

# The normativity of rationality

Dissertation  
zur Erlangung des akademischen Grades  
Doctor philosophiae (Dr. phil.)

Eingereicht an der  
Philosophischen Fakultät I  
der Humboldt-Universität zu Berlin

am 18. Oktober 2013

von Benjamin Kiesewetter, M.A.

Präsident der Humboldt-Universität zu Berlin: Prof. Dr. Jan-Hendrik Olbertz  
Dekan der Philosophischen Fakultät I: Prof. Michael Seadle, Ph.D.

Gutachter:  
Prof. Dr. Thomas Schmidt  
Prof. R. Jay Wallace, Ph.D.



# Contents

---

<i>Contents</i> .....	iii
<i>Acknowledgements</i> .....	vii
1. Introduction: the normativity of rationality	1
1.1 The normative question about rationality.....	2
1.2 Normativity, reasons, and ‘ought’.....	5
1.3 Structural irrationality.....	17
1.4 Structural requirements of rationality.....	22
1.5 A range of views.....	25
1.6 Prospects.....	28
2. Rationality, reasons, and criticism	31
2.1 The criticism argument.....	31
2.2 The criticisability claim.....	34
2.3 Criticisability and the subjective ‘ought’.....	41
2.4 Criticisability and well-functioning.....	44
2.5 Criticisability and virtue.....	49
2.6 Rationality and criticism.....	50
2.7 The obscurity of non-normative requirements of rationality.....	56
2.8 Summary and prospects.....	59
3. Structural requirements of rationality	61
3.1 Wide- or narrow-scope requirements?.....	61
3.2 Proposition- or response-requirements?.....	69
3.3 Conditional or unconditional requirements?.....	79
3.4 Synchronic or diachronic requirements?.....	84
3.5 State- or process-requirements?.....	95
3.6 Summary.....	108
4. Bootstrapping and other detachment problems	109
4.1 The bootstrapping problem.....	109
4.2 The reasons strategy.....	112
4.3 The subjective ‘ought’ strategy.....	116
4.4 The wide-scope strategy.....	118

---

4.5	Wide-scope and <i>normative detachment</i> .....	120
4.6	Wide-scope and <i>liberal transmission</i> .....	129
4.7	Wide-scope and <i>necessary means transmission</i> .....	138
4.8	Summary.....	143
5.	The why-be-rational challenge	145
5.1	Are there instrumental reasons to be rational?.....	145
5.2	Are rational requirements themselves reasons?.....	150
5.3	Buck-passing and the normative question.....	152
5.4	Are there epistemic reasons to be theoretically rational?.....	156
5.5	The transparency account.....	160
5.6	The apparent reasons account.....	167
5.7	Summary.....	175
6.	The myth of structural rationality	177
6.1	Narrow-scope problem (i): bootstrapping.....	177
6.2	Narrow-scope problem (ii): inconsistency.....	182
6.3	Narrow-scope problem (iii): undermining.....	186
6.4	Wide-scope problem (i): asymmetry.....	194
6.5	Asymmetry and the rational process constraint.....	203
6.6	Wide-scope problem (ii): guidance.....	209
6.7	The myth of structural requirements of rationality.....	220
7.	Rationality as responding correctly to reasons	223
7.1	Taking stock: structural rationality and normative reasons.....	223
7.2	A fresh start: rationality as responding to reasons.....	224
7.3	Response rationality.....	229
7.4	The supervenience argument.....	232
7.5	The backup view.....	238
7.6	Evidence, practical reasons, and goodness.....	244
7.7	Theoretical rationality, uniqueness, and justificational fecundity.....	252
7.8	Evidence-based reasons and evidence for ‘ought’-propositions.....	259
7.9	Practical rationality, and reasons for actions and intentions.....	265
7.10	Summary and prospects.....	272
8.	An evidence-relative account of reasons	273
8.1	Objectivism and perspectivism about ‘ought’.....	273
8.2	Three objections to the objective view.....	279
8.3	The misguidance argument.....	284
8.4	The problem of advice.....	290
8.5	Reasons, time, and advice.....	294
8.6	The problem of seeking and sharing new evidence.....	298

---

8.7	Sharing evidence and the good.....	306
8.8	Objections.....	308
8.9	Conclusion.....	313
9.	Explaining structural irrationality	317
9.1	Explaining structural irrationality: the general idea.....	317
9.2	Structural and non-structural requirements, violation and satisfaction.....	323
9.3	An objection: satisfaction in the wrong direction.....	325
9.4	Explaining weakness of will.....	330
9.5	Explaining weakness of belief.....	332
9.6	Explaining <i>modus ponens</i> irrationality.....	336
10.	Instrumental rationality as a system of categorical imperatives	349
10.1	Explaining instrumental irrationality (in determined cases).....	350
10.2	The problem of underdetermination.....	355
10.3	Alternative explanations: normativism and cognitivism.....	363
10.4	A reason to decide.....	368
10.5	Instrumental rationality: final thoughts.....	376
10.6	Conclusion.....	380
	<i>Bibliography</i> .....	383



## Acknowledgements

---

During the years of working on this thesis, I received the help of numerous people and various institutions. First and foremost, I would like to thank Thomas Schmidt for his invaluable advice and steady support over all these years. I find it almost impossible to imagine a better supervisor than him and consider myself very lucky to have been one of his doctoral students. I am also very grateful to Jay Wallace for his commitment as a supervisor of this project during my time as a visiting scholar at UC Berkeley in 2009/10, and for many stimulating discussions.

I have greatly benefited from a multitude of conversations with and comments from many colleagues and friends, who were willing to read and discuss with me the material that entered into this thesis (as well as material that did not enter into it). I have to mention particularly Jan Gertken and Felix Koch for their extensive and extremely helpful comments on several draft chapters, and their constant readiness to exchange ideas and engage in philosophical arguments related (as well as unrelated) to the issues discussed here. I also owe special thanks to Vuko Andric, Katrin Beushausen, Anne Burkard, Niko Kolodny, Thomas Kroedel, and the participants of Thomas Schmidt's ethics colloquium for discussing substantial parts of this thesis with me; their comments have made significant contributions. Besides those already mentioned, I am greatly indebted, for the valuable feedback I received, to Monika Betzler, Tony Bezsylko, Michael Bratman, John Broome, Krister Bykvist, Luis Cheng-Guajardo, Eleonora Cresto, Guido Ehrhardt, Christoph Fehige, Steve Finlay, Branden Fitelson, Max Flaum, Daniel Friedrich, Pete Graham, Alex Gregory, Tim Henning, Frank Jackson, Susanne Mantel, Leo Menges, Elijah Millgram, Andreas Müller, Derek Parfit, Andreas Pittrich, Eduardo Rivera-López, Katrien Schaubroeck, Karsten Schoellner, Mark Schroeder, Sam Shpall, Michael Smith, Ezequiel Spector, Horacio Spector, Jack Woods, and Brandon Woolf.

I have also greatly benefited from discussions on several occasions at which I had the chance to present parts of the material. These occasions include the Berkeley-Stanford-Davis Graduate Philosophy Conference in 2010 (University of California, Berkeley); the “Reasons, Agency, and the Good” workshop in 2011 and the Humboldt-Princeton Graduate Conferences in Philosophy in 2011 and 2012 (all at Humboldt-Universität zu Berlin); the Deutscher Kongress für Philosophie in 2011 (Ludwig-Maximilians-Universität München); and the Kongress der Gesellschaft für analytische Philosophie in 2012 (Universität Konstanz); as well as colloquia, seminars and workshops at Humboldt-Universität zu Berlin; Rheinisch-Westfälische Technische Hochschule Aachen; Universidad Torcuato Di Tella, Buenos Aires; Universität des Saarlandes, Saarbrücken; University of California, Berkeley; and University of Oxford. I would like to thank the participants for their much appreciated feedback.

Work on this thesis would not have been possible without the generous financial support that I received during the last years from the Carl and Max Schneider Stiftung, the Leibniz-Preis Projekt “Transformationen des Geistes”, the Deutsche Forschungsgemeinschaft (SFB 644), and the Studienstiftung des deutschen Volkes.

Parts of Chapter 5 and Chapter 8 are based on previously published articles. Chapter 5.7 contains material from my paper “A dilemma for Parfit’s conception of normativity”, published in: *Analysis* 72 (3), 466–74 (2012). Chapter 8 contains material from my paper “‘Ought’ and the Perspective of the Agent”, published in: *Journal of Ethics and Social Philosophy* 5 (3), 1–24 (2011).

Berlin, October 2013

## Introduction: the normativity of rationality

---

Sometimes people are irrational. Occasionally, we have intentions for actions that we believe to be incompatible with each other, or beliefs that are inconsistent. Or we have beliefs that we do not believe to be supported by our evidence. Or we believe we should act in a certain way, but have no intention to do so. Is there anything wrong with that? *Should* we be rational rather than irrational? In other words, is rationality normative? This is the main question that I seek to answer in this thesis.

Intuitively, the answer to this question is “yes”. Calling someone irrational amounts to a form of criticism. By doing so, we seem to imply that the person in question has made some kind of mistake, that her state is in need of revision. Ordinary attributions of irrationality thus seem to presuppose that rationality is normative. This understanding is also implicit in many philosophical approaches to rationality.<sup>1</sup> In recent discussion, however, the assumption that we ought to be rational has come under attack. In his seminal paper “Why be rational?”, Niko Kolodny (2005) argues that it faces insurmountable problems and has to be given up. While remaining unconvinced by Kolodny’s arguments, John Broome (2007a), too, raises sceptical doubts about the normativity of rationality, and argues that several attempts to establish it must fail.

This thesis defends the normativity of rationality by presenting a new solution to the problems that arise from the common assumption that we ought to be rational. The argument touches upon many other topics in the theory of normativity, such as the form and the content of rational standards or requirements, the preconditions of criti-

---

<sup>1</sup> For one, Kant (1785), in his attempt to show that the moral law is universally binding by establishing that moral requirements are imperatives of rationality, arguably takes it for granted that rationality is normative. For another, Williams (1979), in his argument against the Kantian approach, trying to demonstrate that rational deliberation is constrained by contingent motivations, also takes it for granted that rationality is normative: He presupposes that if it could be established that moral requirements are really requirements of rationality, then morality would be normatively binding for everyone.

cism, and the function of reasons in deliberation and advice. Over and above an extensive and careful assessment of the problems discussed in the literature, the chapters to follow provide a detailed defence of a reason-response conception of rationality, a novel, evidence-relative account of reasons, and an explanation of structural irrationality in terms of these accounts.

My purpose in this introductory chapter is to explain in more detail the main question that I seek to answer in this thesis, to lay out the most important assumptions on which my argument is based, and to describe the basic problem that I aim to solve. I begin by introducing the normative question about rationality and by elucidating what it means to ask this question (1.1). I then outline the basic picture of normativity that I will be presupposing, discussing the concepts ‘ought’ and ‘reason’ and their function in deliberation about what to do or what to believe (1.2). Next, I turn to the relation between the phenomenon of structural irrationality (1.3) and the idea of structural requirements of rationality (1.4). This discussion allows me to state the basic problem that will occupy me in this thesis, and to describe a range of views that have been advocated with respect to that problem (1.5). I end by pointing out the main difficulties of these views, indicating the way in which my account accommodates these difficulties, and giving a brief preview of later chapters (1.6).

## 1.1 The normative question about rationality

Rationality, in the sense that I shall be concerned with, comes with certain norms or standards of correctness. This becomes clear when we note the difference between two senses of the word ‘rational’: one that contrasts with ‘arational’ or ‘nonrational’, and one that contrasts with ‘irrational’. Sometimes, when we describe something as rational we mean that it has the capacity of rationality, or that it is an activity of that capacity. (By ‘rational capacity’ we might just mean a certain set of capacities – such as the capacities of reflection and reasoning – that can be subsumed under the capacity of rationality.) I shall call this the *capacity sense* of ‘rationality’. I am a rational being, and my activity of thinking about the issues in this thesis is a rational activity in this sense. The capacity sense of ‘rational’ contrasts with ‘arational’, not with ‘irrational’: a plant

is not irrational because it lacks the rational capacity, nor are the activities of my digestive system irrational because they are not activities of my rational capacity. And when Aristotle and other philosophers maintained that human beings are rational beings, they did not want to say that human beings are never irrational, but only that they have the rational capacity.<sup>2</sup>

But sometimes we describe beings or activities as ‘rational’ in a sense that contrasts with ‘irrational’ rather than ‘arational’. In such cases, we are not ascribing the property of having (or belonging to) a certain capacity. Instead, we are assessing the entity in question relative to a certain standard that applies only to entities that have the rational capacity (or, derivatively, activities of that capacity) – a standard that beings with a rational capacity can either satisfy or violate. As Donald Davidson puts it, “irrationality is a failure within the house of reason”<sup>3</sup> – it is the violation of a standard that applies to rational beings. The sense of ‘rational’ that contrasts with ‘irrational’ is thus essentially linked to the idea of a standard of correctness; we can call it the *standard-related sense*.<sup>4</sup> Since I shall almost exclusively be concerned with this standard-related sense, I will be using the term ‘rational’ without qualification to refer to this sense.

