability to undergo this kind of absurd reasoning.

At this point, I suppose, Way and Whiting's conciliatory strategy provides no further help. It cannot accommodate the pre-theoretical assumption that our Moore-paradoxical case is a case of a genuine normative reasons. But denying that these considerations are genuine reasons seems to be an ad hoc move, especially after admitting that a number of structurally close, albeit a bit less radical cases (e.g. surprise party) are cases of normative reasons. And again, even if one is ready to deny that Moore-paradoxical considerations are reasons, one cannot, I would like to suggest, plausibly deny that they do count in favour of the relevant F-ings. Now, if one denies that they are reasons but recognises that they still are considerations that count in favour, one is, thereby, introducing a new normative category. But multiplying normative entities and categories should require some caution. A theoretically more parsimonious approach would not classify them as a normative category of its own kind, but rather associate them with the already existing category of normative reasons. That they bear some similarity to the other apparently problematic reasons above introduced (e.g. ice-creams, surprise parties, and Markovits's cases) may be taken as a fact speaking in favour of counting them as reasons as well.

Now, what about the positive arguments in favour of the Response Constraint? Here is a tentative line of thought. I would like to express some doubts about the supposed centrality of guidance for reasons.

According to one prominent argument in contemporary epistemology, the Anti-Luminosity argument (cf. Williamson 2000: 93–113), there is no condition *c* such that one can always know that one is in *c* when one is in *c*. Application of the Anti-Luminosity argument to normativity is not new (cf. Srinivasan 2015; Hughes 2018; Lasonen-Aarnio 2019; see also Section 1.3). Adapting it to the case of guidance by reasons amounts to the claim that, roughly, no reason is such that it can always guide us in F-ing when it is a reason to F, assuming that awareness or knowledge is necessary for being able to act upon a reason. In other terms, Anti-Luminosity appears to put some pressure on the very motivation of the guidance constraint on reasons.

Another line of thought here is a tentative suggestion that part of what people find attractive in the guidance requirement is the underlying thought that one cannot be held responsible for things one cannot possibly have done. And if normative reasons are normative, they are in one way or another connected to criticism and praise. Say, if we think that one ought to do what there is the most reason to do and we think that one can be reasonably criticised for not doing what one ought to do (and acknowledged or praised for doing what one ought to do), then having a reason to F is potentially something that contributes towards one being blamed or praised. And only if one can be held responsible/accountable for F-ing can one be reasonably criticisable or praised for F-ing. Thus, one might wonder: if there is no way for one to F upon a reason to F, how could one be held responsible for F-ing? But there are other possible ways one could deal with this underlying thought than endorsing the Response Constraint. One might, for instance, think that being held responsible and ought are not as tightly connected as one might have initially thought. Or one may think that one might be held responsible for F-ing even if one is in no position to F for a reason *r* to F (see, however, Streumer 2007 against impossible reasons; see also Heuer 2010 for discussion). Finally, one might also attempt to distinguish more clearly responsibility from criticism/praising. Perhaps we have good reasons to keep these categories separate (as some recent work on addictions, for instance, appear to suggest; cf. Pickard 2017). At any rate, this is not to say that there is nothing odd to account for if we accept that Moore-paradoxical and self-undermining considerations can be genuine reasons. It is to say that there might be other ways of accounting for the apparent oddity, ways that don't commit one to the thesis that for a consideration to count as a reason to F, it has to respect the Response constraint (see Chapter 5 for a positive proposal). On the other hand, there doesn't seem to be an utterly

convincing way of dealing with Moore-paradoxical and self-undermining reasons without giving up the Reasoning view of reasons.

2.8 Concluding Remarks

A natural move for those who are dissatisfied with the reasons-first approach is to turn towards reductive accounts of normative reasons. A natural first move for those attracted by an attempt to define normative reasons in other more fundamental terms is to look into what reasons are supposed to do and what are the central functions of our common-sense notion of *a reason to act/hold an attitude*. A natural thought here is that normative reasons are those sorts of things that constitute good reasoning, for undeniably there has to be a connection between reasons and reasoning. This chapter has looked at one popular attempt to work out these natural thoughts into a fully elaborated account of normative reasons; namely, the Reasoning view of normative reasons, according to which a normative reasons r for S to F just is for r to be such that it is a content of a premise-response in a pattern of good reasoning from appropriate premise-responses to an appropriate conclusion-response F. In this chapter, we first explored in more detail what such a view amounts to. We reviewed positive considerations in its favour and then examined what appear to be the most troublesome objections to this view; namely, objections from outweighing reasons, enabling conditions, and reasons to F upon which one is not able to F. We saw that responses exist to these worries. But it is far from obvious whether all these are successful. To the contrary, it appears that the cumulative case of these problems favours looking into potential alternatives to the Reasoning view. Before doing just this and exploring other existing prominent reductive views of reasons, let me note that rejecting the Reasoning view doesn't mean giving up the positive insights it provides. Indeed, our positive view that will be developed in Chapter 5 vindicates exactly this. But before we are able to put our cards fully on the table, we need to go through other, also fundamental, insights about reasons that have not been brought to light by the Reasoning view. Namely, we have to look into the very natural thought that reasons are also crucially linked to explanations. Let us now turn to views that explore precisely this other apparently central aspect of reasons.