John Broome and others call such standards of correctness ‘requirements’, and I shall follow this terminology.<sup>5</sup> Irrationality, we can then say, consists in, or at least entails, the violation of a rational requirement. As Broome points out, there are other such standards or requirements. Etiquette, morality, the law, language, and many games

---

<sup>2</sup> Cf. Aristotle (Pol. I, 1253a7-10).

<sup>3</sup> Davidson (1982, 169).

<sup>4</sup> The capacity of rationality is also called ‘reason’; both terms are used interchangeably. Accordingly, I do not see the need for distinguishing between rationality in the standard-related sense and *reasonableness*, as some other authors do (see e.g. Scanlon 1998, 32–33). Persons are rational or reasonable to the degree they conform to the standards that apply to them as beings endowed with the capacity of rationality or reason. One could understand ‘reasonable’ as stemming from the count term ‘reason’ rather than the mass term ‘reason’, meaning ‘accordance with reasons’ rather than ‘accordance with the standards of rationality’. If accordance with reasons is different from accordance with rationality, this would provide a distinction between rationality and reasonableness. In German, however, ‘reasonable’ is ‘vernünftig’, clearly stemming from the mass term ‘reason’ (‘Vernunft’), not from ‘reasons’ or ‘a reason’ (‘Gründe’/ ‘Grund’). I take this as evidence that reasonableness is primarily related to the capacity of reason. At any rate, in what follows I take the mass term ‘reason’ to be equivalent with ‘capacity of rationality’, and I shall not distinguish between standards of reasonableness and standards of rationality. To avoid confusion, I will dispense with the term ‘reasonableness’ and the mass term ‘reason’, speaking of ‘rationality’ only.

<sup>5</sup> See esp. Broome (2007b).

all set up certain standards, according to which some actions or responses count as correct and others as incorrect with respect to that standard. I shall say more about the nature of such requirements in the next chapter, but for now it is only important to note that there are various other requirements besides the requirements of rationality.

The question I am going to pursue in this thesis is whether the requirements of rationality are normative. By this, I mean that the requirements are necessarily accompanied by *normative reasons* to conform to them. There are many requirements that are not normative in this sense. For example, the rules of the German card game Skat require me to distribute the cards in the order 3, 4, and 3, and the Chicago Manual of Style requires me to use the Oxford comma, but I may have no reason to conform to these requirements. I can accept that the standards in question exist and still sensibly ask whether I should conform to them. Drawing on Christine Korsgaard, we can call this question “the normative question”.<sup>6</sup> To ask the normative question is not to ask whether there is some standard, according to which some course of action counts as correct. It is to ask whether a given standard is authoritative or binding for a person. And this, I think, is best understood in terms of normative reasons. I shall call this claim the *reasons claim*. It states that for all agents A and all kinds of responses  $\phi$  that an agent can give (such as actions, intentions, or beliefs):

*The reasons claim:* A is normatively required to  $\phi$  if, and only if, A has decisive reason to  $\phi$ .

The relevant reasons must be on balance decisive because ‘being normatively required’ is a strict, overall notion rather than a contributory notion.<sup>7</sup> However, I shall allow that a particular requirement can be normative already if it corresponds to a contributory, or *pro tanto*, reason. More precisely, a requirement R is normative if, and only if, necessarily, if R requires A to  $\phi$ , then A has a reason to  $\phi$ .<sup>8</sup>

---

<sup>6</sup> Korsgaard (1996, ch.1). See also Broome (2005a, 325; 2007a, 163).

<sup>7</sup> See Dancy (2004a, ch. 2) for a valuable discussion of the distinction between overall and contributory normative concepts.

<sup>8</sup> The careful reader will note the implication that it is conceptually possible for an agent to be required to  $\phi$  by a normative requirement without being normatively required to  $\phi$ . In such a case, the normativity of the

According to the reasons claim, the normative question asks for reasons. Once such reasons are given, once there is agreement that there are decisive reasons for giving a certain response, the normative question is answered. There is no further sensible question “Why should I do what I have decisive reason to do?”, for this question itself asks for reasons and thus presupposes the very thing it is questioning: the normative authority of reasons.<sup>9</sup>

Some requirements are normative in this sense, and others are not. The requirements of morality are plausibly normative; it seems that we necessarily have reasons to act as morality requires (although this claim is disputed in moral philosophy).<sup>10</sup> The requirements of etiquette, in contrast, are plausibly not normative; it seems that we do not necessarily have reasons to conform to etiquette requirements (which is not to deny that we often have instrumental reasons stemming from moral or prudential considerations to conform to them). Rationality, like morality and etiquette, comes with standards of correctness, but this does not by itself mean that rationality is normative. It leaves open the normative question: Do we have reasons to satisfy the requirements of rationality? Is there anything *wrong* with being irrational?

## 1.2 Normativity, reasons, and ‘ought’

My understanding of the normative question is based on the reasons claim, and thus involves the assumption that normative authority – the bindingness of a requirement –

---

requirement is only contributory even though the requirement itself is strict. Example: Suppose (for the sake of argument) that there is necessarily a reason to conform to the law, but this reason can be outweighed by other reasons. If you are legally required to stop at the red light, you have reason to do it, but if you are taking someone to the hospital, this reason is outweighed (let us assume). In such a case, you are not normatively required to stop, but you are still legally required to stop, and this legal requirement is (*ex hypothesi*) normative in the sense that we necessarily have a reason to conform to it. Note that by a “normative requirement”, I mean a requirement that is normative; I am not referring to those specific wide-scope principles that Broome used to call “normative requirements” (Broome 1999), but no longer does (cf. Broome 2007b, 4).

<sup>9</sup> This is not to say that the existence of reasons, or normative authority itself, cannot coherently be questioned, although a number of authors have presented arguments to this effect (see Husi 2013 for a critical discussion). I am here only making the weaker claim that once the existence of a reason has been accepted, its normative authority cannot coherently be questioned.

<sup>10</sup> For an example of scepticism about the normativity of morality, see Nielsen (1984). For an example of a theory according to which we always have *sufficient* reason to be moral, see Gert (2005, ch. 13). For an example of a theory according to which we always have overall *decisive* reason to be moral, see Baier (1978). (For the distinction between sufficient and decisive reasons, see below).

is a matter of reasons. This can be seen as a variant of what Derek Parfit calls a “reason-involving conception of normativity”, and which is now a very common conception of normativity among philosophers working in the field.<sup>11</sup> It is worth noting that the reasons claim leaves open a great number of controversial questions in the philosophy of normativity. First, it leaves open whether or not Humeans are right to insist that practical reasons are conditional on desires or other conative states.<sup>12</sup> Second, it leaves open whether normative judgements are expressions of cognitive or non-cognitive states, and whether, if they are expressions of cognitive states, what they express should be given a realist or an anti-realist metaphysical interpretation.<sup>13</sup> Third, the reasons claim is also neutral as regards the possibility of a semantic or metaphysical reduction of statements or facts about reasons, be it in naturalist or in non-naturalist terms.<sup>14</sup> Despite all these philosophical disagreements, proponents of very different views can all agree that the normative question asks for reasons, and that sometimes this question has an answer.

There are further controversies over reasons, and the arguments presented in this thesis are not always neutral with respect to all of them. In the remainder of this

---

<sup>11</sup> See Parfit (2011a, 144–148; 2011b, 267–269). Compare Raz (2010a, 5): “All normative phenomena are normative in as much as, and because, they provide reasons or are partly constituted by reasons.” The reason-involving conception of normativity is also explicit in Raz (1999a, 67), Schroeder (2007a, 81), and Skorupski (2007; 2010, ch.4), and somewhat less explicit in Dancy (2004a) and Scanlon (1998, ch.1). For opposition to this view, see Thomson (2008, 125–6). Note that despite his subscription to the reason-involving conception of normativity, Parfit maintains that there are ‘oughts’ that do not correspond to any reasons, which seems incompatible with the reasons claim. As I argue elsewhere, however, this thesis creates a dilemma for his view and should be given up (Kiesewetter 2012).

<sup>12</sup> Schroeder (2007a) and arguably also Williams (1979) can be seen as Humean proponents of the reason-involving conception of normativity, whereas Parfit (1997; 2011a, ch. 1–4), Raz (1998) and arguably also Nagel (1970) can count as anti-Humean proponents.

<sup>13</sup> As I tend towards cognitivism, I frequently speak of normative judgements as beliefs, but I think that the major points of this thesis could be made in more expressivist-friendly terms. Gibbard’s expressivist account of normativity arguably embraces a reason-involving conception (Gibbard 1990, see esp. 160–164). Skorupski (2010, esp. ch. 17) subscribes to a cognitivist anti-realist (or ‘irrealist’) account of normativity, which is reason-involving. See also Street (2008) for a different anti-realist account of reasons. Cognitivist realist accounts include Parfit’s (2011b) and Schroeder’s (2007a).

<sup>14</sup> The reason claim is compatible with the view that the notion of a ‘reason’ is primitive or indefinable (as maintained, among others, by Parfit 2011a, 31; Raz 2009a, 18; and Scanlon 1998, 17), and the view that facts about reasons are irreducible (Parfit 2011b, ch. 27). But it can also be embraced (and has been embraced) by philosophers who think that the notion of a reason can be defined (such as Broome 2004a, who defines reasons in terms of ‘ought’, which he takes to be the primitive normative concept), and those who think that facts about reasons are reducible to natural facts (such as Schroeder 2007a, ch. 3.4–4; and Smith 1994, ch. 5).

section, I shall lay out the picture of reasons that I find most plausible and that serves as a basis for the discussion to follow. It is generally agreed that reasons *count in favour* of (or against) what they are reasons for (or against).<sup>15</sup> I shall say a bit more about the way in which reasons count in favour presently, but let me first focus on the relation involved in the reasons relation.

### *The reasons relation*

I take reasons to be ordinary facts (or true propositions)<sup>16</sup> that stand in a certain relation to an agent and a type of response that the agent can give. By “response”, I mean something that can be done for a reason, such as an action, a believing, or an intending.<sup>17</sup> Here are some examples: The fact that skiing is a lot of fun may be a reason for you to intend to spend your winter holiday in the Alps. The fact that it reduces the risk of a severe injury may be a reason for you to wear a helmet when skiing downhill. The fact that your neighbour has promised to look after your plants while you are away may be a reason for him to go water them. The fact that the plants in your apartment have gone limp may be a reason for you to believe that your neighbour is not very reliable.

The picture underlying this natural way of thinking and talking about reasons is fairly widely accepted, but it involves some substantive assumptions, a satisfactory discussion of which would go beyond the scope of this thesis. Some philosophers might deny that all reason relations are triadic relations involving an agent, holding that there can be reasons unrelated to any agent.<sup>18</sup> I am sceptical that such reasons exist. The fact that wearing a helmet reduces the risk of a severe injury is not a reason to wear a helmet, it seems to me, unless it is a reason *for someone* to wear a helmet. So-called “agent-

---

<sup>15</sup> See e.g. Broome (2013, 54), Dancy (2000a, 1), Finlay (2006, 5), Gibbard (2003, 188–89), Parfit (2011a, 31), Raz (2009a, 18), Scanlon (1998, 17), Thomson (2008, 127). Needless to say, what is agreed upon is that this is a necessary condition, not that this is a sufficient condition, or that this is all that is to be said about the notion of a reason.

<sup>16</sup> I remain neutral on the questions of whether facts are different from true propositions, and whether, if they are, reasons are one or the other. See Alvarez (2010, 151–8) for a helpful discussion.

<sup>17</sup> My focus will be on these three kinds of responses, but this is not meant to exclude the possibility of reasons for e.g. desires or emotions, as long as desires and emotions are seen as possible responses to acknowledged reasons.

<sup>18</sup> Schroeder (2007b) discusses such views, but only to reject them.

neutral” reasons are better understood as reasons for all agents or reasons for all agents that satisfy certain criteria rather than reasons unrelated to any agent.<sup>19</sup> In any case, I assume that reasons unrelated to any agent are irrelevant for answering the normative question. I am here concerned only with reasons related to agents.<sup>20</sup>

Some philosophers think that reasons favour propositions rather than response-types.<sup>21</sup> According to them, the risk of an injury does not actually count in favour of the action of wearing a helmet, but in favour of the proposition *that you wear a helmet*. The issue is only briefly discussed in Chapter 3, but when it is relevant elsewhere I make it explicit.

Furthermore, there are philosophical theories that take reasons to be mental states rather than facts or true propositions, and others allow that false propositions can be reasons.<sup>22</sup> Some of these theories are concerned with motivating or other explanatory reasons rather than normative reasons, i.e. reasons that explain in some way or other why some agent showed a particular response, which need not be reasons that do in fact speak in favour of the response. In this thesis, I am concerned with normative reasons, and I remain neutral with respect to different theories about motivating and explanatory reasons.<sup>23</sup> I take it, however, that the most natural view about *normative* reasons is that they are facts. In ordinary discourse, we frequently refer to reasons by citing propositions or ‘that’-clauses (“that you have promised to look after the plants is a reason to go water them”). Such reasons cannot equally well be expressed by citing some mental state, such as e.g. the belief that one has promised to look after the plants. First, the statement “that you have promised to look after the plants is a reason for you to go water them” has a different content than the statement “your belief that you have

<sup>19</sup> See Schroeder (2007b). Broome (2013, 66) takes the same view.

<sup>20</sup> I should say that I do not distinguish between “There is reason for A to  $\phi$ ” and “A has a reason to  $\phi$ ”. Some authors use the second locution only to refer to reasons that are epistemically accessible in one way or another; see e.g. Lord (2010), Schroeder (2008) and Skorupski (2010, 113). But this use of these expressions is stipulative, it is not suggested by standard English. I discuss questions of epistemic accessibility in Chapters 7 and 8.

<sup>21</sup> See e.g. Broome (2013, 66). The view is explicitly discussed with respect to ‘ought’ rather than reasons in Broome (2013, 15–18).

<sup>22</sup> For an example of the former view, see Davidson (1963). For examples of the latter view, see Henning (forthcoming), Schroeder (2007a, 14; 2008; 2011a) and Vogelstein (2012, 250). The view that false propositions can be normative reasons seems also implicit in claims made by Joyce (2001, 53) and Way (2009, 3–4).

<sup>23</sup> See e.g. Alvarez (2010), Dancy (2000a), Schroeder (2007a, ch. 1.3) and Smith (1994, ch. 4) for discussions of the relation between normative and motivating reasons.

promised to look after the plants is a reason for you to go water them”. Second, the former statement may be true without there being any belief of the sort presupposed by the latter statement. For example, one could truly utter the former sentence in advising someone who has forgotten that he has promised to look after the plants. More generally, ordinary reasons talk presupposes that reasons can be disclosed in advice or discovered in deliberation, which implies that these reasons were there before one formed beliefs about them.<sup>24</sup>

Conversely, however, whenever a mental state is cited as a normative reason, we can without loss reformulate this reason as a proposition or fact.<sup>25</sup> The statement “*your belief* that you have promised to look after the plants is a reason to go water them” can equally well be expressed as “*that you believe* that you have promised to look after the plants is a reason to go water them”. So if normative reasons belong to a single ontological category, then this category will be propositions or facts. This allows facts about mental states to be reasons; whether and in what context this might be so is here understood as a first-order normative question.

Once a proposition  $p$  is accepted as false, a statement to the effect that  $p$  is a reason to  $\phi$  is usually retracted or regarded as falsified as well. If your neighbour has not promised to look after the plants, then that he has promised is also not a reason for anything. Contrary views, according to which false propositions can be normative reasons, are discussed where they become relevant in the course of this thesis.<sup>26</sup> Such views are often motivated by the important insight that normative claims have to take the epistemic limitations of agents into account. Chapter 8 proposes a theory of reasons that embraces this insight while preserving the intuitive idea that reasons are facts.

---

<sup>24</sup> These phenomena are discussed at length in Chapter 8.

<sup>25</sup> This is a general point, emphasised by Schroeder (2007a, 20–21). Whenever something other than a proposition is cited as a reason, a proposition can equally well be cited to make the same point, while the converse is not true. For example, all statements citing persons as reasons (“Clare Daines is a reason to watch *Homeland*”) could just as well be expressed by statements citing propositions (“that Claire Daines is starring in it is a reason to watch *Homeland*”), but obviously not all statements citing propositions as reasons can equally well be expressed by statements citing persons.

<sup>26</sup> See e.g. Chapters 4.3 and 7.4.

*Reasons: pro tanto, sufficient, decisive*

Now, what does it mean to say that a fact *counts in favour* of a certain response, such as an action, an intention, or a belief? Part of what it means is that it matters in deliberation. Deliberation, or at least one important kind of deliberation, is concerned with settling what to do, to intend, or to believe, by finding out what you *ought* to (or *may*) do, intend, or believe. Reasons play a contributory, or *pro tanto*, role in such reflections; they are factors that help to determine what you ought to do or may do.<sup>27</sup> Reasons typically come with different strengths and point in different directions; they often have to be weighed against each other. If you have overall *sufficient reason* to  $\phi$ , then your  $\phi$ -ing is justified or normatively permitted (you may  $\phi$ ). If you have overall *decisive reason* to  $\phi$ , then your  $\phi$ -ing is mandatory or normatively required (you ought to  $\phi$ , or should  $\phi$ ). (I take ‘ought’ and ‘should’ to be equivalent, but for simplicity’s sake, I will mostly use ‘ought’.) This ‘ought’ need not have a moral connotation, as the reasons relevant in deliberation can be reasons of all sorts (encompassing moral, but also prudential, hedonic, and – not to forget – epistemic reasons). It may well be that you ought to do something because it is a lot of fun.

Can the notions ‘sufficient reason’ and ‘decisive reason’ be defined in terms of the relative weight or strength of *pro tanto* reasons? We might say that an agent A has *sufficient reason* to  $\phi$  just when A has a reason to  $\phi$ , and no stronger reason not to  $\phi$ , and A has *decisive reason* (or *ought*) to  $\phi$  just when A’s reasons to  $\phi$  are stronger than A’s reasons not to  $\phi$ .<sup>28</sup> But this is not uncontroversial. First, the definition implies that all

<sup>27</sup> Broome (2004a; 2013, ch. 4) suggests that reasons can be analysed in terms of their explanatory role for ‘oughts’, whereas Dancy (2004a, ch. 2) and Schroeder (2007a, 130–131) suggest an analysis of ‘ought’ in terms of reasons. I accept that there is a strong connection between reasons and ‘ought’, but remain neutral with respect to the question of whether either of the concepts can be analysed in terms of the other. Kearns and Star (2008; 2009) and Thomson (2008, ch. 8-9) propose that the notion of a reason can be analysed in terms of *evidence* for an ‘ought’-proposition. I briefly discuss (and reject) this suggestion in Chapter 7.8.

<sup>28</sup> See e.g. Schroeder (2007a, 193): “For it to be the case that X ought to do A, is for it to be the case that  $S_{X,A} > S_{X,-A}$ , where  $S_{X,A}$  is the set of all the reasons for X to do A, and  $S_{X,-A}$  is the set of all the reasons for X to not do A.” As the definitions mentioned above, this is plausible only on the assumption that a reason for a response that is incompatible with  $\phi$ -ing is also a reason against  $\phi$ -ing. If this assumption is denied, a similar definition can be given in terms of reasons for incompatible alternative responses. See e.g. Schroeder (unpublished, 6): “A set S of reasons for X in favor of A is *sufficient* just in case for each alternative B to A, S is at least as weighty as the set of reasons for X in favor of B. A set S of reasons for X in favor of A is *conclusive* just in case for each alternative B to A, S is weightier than the set of reasons for X in favor of B.” Compare also Parfit (2011a, 32).

reasons are *presumptively sufficient*, i.e. that every reason is sufficient unless it is defeated by a countervailing reason. While this seems plausible in the case of practical reasons, it is disputable in the case of epistemic reasons. Some authors hold that one can have an insufficient reason for a belief even if no reason against the belief is present.<sup>29</sup> Second, the definition also implies that all reasons are *presumptively requiring*, i.e. that every reason is decisive or amounts to an ‘ought’ unless it is counterbalanced by a countervailing reason. This can already be disputed in the case of practical reasons. One might question that every little reason that makes it the case that you have on balance slightly more reason to  $\phi$  than not to  $\phi$  therefore makes it the case that you ought to  $\phi$ . Perhaps there are reasons that merely justify but never require responses. Or perhaps in order to require responses, reasons need to be more than merely slightly stronger. For example, perhaps it is possible for you to have slightly stronger reason to go to the movies tonight than reason to do something else, and yet your reasons are not decisive (perhaps you would still have sufficient reason, or would be justified, to do something else).<sup>30</sup> Possibly the definitions mentioned above can be defended against such objections; I am not committed to rejecting them. But I will not presuppose them either. I will assume that we have an intuitive grasp of the notions ‘sufficient reason’ (i.e. the idea that reasons justify a response) and ‘decisive reason’ (i.e. the idea that reasons require a response).

### *The deliberative ‘ought’*

We can call the ‘ought’ that is provided by what we have decisive reason to do the *deliberative ‘ought’*, because it is the notion of ‘ought’ that figures in the question “what

---

<sup>29</sup> See Raz (2009b, 45–47); I borrow the term “presumptive sufficiency” from him. For a defence of the view that epistemic reasons are presumptively sufficient, see Schroeder (unpublished). Raz’s position seems inevitable given his assumption that “any reason against believing that a proposition is true is a reason for believing that it is false” (2009b, 36, n. 3), for it is clear that one can have an insufficient reason to believe  $p$  even if one has no reason for believing non- $p$ . Thus, much in this dispute hangs on the question of what one allows as a reason *against* belief. Some authors are even sceptical that there are any reasons against believing (as contrasted with reasons for believing the negation) at all (see Skorupski 2010, 39, n. 9).

<sup>30</sup> Taking up a suggestion that Raz (1999b, 101–2) brings up but discards, Dancy (2004b) defends the idea that there are reasons (“enticing reasons”), which by their very nature make responses attractive without ever requiring them or generating ‘oughts’. For a reply on behalf of the view that all reasons are presumptively requiring, see Broome (2013, 60–61).

ought I to do?”, as it is asked from the standpoint of deliberation, and in deliberative conclusions or all-things-considered judgements of the form “I ought to  $\phi$ ”.<sup>31</sup> As the notion of a ‘reason’, this notion involves a relation between an agent and a response-type. There are other senses of ‘ought’, some of which do not involve such a relation at all, but refer to very different ideas. For one, there is an epistemic sense of ‘ought’, as it occurs in “the train ought to arrive at 8 a.m.”. This statement is badly understood as relating an agent (the train) to a response (arriving at 8 a.m.). Rather, it means roughly that it is likely that the train will arrive at 8 a.m.<sup>32</sup> For another, there is an evaluative sense of ‘ought’, as it occurs in “people ought not to starve”. This statement is also badly understood as relating agents (people) to a response (not starving) they have reason to give. Rather, it says that it would be good or desirable if a certain state of affairs occurred, namely the state of affairs in which people would not starve.<sup>33</sup>

There are also qualified senses of ‘ought’ such as the moral ‘ought’ or the prudential ‘ought’, which *do* relate agents to responses. A statement involving such a qualified sense of ‘ought’ (such as “morally speaking, one ought not to travel by air” or “from the standpoint of your own personal interest, you ought to file a tax return”) might be understood in different ways. It could be interpreted as stating that a response is favoured by the balance of reasons of a particular kind (moral reasons, prudential reasons); or that it is favoured by the balance of reasons of a particular kind and not counterbalanced by any other reasons; or that it is required by some standard of correctness for which the normative question is essentially open.<sup>34</sup> In any case, judge-

<sup>31</sup> The term “deliberative ‘ought’” is used, among others, by Schroeder (2011b, 2).

<sup>32</sup> See Broome (2013, 9–11) and esp. Thomson (2008, 199–205) for discussion.

<sup>33</sup> In line with the so-called Meinong/Chisholm analysis, some authors think that the deliberative ‘ought’ can be reduced to the evaluative ‘ought’: “A ought to  $\phi$ ” means the same as “It ought to be the case that A  $\phi$ -s”; the deliberative ‘ought’ can be interpreted as a sentential operator and does not require an argument place for an agent (see e.g. Chisholm 1964; Williams 1980; for extensive discussion, see Horty 2001). Broome (2013, 12–24), J. Ross (2010), and Schroeder (2011b) – partly drawing on considerations first put forward by Harman (1973) and Geach (1982) – argue convincingly that the reduction fails; the deliberative ‘ought’ essentially involves an argument place for an agent. However, a reduction might work the other way round. Supposing that value can be analysed in terms of reasons for attitudes of appraisal, as the so-called “buck-passing” or “fitting attitude” account of value claims, the evaluative ‘ought’ could be analysed in terms of the deliberative ‘ought’: “it ought to be the case that p” might be analysed as “we ought to wish that p is the case”, for example.

<sup>34</sup> The first of these interpretations is suggested by Schroeder (2007a, 130–31, n. 6); the third by Broome (2013, 25–29). I have sympathies with the second interpretation.

ments involving a qualified sense of ‘ought’ are not conclusions of deliberation as long as it is not assumed that the particular qualified ‘ought’ entails decisive reasons.

This becomes clear if we pay attention to an important feature of deliberative conclusions, namely that their point is to guide responses. A rational agent must be capable of guiding her responses on the basis of judgements about what she ought to do in the deliberative sense. More specifically, the deliberative ‘ought’ is the ‘ought’ appealed to in the common idea (which I will return to presently) that it is irrational not to intend what one believes one ought to do.<sup>35</sup> This need not be true for qualified senses of ‘ought’. For example, it might not be irrational for you to refrain from intending to  $\phi$  if you believe that you *morally* ought to  $\phi$ , as long as you do not believe that you have decisive reason to do what you morally ought to do.<sup>36</sup>

#### *The guiding role of reasons and the state-given/object-given distinction*

The role of reasons is not restricted to their contributory function for all-things-considered judgements about what we may or ought to do. Acknowledgement of a single reason can guide a response, and thereby make it intelligible, independently of a judgement to the effect that there is sufficient or decisive reason for the response; in fact, it can do this even when the reason is acknowledged to be insufficient.<sup>37</sup> As we might put it, reasons do not only *justify* and *require*, but also *render intelligible*, and they might do the latter without doing the former.<sup>38</sup> If  $p$  is a reason for you to  $\phi$ , then your acknowledgement of  $p$  can guide your  $\phi$ -ing, you can  $\phi$  *for the reason* that  $p$ , and if you do, your  $\phi$ -ing has a point, or is intelligible. For example, if we say that the fact that your plants have gone limp is a reason for you to believe that your neighbour is unreli-

<sup>35</sup> This is how Broome (2013, 22–25) identifies what he calls the “central ought”; see also J. Ross (2012a, 163–64).

<sup>36</sup> See also Broome (2013, 28).

<sup>37</sup> In such a case, the response is arguably irrational, but irrationality does not imply unintelligibility.

<sup>38</sup> Cf. Raz (2005a, 8–9): “One aspect of reasons for actions is that they make choices, intentions, and actions intelligible both to the agent and to others”. My claim is that every normative reason has the potential to render a response intelligible, but I agree with Broome (2005b, 8) that the intelligibility of a response does not conversely entail that there was a normative reason for the response. However, it might still be true that the intelligibility of a response is necessarily related in some way or other to normative reasons. For example, it might be true that a response cannot be intelligible unless the agent took himself to have a reason for it (though I do not here assume that this is true).

able, we partly seem to be saying that your believing this would make sense in light of this fact, and would be intelligible if guided by your acknowledgement of this fact, while we would not usually think that e.g. believing that God exists makes sense in light of this fact or is rendered intelligible by your acknowledgement of it.

That reasons are capable of guiding the responses they are reasons for entails that there is an important connection between normative and explanatory reasons, as Bernard Williams and, more recently, Joseph Raz have emphasised. As Raz puts it:

If that R is a reason to  $\phi$  then it must be possible that people  $\phi$  for the reason that R and when they do, that explains (or is part of an explanation of) their action.<sup>39</sup>

This has a crucial implication for the ongoing debate over the question of whether there are *state-given* reasons for attitudes or whether all such reasons are *object-given*. Roughly, a reason for an attitude is object-given if it is provided by facts that bear on the attitude's *object*, while it is state-given if it is provided by facts that bear on having the attitude, or being in the *state*, itself.<sup>40</sup> Proponents of state-given reasons typically hold that the benefits of being in a mental state provide a reason for it. Accordingly, if you have been promised a reward for having a belief or an intention, then you have a state-given reason for the belief or intention, no matter whether the content of the belief is supported by evidence, and no matter whether the object of the intention – the action – is supported by any reason. I side with those who deny that such reasons exist. The benefits of having an attitude may provide a reason for *wanting* us to have this attitude, or *doing* something that brings about that we have this attitude (such as taking a

---

<sup>39</sup> Raz (2009a, 27). Williams (1979, 102) makes the same point, but with respect to reasons for action only.

<sup>40</sup> The distinction was first introduced by Parfit (2001, 21–25). See also Parfit (2011a, App. A) and Piller's (2006) distinction between "content-related" and "attitude-related" reasons for attitudes. As Schroeder (2012, 462–64) and others have pointed out, the way in which facts have to bear on the object of an attitude or the attitude itself in order to count as object-given or state-given reasons is in need of further clarification. I here have to rely on an intuitive understanding of the distinction. The distinction between state- and object-given reasons is often invoked in the context of what is known as the wrong kind of reasons problem for the buck-passing, or fitting attitude, account of value. For a discussion of these issues, see esp. Rabinowicz and Rønnow-Rasmussen (2004; 2006), but also Olson (2004) and Stratton-Lake (2005).

pill that causes us to have the attitude), but not a reason for the attitude itself.<sup>41</sup> My reason for thinking this lies in the connection mentioned above between normative and explanatory reasons. I take it to be a conceptual truth about normative reasons that they are capable of guiding the responses they are reasons for. Alleged state-given reasons do not satisfy this constraint: we cannot believe or intend something on the basis of the consideration that it would be beneficial for us to have this attitude, and therefore our having this attitude would also not be intelligible solely in the light of our acknowledging that it would be beneficial to have this attitude. We could *want* to have this attitude, or *do* something to bring it about for the reason that it would be beneficial to have the attitude; hence, I have no objection to the claim that the benefits of having an attitude can provide reasons for *these* responses. Yet such reasons would not be state-given, i.e. they would not be reasons for attitudes provided by facts that bear on having the attitude itself rather than its object.

I shall thus assume that reasons for attitudes are generally provided by facts that bear on the object of these attitudes. Roughly speaking, the picture is that reasons for belief are provided by evidence for the truth of their content, and reasons for intention are provided by reasons for the intended action. This is independently supported by the

---

<sup>41</sup> Here I find myself in agreement with a number of other authors, including Parfit (2011a, App. A), Skorupski (2010, 54–55), and Way (2012a). Raz (2009b) seems to take the same view with respect to reasons for beliefs and emotions, but (apparently) not with respect to intentions. He first distinguishes *standard* and *non-standard reasons*: “Standard reasons are those that we can follow directly, that is have the attitude, or perform the action, for that reason. Non-standard reasons for an action or an attitude are such that one can conform to them, but not follow them directly” (2009b, 40). In a section entitled “Special practical reasons for beliefs and emotions?”, he argues that “all non-standard reasons are practical reasons. [...] a non-standard reason to believe that P is a standard reason to bring it about that I believe P. A non-standard reason to  $\phi$  is a standard reason to bring it about that I  $\phi$ ” (2009b, 57). In the case of beliefs and emotions, he seems to take this identification to be eliminative: “the reasons are not *really* reasons for the emotions, since they are reasons for action to bring it about that one has the emotions” (2009b, 57). It follows that there are “not really” any non-standard reasons for emotions (and, by the same token, for beliefs). And yet Raz seems to insist that there are non-standard reasons for actions and intentions, for he begins the section by stating: “we saw that there are standard and non-standard reasons for actions and standard and non-standard reasons for having intentions” (2009b, 56, see also the discussion on pp. 50–53). So Raz’s overall view seems to be that there are no non-standard reasons (and hence no state-given reasons) for beliefs and emotions, but there are non-standard reasons for action and intention (including state-given reasons for intentions). This is quite surprising given what he says about the connection between normative and explanatory reasons. It seems to have slipped his attention that the assumption of any non-standard reason – which is, to recall, a reason that one cannot “follow directly, that is have the attitude, or perform the action, for that reason” (2009b, 40) – is logically inconsistent with his claim that “if that R is a reason to  $\phi$  then it must be possible that people  $\phi$  for the reason that R” (2009a, 27).

phenomenology of deliberation: in deliberation we do not seem to treat reasons for believing  $p$  as being different from evidence for  $p$ , and reasons for intending to  $\phi$  as being different from reasons for  $\phi$ -ing.<sup>42</sup> The question of whether there are state-given reasons arises at a number of places in this thesis, and I shall make it explicit when my rejection of them plays a role in the argument.

### *Three dimensions of normativity*

In this thesis, I take the first-personal, deliberative dimension of normativity to be fundamental: a claim to normative authority must ultimately be grounded in reasons that matter in first-personal deliberation. But there are other important dimensions of normativity. Most notably, there is the second-personal dimension of *advice* and the third-personal dimension of *criticism*.<sup>43</sup> In both advice and criticism, ‘ought’ is employed in a way that seems genuinely normative, relates agents to responses, and is not further qualified. According to some theories, the ‘oughts’ of deliberation, advice, and criticism are different concepts.<sup>44</sup> If this were true, then conflicting ‘oughts’ could apply to agents at a particular time, and it would be in principle impossible to resolve this conflict in deliberation. There could be valid ‘ought’ claims on agents that they would have no reason to act upon. I believe that these implications should make us cautious about accepting the theories in question. According to the view that I shall argue for, the ‘ought’ used in advice and criticism is the same as the deliberative ‘ought’.<sup>45</sup> I take

<sup>42</sup> Moran (1988, 142–148; 2001, ch. 2) famously calls this feature “transparency” in the case of doxastic deliberation. See also Shah (2003; 2008).

<sup>43</sup> Needless to say, criticism can also be raised in the first or the second person. The point of calling the normative dimension of criticism “third-personal” is merely that it is not essentially first-personal (as deliberation) or second-personal (as advice).

<sup>44</sup> See e.g. Gibbard (1990, 162) and Jackson (1991, 471–72) for remarks that suggest a difference between the deliberative ‘ought’ and the ‘ought’ of advisability. Parfit, who distinguishes numerous senses of ‘ought’ (2011a, 150–74), maintains that the ‘ought’ of criticism is different from the deliberative ‘ought’, as well as from the ‘ought’ used in advice (see esp. 2011a, 35–36). Kolodny and MacFarlane (unpublished, §3.1 and §3.3) maintain that the ‘ought’ in deliberation and advice is the same concept, but differs from the ‘ought’ of criticism (see also Kolodny, forthcoming, n. 27). Smith (2008, 255), in contrast, seems to identify the ‘ought’ of criticism with the deliberative ‘ought’, but distinguishes it from the ‘ought’ of advice (cf. Smith 2006, 147). Schroeder (2009, 229) distinguishes the ‘ought’ of advice from the ‘ought’ of criticism, but is silent with respect to how these two ‘oughts’ relate to the ‘ought’ of deliberation.

<sup>45</sup> In Chapter 2, I argue that criticism involves the assumption that decisive reasons have been violated; in Chapter 8 I develop a theory of reasons that accounts for this fact as well as the plausible idea that advice is concerned with answering the same question as deliberation.

it to be a considerable advantage of my view that it can integrate the normative dimensions of deliberation, advice, and criticism, and I do not know of many other theories that can claim this advantage for themselves. However, that the ‘ought’ of deliberation is the same as the ‘ought’ of advice and the ‘ought’ of criticism is not a premise of the argument of this thesis, but one of its conclusions. My starting point is that the normative question is a question that can occur, and naturally does occur, in first-personal deliberation, and that this question asks for reasons.

### 1.3 Structural irrationality

The question of whether rationality is normative is here understood as the question of whether we necessarily have reasons to conform to the standards of rationality. This question may seem puzzling. Often irrationality consists in a failure to respond appropriately to reasons that are available to one.<sup>46</sup> It is irrational, for example, to wishfully believe something in the face of strong counterevidence.<sup>47</sup> Evidence provides reasons for beliefs, and insofar as our irrationality consists in a failure to form our beliefs in accordance with our evidence, we are at the same time failing to respond to our reasons. At least *prima facie*, this kind of irrationality does not pose a problem for the normativity of rationality, since what rationality requires in such cases seems to coincide with our reasons in an obvious way. (I shall come back to such cases of irrationality in Chapter 7.)

The problem begins to arise when we consider the phenomenon of what we might call *structural irrationality*.<sup>48</sup> Structural irrationality is a kind of irrationality that we can detect simply by looking at a particular combination of attitudes that a person holds, independently of any assumptions about the reasons that this person has for holding the particular attitudes in question. It is this kind of irrationality that raises a

---

<sup>46</sup> See e.g. Raz (1999a, 68).

<sup>47</sup> This is a commonplace in the philosophical literature; see e.g. the entry on “evidence” in the *Stanford Encyclopedia of Philosophy*: “insofar as one is rational, one is disposed to respond appropriately to one’s evidence” (Kelly 2006, §2). See also Chapter 7.3 for further discussion and references.

<sup>48</sup> I take this term from Scanlon (2007). Drawing on Kolodny (2005, 515–516), Way (2011, 227) calls patterns such as (A)-(D) “local irrationalities”. My impression is that the word “local” is used in slightly different senses by different participants of the debate, which is why I avoid it.

substantive question about the normativity of rational standards, for it is not obvious that structural irrationality can be understood as a failure to respond correctly to reasons.

I shall focus my discussion on the following four kinds of structural irrationality, which have been much discussed in recent literature:

- (A) *Weakness of will*: If A believes that she (herself)<sup>49</sup> ought to  $\phi$ , and A does not intend to  $\phi$ , then A is irrational.<sup>50</sup>
- (B) *Instrumental irrationality*: If A intends to  $\phi$ , and A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A does not intend to  $\psi$ , then A is irrational.<sup>51</sup>
- (C) *Weakness of belief*: If A believes that she has sufficient evidence that  $p$ , and A does not believe that  $p$ , then A is irrational.<sup>52</sup>

<sup>49</sup> This qualification guarantees that A ascribes the ‘ought’ to herself, i.e. that she is in a position to express her belief by using the first-personal pronoun ‘I’. This is necessary to avoid counterexamples to (A) in which the agent believes that some particular person ought to  $\phi$ , but does not know that she is this person herself (cf. Broome 2013, 22–23). I henceforth take this qualification as implicit.

<sup>50</sup> Compare: “Rationality requires people to intend to do what they believe they ought to do” (Broome 2013, 2, see p. 170 for a more exact statement). “Rationality requires one to intend to X, if one believes that there is conclusive reason to X” (Kolodny 2005, 521). “Irrationality in the clearest sense occurs when a person’s attitudes fail to conform to his or her own judgements: when, for example [...] a person fails to form and act on an intention to do something even though he or she judges there to be overwhelmingly good reason to do it” (Scanlon 1998, 25). “If an agent believes that she has normative reason to  $\phi$ , then she rationally should desire to  $\phi$ ” (Smith 1994, 148). “Necessarily, if one is rational, then, if one judges ‘I ought to  $\phi$ ’, one also intends to  $\phi$ ” (Wedgwood 2007, 25). Similar claims are made by Coates (2013), Cullity (2008, 78), Dancy (2009, 95–96), Davidson (1969, 41), Korsgaard (1986, 11), Mele (1987, 7), Parfit (2011a, 113), J. Ross (2012b, 79–80), Way (2011, 227), Williams (1979, 107), and many others. Arpaly (2000) appears to be arguing against (A), but as she notes herself, she actually endorses the view “that every agent who acts against her best judgment is, as an agent, less than fully rational” (2000, 491). I take her arguments to be directed at certain *versions* of the claim that defying one’s normative judgements is irrational. The condition described in (A) is often called “akrasia”, but as I have argued in Kiesewetter (2011a), it differs significantly from the conception of *akrasia* in Aristotle (E.N. VII ; M.M. II 4–6).

<sup>51</sup> Compare: “There is a rational demand that one’s intentions be means-end coherent in the sense, roughly, that it not be true that one intends E, believes that E requires that one now intends necessary means M, and yet not now intends M” (Bratman 2009, 412). “A person is necessarily irrational if she does not intend whatever she believes is a necessary means to an end she intends” (Broome 2005b, 2). “Wer den Zweck will, will (sofern die Vernunft auf seine Handlungen entscheidenden Einfluß hat) auch das dazu unentbehrlich notwendige Mittel, das in seiner Gewalt ist” (Kant 1785, 46 [BA 44–45]). “People who fail to pursue the means to their ends display or manifest a form of malfunctioning criticisable as a form of irrationality” (Raz 2005a, 18). “Intending to E, believing that M-ing is necessary for E-ing, but failing to intend to M is a form of irrationality” (Way 2012b, 487). Versions of (B) can also be found in Brunero (2005; 2012), Korsgaard (1997), Rippon (2011), Schroeder (2009), Searle (2001, 266), Setiya (2007a), and Wallace (2001), among others.

(D) *Modus ponens irrationality*: If A believes that p, and A believes that  $p \rightarrow q$ , and A does not believe that q, then A is irrational.<sup>53</sup>

It is plausible to think that at least some of these claims are in need of further qualification. For example, Gilbert Harman has famously argued that it is not always irrational to refrain from believing the logical consequences of what one believes, since there is no point in cluttering one's mind with an endless array of irrelevant beliefs.<sup>54</sup> Plausibly, then, (D) needs to be complemented with a condition to the effect that A takes it to be relevant or at least attends to the question of whether q is the case. For another example, Frances Kamm has argued that one can rationally abstain from intending the means to one's end if one believes that one will take the means no matter whether one intends them or not.<sup>55</sup> A similar point seems to apply to weakness of will as well. Plausibly, then, (A) and (B) need to be complemented with a condition that excludes such cases.<sup>56</sup>

These questions will receive more attention in Chapters 7 and 9, but until then I shall work with the simplified versions stated above. Particular amendments of the sorts mentioned are subject to extensive debate, while my concern is with the general phenomenon of structural irrationality rather than with the task of a complete and accurate description of it.<sup>57</sup> What matters for now is only that there are truths in the vicinity of the claims mentioned – truths to the effect that some attitude, or some

<sup>52</sup> Compare: "Rationality requires one to believe that p, if one believes there is conclusive evidence that p" (Kolodny 2005, 521). "If an agent believes that she has (most) reason to believe that p then she rationally should believe that p" (Smith 1994, 178). See also Mele (1987, 112), Scanlon (1998, 25), and Wallace (2001, 13).

<sup>53</sup> Compare: "Rationality requires of N that, if N believes p and N believes that if p then q, then N believes that q" (Broome 2006, 183). "One ought rationally to be such that (if one believes that p and one believes that (if p, then q) then one believes that q" (J. Ross 2009, 278). See also Reisner (2011, 47), Setiya (2007a, 665–66), and Way (2011, 228).

<sup>54</sup> See Harman (1986, 12).

<sup>55</sup> See Kamm (2000, §4).

<sup>56</sup> As Broome (2013, 171–72) points out, (A) needs also to be complemented with a condition that guarantees that A believes that she can intentionally control whether she  $\phi$ -s. For example, A might rationally abstain from intending to believe something that she believes she ought to believe if she does not think that intending to believe it will have an effect on her believing it. Alternatively, ' $\phi$ -ing' in (A) might be restricted to actions.

<sup>57</sup> The most thorough and meticulously argued statement of sufficient conditions for structural irrationality I know of is to be found in Broome (2013, ch. 9-10).

combination of attitude-states, is sufficient for irrationality, independently of whether any of the particular attitude-states are favoured or disfavoured by a reason (I use ‘attitude-state’ as the generic term for both attitudes and *lack* of attitudes)<sup>58</sup>. I shall refer to such claims as *structural irrationality claims*. Even though there is a lot of disagreement about the exact content of such claims, there is wide agreement that there are truths of this general form, and I shall take this for granted as an intuitive starting point.<sup>59</sup>

I have chosen the examples (A)-(D) for two reasons. First, they seem to address paradigmatic cases of irrationality about which there is wide agreement. Second, they represent four different kinds of structural irrationality. Weakness of will and instrumental irrationality are practical forms of structural irrationality, because they are (at least primarily) concerned with intentions. In contrast, being concerned with beliefs only, weakness of belief and *modus ponens* irrationality are theoretical forms of structural irrationality. At the same time, weakness of will and weakness of belief exhibit a specific higher-order structure that distinguishes them from instrumental irrationality and *modus ponens* irrationality. All kinds of structural irrationality are due to some form of incoherence between attitude-states (I use ‘incoherence’ in a broad sense, which covers conflicts not only between attitudes, but also between attitudes and lacks of attitudes). A natural conjecture is that the incoherence involved in instrumental irrationality and *modus ponens* irrationality can be traced back to purely formal relations

---

<sup>58</sup> Here, I follow Brunero (2010a, 28).

<sup>59</sup> As I understand the so-called *myth view* put forward by Kolodny (2008a) and Raz (2005a), it is compatible with the truth of structural irrationality claims such as (A)-(D). As I explain below, I take the myth view to deny certain *requirements* of structural rationality, when “requirement” is understood as a code or principle rather than a necessary condition for being rational. The denial of such requirements is compatible with the claim that certain combinations of attitudes are always irrational; it just poses the challenge to explain this fact without invoking structural requirements of rationality (which is exactly what both Kolodny and Raz attempt to do). Some of Raz’s remarks (see esp. 2005b, §2) seem to suggest that he also questions the structural irrationality claims themselves, though he is not explicit about that, and I do not think this is essential to his view. Other authors have explicitly denied the possibility of structural irrationality, at least for particular areas such as means/end incoherence. For example, Finlay (2009) seems to argue that every combination of attitudes that can plausibly be regarded as instrumentally irrational is at the same time impossible. This is clearly not Kolodny’s or Raz’s concern, however.

among the contents of the (possible) attitudes concerned.<sup>60</sup> In any case, at least *prima facie* these kinds of irrationality are constituted by conflicting attitude-states on a simple, first-order level. In contrast, the incoherence involved in weakness of will and weakness of belief is constituted by a conflict between an attitude-state and a *normative belief* the content of which has implications for the normative status of this attitude-state. In the case of weakness of will, the belief that one ought to  $\phi$  has implications for the normative status of a lack of intention to  $\phi$ . In the case of weakness of belief, the belief that one has sufficient evidence that  $p$  has implications for the normative status of a lack of belief that  $p$ . A natural conjecture is that once it is fully spelled out what a person has to believe in order to be subject to an irrational weakness of either form, these beliefs will entail that the agent has decisive reasons for the attitude-state that she lacks.<sup>61</sup> In any case, since the incoherence in both cases can be described as a conflict between an attitude-state and a belief with implications for the normative status of this attitude-state, I shall call weakness of will and belief *second-order* forms structural irrationality.

Hence, we have practical and theoretical as well as first-order and second-order kinds of structural irrationality. The following table illustrates these distinctions:

<i>Structural irrationality</i>	Practical	Theoretical
Second-order	(A) <i>Weakness of will</i>	(C) <i>Weakness of belief</i>
First-order	(B) <i>Instrumental irrationality</i>	(D) <i>Modus ponens irrationality</i>

<sup>60</sup> This seems clear in the case of *modus ponens* irrationality, but is more controversial in the case of instrumental irrationality. See Broome (2002) for an account that treats these cases of irrationality analogously.

<sup>61</sup> Illustration: For the consideration attributed to Kamm above, it might seem plausible that (A) has to be complemented by the condition that A believes that she will  $\phi$  only if she intends to  $\phi$ . Together, her beliefs that she ought to  $\phi$  and that she will  $\phi$  only if she intends to  $\phi$  plausibly entail that she has decisive reason to intend to  $\phi$ . (The issue is briefly discussed in Chapter 7.9). The idea that the form of irrationality involved in (A) and (C) can be traced back to a conflict between a normative belief about reasons for an attitude and lack of this attitude is present in Scanlon (1998, 25), and especially in Kolodny's conception of what he calls the "core requirements" of rationality (2005, 524).

## 1.4 Structural requirements of rationality

If there is structural irrationality of the kind described in (A)-(D), and if, as suggested above, irrationality consists in the violation of rational requirements, then it seems natural to suppose that there are *structural requirements of rationality* which correspond to (A)-(D):<sup>62</sup>

- (A)<sub>1</sub> *Self-governance requirement*: Rationality requires A to intend to  $\phi$  if A believes that she ought to  $\phi$ .
- (B)<sub>1</sub> *Instrumental requirement*: Rationality requires A to intend to  $\psi$  if A intends to  $\phi$ , and A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing.
- (C)<sub>1</sub> *Evidence-belief requirement*: Rationality requires A to believe that  $p$  if A believes that she has sufficient evidence that  $p$ .
- (D)<sub>1</sub> *Modus ponens requirement*: Rationality requires A to believe that  $q$  if A believes that  $p$ , and A believes that  $p \rightarrow q$ .

It is important to note, however, that the step from (A)-(D) to (A)<sub>1</sub>-(D)<sub>1</sub> is not trivial. Philosophers working on rational requirements have sometimes overlooked this fact. For example, John Broome writes:

Rationality requires of us that, if we intend an end and believe some means is necessary to that end, then we intend the means. This has to be a requirement of rationality. A person is necessarily irrational if she does not intend whatever she believes is a necessary means to an end she intends.<sup>63</sup>

Broome here seems to argue that since (B) is true, so must be (B)<sub>1</sub>. This argument, however, equivocates on ‘requirement’. In one sense of ‘requirement’, a sentence such as “rationality requires A to  $\phi$ ” states only that A’s  $\phi$ -ing is a necessary condition for A’s being rational. In this sense of ‘requirement’, the claims (A)<sub>1</sub>-(D)<sub>1</sub> are indeed equivalent

<sup>62</sup> I intend these formulas to be neutral between so-called *narrow-scope* and *wide-scope* interpretations of requirements, which are introduced in Chapter 3.1.

<sup>63</sup> Broome (2005b, 2).

to (A)-(D). But when Broome concludes that there must be an instrumental requirement of rationality, he means that there is a standard of rationality, the violation of which constitutes irrationality, and this is not quite the same as stating a necessary condition for being rational.

Here is an example: Having the capacity of rationality is a necessary condition of being rational in the standard-related sense. But it is not irrational not to have the capacity of rationality, and there is no standard of rationality that requires us to have that capacity, for rational standards only apply to beings that have the rational capacity in the first place. Similarly, the functioning of your cardiovascular system plausibly is a necessary condition for your being rational, but you are not irrational if this system breaks down, and there is no standard of rationality that requires its functioning.

It might be argued that while not every necessary condition of being rational entails a corresponding rational requirement, every necessary condition for not being irrational does. Claims (A)-(D) state sufficient conditions for irrationality, which means that not meeting these conditions is a necessary condition for not being irrational. So on the assumption that necessary conditions for not being irrational entail corresponding requirements of rationality, (A)<sub>1</sub>-(D)<sub>1</sub> could be derived from (A)-(D). The assumption still has implausible implications, however. Suppose that a necessary condition for your not being irrational is that it is not the case that you believe p and believe not-p. As a trivial consequence, it follows that a necessary condition for your not being irrational is that it is not the case that your name is Fred and you believe p and you believe not-p. But it is implausible to assume that there is a rational standard that requires of you that it is not the case that your name is Fred and you believe p and you believe not-p. Suppose your name is not Fred. Does it follow that because of this you satisfy some kind of rational standard? That is absurd.

The problem cannot be avoided by restricting in some way the relevant necessary conditions to those that mention only possible rational responses of you, such as your beliefs or intentions. If your not [believing p and believing not-p] is a necessary condition for your not being irrational, then so is your not [believing p and believing not-p and believing q]. Assume that q is the proposition that one plus one equals two, that this proposition is maximally independent of p, and that belief in it is entirely

rational for you. It is simply implausible to think that there is a rational requirement not to [believe p, believe not-p, and believe q] *over and above* the rational requirement not to [believe p and believe not-p]. If you believe that p, believe that not-p, and believe that one plus one equals two, and then give up your belief that one plus one equals two, it is absurd to think that you have thereby satisfied some kind of rational standard.<sup>64</sup>

So the inference from necessary conditions of rationality, or absence of irrationality, to rational requirements understood as rational standards, is invalid. The introduction of requirements of structural rationality is a substantial step; statements such as (A)<sub>1</sub>-(D)<sub>1</sub> do not simply follow from statements such as (A)-(D). It was Broome himself who (in later work) emphasised this point. He distinguishes a “property sense” of ‘requirement’, in which a requirement is just a necessary condition for one to have a certain property, from a “code sense”, and he adopts the code sense for ‘rational requirement’.<sup>65</sup> He illustrates this distinction helpfully by contrasting the requirements of survival with the requirements of the law: To say that survival requires you to breathe is to say nothing but that your breathing is a necessary condition for your survival. There is no code or system of rules here that you *violate* if you do not breathe or *satisfy* if you do. To say that the law requires you not to cross the street at a red light is different; it essentially involves the idea of a code or a system of rules that prohibits or demands certain actions – a standard that you can satisfy or violate. These requirements are not to be understood in terms of necessary conditions of your having the

---

<sup>64</sup> In deontic logic, this problem is known as “Ross’s paradox”. It occurs whenever an assumption about the logic of requirements validates the principle known as *disjunction introduction*, which licenses the inference from “you are required to  $\phi$ ” to “you are required to  $\phi$  or  $\psi$ ”. As A. Ross (1941, 61–62) first pointed out, this is very counterintuitive: That you are required to mail the letter should not be taken to entail that you are required to mail or burn the letter. It is sometimes argued that the paradox is harmless and should not be taken seriously (see e.g. Føllesdal and Hilpinen 1971, 21–23; Sinnott-Armstrong 1987, 136–37). But note that according to *disjunction introduction*, “you are required to mail the letter” entails not only “you are required to mail or burn the letter”, but also “if you burn the letter, you satisfy a requirement”. This reduces the idea of satisfying a requirement to absurdity. It would follow that every claim about a requirement or ‘ought’ entails a claim to the effect that a requirement or ‘ought’ has been satisfied for *every possible* response – including the *omission* of the response originally required. I conclude that everyone who thinks that statements or intuitions to the effect that some response satisfies a requirement have any content at all are committed to rejecting *disjunction introduction*. In any case, I do not accept it, and neither does Broome (2007b, 19–22).

<sup>65</sup> See Broome (2007c). He makes the same point in Broome (2007b), where he calls the “code sense” the “source sense”.

property of being law-abiding; on the contrary, whether you have this property is determined by the requirements of law.

As Broome notes, the sense of ‘rational requirement’ that is relevant for our discussion is the code sense, not the property sense. It is the idea of a standard or code that we can satisfy or violate, and which determines whether our actions are irrational or not. And once again, this means that the step from the existence of structural irrationality to structural requirements of rationality is not trivial.

Structural irrationality claims such as (A)-(D) nevertheless suggest that there are such structural requirements of rationality. For given that the structural conditions depicted by (A)-(D) really are sufficient for irrationality, and given that irrationality involves the violation of a rational requirement, we can conclude that there are rational requirements which are violated in (A)-(D). And at least *prima facie*, a natural explanation for this fact is that there are structural requirements of rationality such as (A)<sub>1</sub>-(D)<sub>1</sub>. So even though the existence of structural irrationality does not *entail* the existence of structural requirements of rationality, it clearly supports the assumption that there are such requirements.

### 1.5 A range of views

I am now in a position to state the basic problem that I shall be occupied with in this thesis. As I have said above, we intuitively think of rational requirements as normative. That is, we usually take it for granted that at least one of the following claims are true:

*The normativity of rationality, contributory:* If A is rationally required to  $\phi$ , then A has a reason to  $\phi$ .

*The normativity of rationality, strict:* If A is rationally required to  $\phi$ , then A has decisive reason (or ought) to  $\phi$ .

I shall explore the intuition behind these claims further in Chapter 2; my purpose here is only to emphasise their initial plausibility. We also, I think, are drawn by the phe-

nomenon of structural irrationality to accept that there are structural requirements of rationality. Thus, both of these claims seem plausible:

- (1) There are structural requirements of rationality.
- (2) Rational requirements are normative.

The problem is that (1) and (2) are incompatible with another claim that has been forcefully defended in the literature:

- (3) Structural requirements of rationality (if there are any) are not normative.

This claim is supported by two major arguments, which I shall discuss at length in Chapters 4 and 5. The first is that normative requirements of structural rationality would lead to implausible “bootstrapping” of reasons; they would allow us to create reasons for intentions and beliefs simply by adopting an attitude we have no reason to adopt. The second argument is that it is obscure what the reason to do what structural rationality requires of us would be.

Since (1), (2), and (3) cannot all be true, at least one of these claims has to be given up. Three views on the matter suggest themselves:

*The structuralist view:* There are structural requirements of rationality, and these requirements are normative.<sup>66</sup>

---

<sup>66</sup> This is the position that Broome once took for granted (Broome 1999), before he became agnostic about the normativity of rationality (Broome 2005a; Broome 2007a), which (notwithstanding his concession that he lacks an argument for it) he now embraces again (Broome 2013, 204). Other proponents of the structuralist view include Bratman (2009a), Hussain (unpublished), Korsgaard (2009a), Rippon (2011), and Southwood (2008) adopt versions of this view. Reisner (2011) holds a structuralist view with respect to theoretical rationality only, tending towards the mixed view regarding practical rationality. Proponents of what is sometimes called the “subjective reasons account” of rationality, such as Schroeder (2009) and Way (2009), might also count as proponents of the structuralist view, provided that they take subjective reasons to be genuinely normative reasons that matter in deliberation, and not merely apparent reasons (i.e. reasons one *would* have if one’s beliefs were true). If subjective reasons are supposed to be merely apparent reasons, or reasons in some sense that is irrelevant from the deliberative perspective, the subjective reasons account should rather be understood as a version of the mixed view. See Chapters 2.3, 4.3 and 5.6 for discussion of subjective reasons and ‘oughts’.

*The mixed view:* There are structural requirements of rationality, but these requirements are not normative.<sup>67</sup>

*The anti-structuralist view:* There are no structural requirements of rationality.

While we find proponents of both the structuralist and the mixed view in the current literature, the anti-structuralist view has not been advocated as far as I can see.<sup>68</sup> What has come to be called the *myth view* (after Joseph Raz's claim that the idea of an instrumental requirement of rationality is a myth, and Niko Kolodny's extension of this claim to other structural requirements of rationality) might seem to come close, but differs from it in two important respects.

First, the target of both main proponents of the myth view is not requirements of structural rationality in general, but only what Kolodny calls "requirements of formal coherence as such"<sup>69</sup> – requirements that correspond to the first-order cases of structural irrationality that I have distinguished from the second-order cases above. The basic idea is that rational requirements inherit their content from the structure of normativity – that there is no independent standard of formal coherence, but only the structure of reasons and the "shadow" it casts when we rationally respond to our *beliefs* about reasons (to use Kolodny's metaphor).<sup>70</sup> Hence, the myth view denies first-order structural requirements of rationality, but not necessarily second-order ones.

Second, especially Kolodny, but also Raz, have famously denied that rationality is normative, and this claim is now associated with the myth view.<sup>71</sup> This part of the

---

<sup>67</sup> See e.g. Brunero (2010a, 30), Dancy (2009), Kolodny (2005), Way (2010a, 220). The view is also implicit in Cullity (2008), and Scanlon (2007) – although unwilling to affirm it explicitly – seems to me to be committed to it as well.

<sup>68</sup> Although it is entailed by Price's view that "rationality is a capacity; it is not directed by requirements" (2008, 175). Price, I think, ignores the difference between the capacity sense and the standard-related sense of rationality.

<sup>69</sup> Kolodny (2007a, 229).

<sup>70</sup> Kolodny (2008a, 390).

<sup>71</sup> For example by Bratman (2009a, 419), who interprets the myth view as the claim that "it is a myth to think of coherence and consistency as themselves having a distinctive, noninstrumental normative significance" (cf. also Bratman 2009b, 239, n. 2). For the denial of the normativity of rationality, see esp. Kolodny (2005), but also Raz (2005a, 15): "The fallacy to avoid is the thought that irrationality in thought or action occurs only if one fails to conform to a reason [...]. It can consist in [...] ways of thinking and of forming beliefs or intentions, and so on, which do not conform to standards of rationality." In Raz (2005b,

promised enlightenment is relevant to the main challenge for the myth view, namely to explain appearances of (first-order) structural irrationality without recourse to (first-order) requirements of structural rationality. Since irrationality involves the violation of a rational standard, proponents of the myth view have to bring in other kinds of requirements: for Kolodny, these are at least in part the second-order requirements, while Raz understands rational requirements as standards of the well-functioning of our capacity of rationality.<sup>72</sup> And in both cases, Kolodny and Raz deny that these requirements are normative. The myth view is thus concerned with two myths: the idea that there are first-order requirements of structural rationality, *and* the idea that rational requirements are normative:

*The myth view:* There are no first-order requirements of structural rationality, and rational requirements are not necessarily normative.<sup>73</sup>

Note that at least as Kolodny develops the myth view, it *entails* the mixed view, which accepts the existence of certain structural requirements, while maintaining that they are not normative.<sup>74</sup>

## 1.6 Prospects

I believe that none of the views that are actually advocated in the literature are satisfactory. Both the mixed view and the myth view deny that rationality is normative. In Chapter 2, I argue that this denial deprives proponents of these views of a basis for making sense of the common understanding that the attribution of irrationality is a

---

9), he explicitly states that the claim “irrationality need not be acting against reason” is one of the central claims of his article “The Myth of Instrumental Rationality”.

<sup>72</sup> See e.g. Kolodny (2007a, 242–244) and Raz (2005a, 14–18). Raz’s well-functioning account is discussed in Chapter 2.

<sup>73</sup> See Kolodny (2005; 2007a; 2008a; 2008b; unpublished), Raz (2005a; 2005b; 2010a, esp. §§4-5). I believe that Scanlon (2007) is also best interpreted as holding exactly this view.

<sup>74</sup> Things are a bit more complicated, for Kolodny’s “transparency account” (2005, §5) aims to reduce structural requirements to descriptive statements of psychological fact; hence the question turns on whether this reduction is supposed to be eliminative. But note that Kolodny’s main argument against the normativity of rationality depends essentially on there being such structural requirements. I thus take the reduction to be intended as non-eliminative.

form of criticism, and thus commits them to a highly revisionary view of ordinary ascriptions of rationality and irrationality. I shall also argue that the assumption of non-normative requirements of rationality is obscure. The very idea of a mixed view, whose proponents eagerly discuss the content and logical structure of rational requirements while at the same time accepting that these requirements have no normative force, seems to me misguided.

Yet the structuralist view comes with problems, too. Chapter 3 introduces and explores the idea of structural requirements of rationality. Chapters 4 and 5 discuss various arguments to the effect that structural requirements (if there are any) are not normative. To anticipate, my conclusion will be that these arguments are forceful and have not yet been successfully dismissed. So there is pressure on both the structuralist and the mixed view. This result is further substantiated in Chapter 6, which argues that the assumption of structural requirements of rationality itself poses several severe problems quite independently of whether they are understood normatively or not.

In Chapter 7, I begin to develop my own view on the matter, which embraces the normativity of rationality (2), and accepts the arguments that structural requirements of rationality are not normative (3), at the cost of denying that there are any structural requirements of rationality (1). The last four chapters of the thesis take up the most important challenges that this view faces. The first is to develop a positive account of non-structural requirements of rationality that is compatible with the normativity of rationality. The second is to explain the phenomenon of structural rationality in terms of these non-structural requirements. The first challenge is taken up in Chapters 7 and 8. Chapter 7 defends the view that rationality requires us to respond correctly to the reasons available to us (where a reason is available if it is part of our evidence). Chapter 8 argues that what we have decisive reason, or ought, to do is determined solely by the reasons available to us. The hypothesis that these chapters explore is that once we see that rationality can take us beyond subjective coherence and that reasons are sensitive to available evidence, we will also see that (non-structural) requirements of rationality and requirements of reasons coincide. Chapters 9 and 10 finally take up the challenge of explaining the phenomenon of structural irrationality in terms of these non-structural requirements of rationality.

On the one hand, the theory I will be suggesting is even more radical than the myth view; it claims, one could say, that the whole idea of structural standards of rationality is a myth. And it shares with the myth view the challenge of explaining the phenomenon of structural irrationality without recourse to such structural requirements. On the other hand, however, I believe that my view is much less revisionary than the myth view, for the rational requirements that I will invoke in order to explain the appearance of structural rationality are, as I argue, necessarily also normative requirements. I therefore think that my view is better equipped to meet what I take to be an important desideratum of a good theory of reasons and rationality, namely to account for the appearance of structural irrationality without sacrificing the normativity of rationality.

## Rationality, reasons, and criticism

---

The aim of this chapter is to elucidate why the normative question about rationality is important and why the normativity of rationality is plausible. Starting with a comparison to the normative question about morality, I suggest an argument that supports a strong version of the normativity of rationality (2.1). I go on to explain and support the crucial premise of this argument, which states that a person is criticisable only if she violates a decisive reason (2.2). I defend this premise against three possible alternative conceptions of criticism, which hold that criticisability can be grounded in the violation of a subjective ‘ought’ (2.3), in malfunctioning (2.4) or in acting in a less-than-virtuous way (2.5). Next, I discuss the other premise of the argument, which states that irrationality is criticisable. My main conclusion is that the sceptic about the normativity of rationality is committed to a highly revisionary error theory about ordinary attributions of rationality and irrationality (2.6). Finally, I present a further argument for the normativity of rationality, which claims that it is simply obscure what non-normative rational requirements are (2.7).

### 2.1 The criticism argument

In the last chapter, I argued that the sense of ‘rationality’ that contrasts with ‘irrationality’ is related to standards of correctness or requirements. Since not all such requirements are authoritative or normatively binding for a person, however, the assumption of rational requirements leaves conceptual space for asking the normative question – a question that I have been understanding as asking for normative reasons. Why is this question important?

Consider the analogous question about morality. Morality certainly involves standards, according to which we can assess agents and their responses as moral or immoral. On the reasons claim, these norms are authoritative for us only if we have

reasons to conform to them. We can thus ask the normative question about morality: Should we really do what morality requires of us? What reasons do we have to act morally? In other words: *Why be moral?*

The challenge posed by this question is a serious one. If it cannot be met, I believe we are forced to conclude that a huge part of the moral discourse actually rests on an illusion. It is a central feature of this discourse that to call some person's action morally wrong is to criticise that person. Critique, however, involves substantial normative judgements; it cannot be captured by the idea that a formal standard has been violated alone. It thus seems that insofar as we engage in moral discourse at all, we are presupposing the normativity of morality.

The following example may help to illustrate this point. According to newspaper reports, the well-known US Republican politician and former priest Mike Huckabee *criticised* the actress Natalie Portman for "glorifying" out of wedlock pregnancy in a discussion of Portman's appearance at the Academy Awards.<sup>1</sup> What does understanding his statement as a form of criticism involve? If Huckabee simply alluded to the fact that there is some standard which bans premarital pregnancy, for example the standard of his Baptist church, everyone could have happily agreed with him without thinking that the standard is in any way binding for Portman, and there would have been no reason to understand Huckabee as expressing any kind of criticism. Insofar as we understand Huckabee as *criticising* Portman, we must take him to claim that the standard in question is indeed authoritative for Portman; and by the reasons claim this entails that Portman must have reasons, decisive reasons indeed, to conform to the standard. This in turn explains why many people thought of his criticism as inappropriate; they disagreed with Huckabee that Portman really is bound by a norm forbidding premarital pregnancy.

So if calling an action morally wrong is a form of criticism, and if criticism involves the judgement that the criticised person had decisive reason not to act as she did, it follows that insofar as we engage in moral discourse at all, we are supposing that the norms of morality are binding, that if they apply to a person, this person must also

---

<sup>1</sup> See e.g. *Süddeutsche Zeitung*, 5-6 March 2011, p. 12.

have decisive reasons to conform to them. And that in turn means that in order to make sense of moral discourse, we need to account for the normativity of morality.

Similar considerations, I believe, apply to rationality. Calling a person irrational involves criticising that person. The criticism is different in character from moral critique, but it is still criticism. Criticising someone involves more than the judgement that the criticised person has violated some standard, it also involves the judgement that the standard is authoritative for her. And by the reasons claim, this means that the person has decisive reasons to conform to this standard. Hence, ascriptions of rationality and irrationality seem to presuppose that we have decisive reasons to be rational. And if we cannot make sense of this, it seems we must give up on an essential feature of the discourse about rationality ascriptions and admit that it rests on an illusion.

The argument behind this reasoning may be summarised as follows:

- (1) If A is rationally required to  $\phi$ , then A would be criticisable for not  $\phi$ -ing.
- (2) If A would be criticisable for not  $\phi$ -ing, then A has decisive reason to  $\phi$ .
- (3) Therefore, if A is rationally required to  $\phi$ , then A has decisive reason to  $\phi$ .

I shall call this *the criticism argument*. It is a simple, but, as I believe, forceful argument. It suggests that the normativity of rationality that we are intuitively presupposing is *strict* rather than *contributory* (to come back to the distinction I made at the ending of the last chapter). People are not criticisable for not conforming to just any *pro tanto* reason; the reasons must be on balance decisive. And this means that insofar as irrationality ascriptions involve criticism, they presuppose strict normativity of rationality.<sup>2</sup>

The criticism argument shows that the sceptic about the normativity of rationality has to give up one of two initially plausible claims about criticisability: He has to deny either that irrationality is criticisable, or that criticisability presupposes the

---

<sup>2</sup> This is often overlooked by both sceptics and proponents of the normativity of rationality. For example, Broome (2005a; 2007a; 2013, ch. 11) and Rippon (2011) are merely looking for a defence of contributory normativity, even though both maintain that irrationality ascriptions involve criticism.

violation of a decisive reason.<sup>3</sup> The forthcoming sections discuss these possibilities in reverse.

## 2.2 The criticisability claim

The second premise of the criticism argument states that one would be criticisable for not giving a response only if one were to have decisive reason for it. By the same token, one would be criticisable for a response only if one were to have decisive reason against it. I shall call this the *criticisability claim*: An agent is criticisable for a response (or lack of response) only if she had decisive reason for (or against) it.

### *Preliminaries*

Let me begin by clearing away a number of possible misunderstandings of this thesis. First, the words ‘critique’ and ‘criticism’ can be used to express several different concepts, and not all of these concepts are connected to reasons in the way the criticisability claim suggests. In one sense, ‘criticism’ means something like ‘examination’ or ‘assessment’, which need not even have a negative connotation (compare ‘film criticism’ or ‘critique of the power of judgement’). In another sense, a criticism is just about any judgement of negative evaluation. In this sense, we can criticise a knife as blunt. In contrast, the sense of ‘criticism’ at issue in the present discussion is essentially directed at agents and concerns particular responses of them. It is the sense in which we criticise *someone* for *something* that she has done (in the wider sense of ‘doing’ that encompasses responses such as intentions and beliefs, as well as lacks of responses).<sup>4</sup> I shall call this *agent criticism*.

---

<sup>3</sup> For convenience, I sometimes use the expression “violation of a decisive reason” as short for “not conforming to what one has (on balance) decisive reason to do”. I do not mean to imply that there must be one particular reason that is the decisive one, but only that the preponderance of reasons decisively counts in favour of (or against) a response.

<sup>4</sup> Cf. Williams (1993, 72): “Blame needs an occasion – an action – and a target – the person who did the action and who goes on from the action to meet the blame. That is its nature; as one might say, its conceptual form.” The same can be said about agent criticism more generally (except that I think that responses other than actions can be criticisable).

It is important to recognise that not every kind of negative assessment of actions amounts to agent criticism. Consider an example that Gideon Rosen puts forward to make a similar point:

Your daughter is practicing the piano, but she is getting frustrated. So you decide to cheer her up with a laughable performance of her piece. You start in Mozart's C-major sonata, but you play a C# where a B should be as the first note of the second measure.<sup>5</sup>

Suppose that I judge your performance to be faulty or incorrect. Then in one sense I could be said to criticise your performance. But that would not necessarily mean that I criticise *you* for performing as you are performing, or that I think that you are *criticisable* for what you are doing. A negative assessment of an action relative to some standard of evaluation does not necessarily involve agent criticism.

In a very broad sense of 'blame', agent criticism might be identified with blame. But we often use 'blame' in a narrower sense, in which it has particularly moral connotations, is connected with feelings of resentment and indignation, and indicates that the relationship between the blaming and the blamed party is impaired.<sup>6</sup> These are not necessary features of agent criticism in general. For example, I may criticise a person epistemically for having an unjustified belief, but this need not have a particularly moral connotation, and it need not go along with feelings of resentment or an impairment of our relationship. So I prefer to speak of criticisability here, taking blameworthiness to be a particular kind of criticisability.<sup>7</sup>

---

<sup>5</sup> Rosen (2001, 620).

<sup>6</sup> Strawson (1962) and – following him – Wallace (1994) emphasise the connection between blame and reactive attitudes such as resentment and indignation; Scanlon's account of blame (2008, ch. 4) stresses the relevance of personal relationships.

<sup>7</sup> It follows that the criticisability claim entails an analogous claim about blameworthiness: a person is blameworthy for not  $\phi$ -ing only if she has decisive reason to  $\phi$ . This is implicit in some conceptions of blameworthiness. For example, according to Wallace, "it is a condition for moral blameworthiness of an action [...] that the action violates some moral obligation that we accept" (1994, 134), where "the stance of holding people to moral obligations one accepts in this way incorporates a commitment to [...] identifying reasons [...] to comply with the obligations in question" (1994, 130). Similarly, according to Sher, "the appropriateness of blame depends [on the truth of] the belief that the agent has acted badly in the sense of flouting or ignoring a compelling moral reason" (2006, 118), where moral reasons "purport to override all

The second point I would like to emphasise is that I am here primarily concerned with the *correctness conditions of criticism*, not with the *ethics of criticism*. By saying that an agent is criticisable, I mean that it would be correct or appropriate to criticise this person, and this should be distinguished from the claim that we ought, or have sufficient reason, to do this. It is possible that there are good independent reasons against criticising a person who is in fact criticisable (at least when ‘criticism’ refers to an action rather than an attitude)<sup>8</sup>. For example, the fact that a criticism would upset a person and cause her to make matters worse might provide a decisive reason against confronting her with a criticism even if she is in fact criticisable. We might even have decisive reason to criticise someone who is not criticisable at all, for example when this is our only way to save some person’s life or achieve some other important value. There are different views about the ethics of criticism; what is important in the present context is that the criticisability claim does not preclude any of them.<sup>9</sup>

Third, the claim that criticisability implies the violation of decisive reasons or ‘oughts’ should not be confused with, and in particular does not entail, the claim that one can be criticisable only for voluntary actions or whatever one can intentionally control. Quite the contrary, throughout this thesis I assume that ‘oughts’ and reasons

---

others” (2006, 15). It is less clear how the accounts of Graham (forthcoming) and Scanlon (2008) relate to the criticisability claim, for both deny that the wrongness of an action is a condition for blameworthiness. However, according to Graham we are (strictly speaking) not blameworthy for actions but only for attitudes, and Graham accepts that we are blameworthy for an attitude only if we violate a moral requirement in having it. So given that moral requirements are accompanied with decisive reasons, Graham’s account entails the criticisability claim. For Scanlon’s position, see note 12 below.

<sup>8</sup> Like ‘blame’, ‘criticism’ can refer to both an attitude and a speech act. The correctness conditions are the same in both cases, but since reasons for attitudes differ from reasons for action, the ethics of act-criticism differs from the ethics of attitude-criticism. In particular, it is arguable that the only reasons for and against attitude-criticism bear on the criticisability of the agent. This is because it is plausible to think that reasons for attitudes must be object-given, and that reasons for attitude-criticism that do not bear on the criticisability of the agent would not be object-given but state-given reasons (see Chapter 1.2 for this distinction). However, according to the evidence-relative view about reasons that I shall defend in Chapter 8, reasons for attitude-criticism and criticisability could still come apart. One might have no reason to criticise a criticisable person because one has no evidence for the criticisability. And one’s evidence might provide one with decisive reasons to criticise a person who is not in fact criticisable.

<sup>9</sup> Most notably (perhaps), the criticisability claim is compatible with the consequentialist claim that “each instance of praising and blaming is itself an act that should be assessed in terms of its consequences” (De Lazari-Radek and Singer 2010, 57). These claims are compatible because the latter is about the *ethics* of criticism and not about the correctness conditions of criticism. It is of course arguable that since incorrect criticism is undeserved criticism, incorrect criticism is necessarily morally problematic. But again, this is a substantial thesis about the ethics of criticism; it does not follow from claims about the correctness conditions of criticism alone, or at least requires a substantial argument.

also apply to attitudes that we cannot intentionally control, such as beliefs. In my view, this is compatible with the principle that ‘ought’ implies ‘can’ when this principle is correctly interpreted. If you ought to  $\phi$ , it follows that you have reasons to  $\phi$ , and that it is possible for you to  $\phi$  for these reasons. But that it is possible for you to  $\phi$  for a reason does not entail that it is possible for you to  $\phi$  intentionally. Reasons for belief are a case in point. Hence, arguments to the effect that we can be criticisable for what lies out of our intentional control do not by themselves provide any reasons to question the criticisability claim.<sup>10</sup>

Fourth, it is worth noting that the criticisability claim does not rule out the possibility of *blameless wrongdoing*, i.e. it does not entail that every violation of an ‘ought’ or decisive reason is criticisable.<sup>11</sup> Fifth, the criticisability claim also does not rule out the possibility of what might be *blameworthy ‘rightdoing’* in the following sense: An agent could be criticisable for the way in which she  $\phi$ -s, even though she has decisive reason to  $\phi$  (and thus no decisive reason not to  $\phi$ ). For example, a person can have decisive reason to apologise for something, and still be criticisable for the way in which she apologises; say, if she apologises for the wrong reasons. The criticisability claim only insists that in such a case, the person must have decisive reasons not to  $\phi$  *in this criticisable way*. This poses no problem. Just as it is possible that a person is not criticisable for  $\phi$ -ing, but criticisable for  $\phi$ -ing in a certain way, it is possible that a person has decisive reason to  $\phi$ , but decisive reason not to  $\phi$  in a certain way.

*First argument: sufficient reasons defeat criticism*

This concludes my preliminary remarks about the criticisability claim. I shall now turn to the question of whether it is true. Agent criticism involves at least some kind of negative evaluation, so there can really be no question whether criticism involves nor-

---

<sup>10</sup> Adams (1985), one of the most prominent proponents of the view that one can be blameworthy for involuntary states, does not seem to object to the criticisability claim, and does apply ‘ought’ to such states himself (see. e.g. 1985, 5).

<sup>11</sup> I say this even though I am sceptical that a person could be immune to criticism despite the fact that she violates a decisive reason. It seems to me plausible that the very conditions that undermine criticisability undermine the relevant ‘ought’-claim as well. For an argument to this conclusion with respect to moral requirements, see Rivera-López (2006).

mative judgements in the broad sense of ‘normative’ which contrasts with ‘merely descriptive’. The question thus turns on what this judgement is, and the criticisability claim states that it is the judgement that the criticised person violated a decisive reason. My first argument for this claim is that one cannot coherently criticise a person for a response if one accepts that the person had sufficient reasons for it. Suppose that I criticise you for something, say, for not showing up at my party even though you promised you would. You then tell me that a close friend of yours asked you for help that night, and that you took this to be a sufficient reason not to come to the party. I reply as follows: “You are right. That your friend needed help gave you sufficient reason not to show up – but I maintain my criticism nonetheless.” This would be a *very* strange thing to say. Maintaining a criticism while accepting that the person had sufficient reasons for what she is criticised for seems incoherent.

It follows that the kind of normative judgement involved in criticism should entail that the person did not have sufficient reasons for giving the criticised response. It cannot be the judgement that the person did not have decisive reason for the response itself, for the absence of a decisive reason is no grounds for criticism. Neither can it be the judgement that the person did not have sufficient reason for the response, for the absence of a sufficient reason only provides grounds for criticism when the absence of a sufficient reason *for* a response is also a decisive reason *against* it (as is plausible in the case of belief, but controversial in the case of action). The only judgement that both provides grounds for criticism and has the implication that the agent did not have sufficient reason for a response is, I think, the judgement that this person had decisive reasons *not* to give that response. Hence, this is the judgement on which agent criticism must ultimately rest.

Some authors would object that it *can* be coherent to criticise a person for a response for which she had sufficient reason, namely if she does not give the response for the sufficient reason, but for other reasons, which are insufficient.<sup>12</sup> Suppose you did

---

<sup>12</sup> Scanlon maintains that “it can be appropriate to blame a person who has done what was in fact the right thing if he or she did it for an extremely bad reason” (2008, 125). Supposing that (i) the action in question might also be supported by sufficient reasons, and that (ii) the person in question would nevertheless be blameworthy *for the action* (rather than for acting for the bad reason, for instance), this claim is incompatible with the thesis that blameworthiness for  $\phi$ -ing entails decisive reasons against  $\phi$ -ing. Assumption (i) seems

not come to my party not because you helped your friend (which you had sufficient reason to do), but because you watched TV all night and were just too lazy to get up from your armchair. In this case, it may seem that I could coherently criticise you for not coming to my party while accepting that you did in fact have sufficient reasons not to come (and thus lack decisive reason to come).<sup>13</sup>

I am not convinced of this objection. To begin with, there are a number of things that you do in the case described even though you do not have sufficient reason to do them, and indeed have decisive reason not to do them. For example, you watch TV instead of going to my party, you break a promise for insufficient reasons, and you (arguably) lack appropriate concern for me. So the example does not show that you can be criticisable without violating a decisive reason; if anything, it shows that you can be criticisable *for  $\phi$ -ing* without having decisive reason *not to  $\phi$* .<sup>14</sup> Now the criticism argument stated above does indeed require that the criticism and the corresponding reason have the same object. But a weaker version of this argument would still yield the result that the violation of a rational requirement involves the violation of a decisive reason, and this would still put a lot of pressure on the sceptic about the normativity of rationality. So I doubt that the objection is of any significant help for the sceptic.

Second, and more importantly, the example conflicts with the criticisability claim only on the assumption that in such a case you are criticisable *for not coming to*

---

inevitable given Scanlon's view that we always have sufficient reasons to be moral (cf. Scanlon 1998, 148). Assumption (ii) is strongly suggested by Scanlon's general claim that "to claim that a person is blameworthy *for an action* is to claim that the action shows something about the agent's attitudes toward others that impairs the relations that others can have with him or her" (2008, 128, my emphasis). Hence, assuming that blameworthiness entails criticisability (which is not entirely evident on Scanlon's account of blame), Scanlon's view is incompatible with the criticisability claim. The argument to follow suggests that Scanlon's account should be revised such that, firstly, it guarantees that the blameworthy person has decisive reason against holding the relevant attitudes toward others, and secondly, that it no longer entails (ii). If revised in this way, it is compatible with the criticisability claim.

<sup>13</sup> Capes (2012, §5) presents a structurally equivalent example in order to show that "a person can be blameworthy for A-ing even if it was not objectively morally wrong for her to A" (2012, 434), where by "objectively morally wrong" he means that the action is "morally required all things considered" rather than believed (by the agent) to be morally required all things considered (2012, 418 and 434, n.1). Everything I go on to say applies to his example as well.

<sup>14</sup> Capes seems to agree. On his account, "if S is blameworthy for A-ing, then A manifested a morally objectionable quality of will" or "morally unjustifiable degree of ill will" (2012, 432). Assuming that we have decisive reasons against attitudes which constitute a morally unjustifiable degree of ill will (which seems plausible), Capes' account entails that we are blameworthy only if we violate a decisive reason (even though on his account, the blame and the reason might have different objects).

*the party*, and this assumption does not stand up to reflection. There are a number of alternative descriptions that the case itself suggests as more plausible candidates for the object of appropriate criticism. For example, we might say that you are criticisable for watching TV instead of coming to my party, or for breaking a promise for insufficient reasons, or for your lack of appropriate concern for me. And as mentioned above, of all of these responses we could plausibly say that we have decisive reason against them.

This point can be substantiated by the following principle:

*Joint avoidability*: If A is criticisable for  $\phi$ -ing, and A is criticisable for  $\psi$ -ing, then it was possible for A to [not  $\phi$  and not  $\psi$ ].

This seems surely plausible: if it was not possible for you to omit both of two responses, then it cannot be true that you are criticisable for each of them. If this is true, however, we must reject the assumption that you are criticisable for not coming to the party. For recall your friend who needed your help at the evening of my party. Plausibly, you are criticisable for not helping your friend. Yet we assumed that you could not both have helped your friend and have come to my party; this was exactly why we accepted that you did have sufficient reason for not showing up. So it follows from *joint avoidability* that you are not criticisable for not coming to my party; the object of the criticism (insofar as it is appropriate) must be something else. I conclude that the example provides no reason to think that one could coherently criticise someone for a response while accepting that he had sufficient reason for the response.

The argument was based on the assumption that you are criticisable for not helping your friend. Why could my opponent not reject this assumption instead? Because rejecting this assumption while insisting that you are criticisable for not coming to my party would be entirely *ad hoc*. In contrast, I can reject the assumption that you are criticisable for not coming to the party while insisting that you are criticisable for not helping your friend on grounds of the assumption that you had decisive reason to help your friend, but no decisive reason to come to the party – this is not *ad hoc* at all. The moral is that if one denies the criticisability claim, one can get around violating

*joint avoidability* only at the expense of an *ad hoc* exclusion of certain responses as criticisable.

*Second argument: no other normative judgement grounds criticism*

I shall next turn to my second argument for the criticisability claim, which will be the focus of the next three sections. The argument is that there is simply no plausible alternative conception of the normative judgement involved in agent criticism. Someone who denies the criticisability claim has to tell us which broadly normative claim other than one about a decisive reason grounds criticism, and I think there is no good answer to this question. The options seem to be this: Either we understand criticism as being grounded in the violation of an ‘ought’ or normative requirement that is different from a claim about decisive reasons, or we understand criticism as being grounded in the violation of an evaluative standard rather than an ‘ought’. I shall discuss these options in turn.

### 2.3 Criticisability and the subjective ‘ought’

I am out to defend the claim that no (broadly) normative judgement other than the judgement that an agent has violated a decisive reason could figure as ground for appropriate criticism. This section considers a possible alternative, according to which criticism involves a judgement to the effect that the agent has violated an ‘ought’ other than the ‘ought’ of decisive reasons. This amounts to a denial of what I have called *the reasons claim*, according to which we are normatively required to  $\phi$  only if we have decisive reason to  $\phi$ .<sup>15</sup> Derek Parfit and, with some variations, Mark Schroeder seem to hold this view. Both authors argue that in criticism, we generally need to appeal to a notion of ‘ought’ that does not correspond to what we have decisive reason to do, but rather to what we *would* have decisive reason to do *if* things were as we believe them to be. Following Parfit, I shall call a reason that a person would have if things were as she believed them to be an *apparent reason*. Following Schroeder, I shall call the ‘ought’ that

---

<sup>15</sup> See Chapter 1.1.

is given by an agent's decisive apparent reasons the *subjective 'ought'*.<sup>16</sup> According to the view at issue, criticisability is grounded in the violation of a subjective 'ought', which is given by an agent's decisive *apparent* rather than actual reasons.

There are good reasons to be sceptical that there is a genuinely normative notion of 'ought' that corresponds to our apparent reasons, some of which will be discussed in later chapters.<sup>17</sup> I agree with those who find it problematic to suppose that normativity is fundamentally divided between different notions of 'ought' that can conflict but not weighed against each other.<sup>18</sup> It is not easy to make sense of the proposition that one and the same agent is subject to different 'oughts' that at least sometimes issue incompatible directives. In this chapter, however, I shall not question the existence of the subjective 'ought'. Rather, I shall argue that regardless of whether truths about decisive apparent reasons ground truths about what we subjectively ought to do, they do not ground truths about what we are criticisable for.

I shall argue this point by way of an example. Suppose you are planning a weekend camping trip, and out of wishful thinking, you believe that the weather will be splendid. You also believe that this gives you decisive reasons to leave your rain gear at home, because you would then be carrying unnecessary extra weight. If your beliefs were true, you would have decisive reason to leave your rain gear at home; hence, this is what you have decisive apparent reason (or ought subjectively) to do. So if apparent reasons ground criticisability, then bringing your rain gear would be criticisable. But suppose that you do not indeed have any reason to believe the weather will be splendid. In fact, several people have warned you that it is very likely going to rain. In this case,

---

<sup>16</sup> See Parfit (2011a, 162–163) and Schroeder (2009). In contrast to Parfit, Schroeder seems to think that apparent reasons are genuinely normative reasons; not only "ought" but "reason" as well is ambiguous between different readings, one of which (the subjective one) is given by what would be the case about the other if one's beliefs were true. Consequently, Schroeder's way of putting things is literally speaking compatible with both the reasons claim as well as the criticisability claim. However, the notions of 'ought' and 'reason' that I refer to when I use them without qualification are the ones that are primary in deliberation (see Chapter 1.2). These notions, I assume, are not Schroeder's subjective notions: otherwise the inference from 'I believe that I have decisive reason to  $\phi$ ' to 'I have decisive reason to  $\phi$ ' would have to qualify as correct deliberation, which it is not. So in the deliberative sense in which I am using the notion of a 'reason', Schroeder's "subjective reasons" are not (necessarily) actual reasons, the assumption of a subjective 'ought' is not compatible with the reasons claim, and the view that criticisability entails the violation of a subjective 'ought' is an alternative to the criticisability claim.

<sup>17</sup> See esp. Chapter 4.3 and 5.6.

<sup>18</sup> See e.g. Dancy (2000a, 61); Kolodny (2005, 512); Zimmerman (2008, 6–8).

you seem to be criticisable for *not* bringing your rain gear. You cannot be criticisable for both bringing and not-bringing your rain gear. That you have decisive apparent reason against a response is thus neither necessary nor sufficient for the criticisability of this response.

Note that this argument does not imply that you might act against your decisive apparent reasons and still not be criticisable at all. Perhaps you are criticisable for your beliefs, or for the combination of your beliefs and your omission to respond in the way your apparent reasons suggest. If this is so, and in what way, is just part of the very question at issue in this thesis. The view presently under discussion claims that the intuition that you are criticisable in such a situation is grounded in the fact that you ought subjectively to give the response that is favoured by your decisive apparent reasons, and that one is generally criticisable for refraining from giving responses that one ought subjectively to give. The example shows that this explanation fails: You might be criticisable for not bringing your rain gear even if you ought subjectively not to bring it, and you might not be criticisable for bringing your rain gear even if you ought subjectively leave it at home.

Parfit puts forward a different example that is designed to show that we need to appeal to apparent reasons in criticism. A person who believes in sorcery and tries to kill someone by sticking pins into a wax dummy seems blameworthy, no matter whether she has any actual reasons that count against sticking pins into a wax dummy.<sup>19</sup> It thus seems that the ‘ought’ that grounds the criticism in question cannot be the ‘ought’ of decisive reason; according to Parfit, it is a subjective ‘ought’, which corresponds to what this person’s reasons *would* be if her beliefs were true rather than her actual reasons.

The example does not actually support this conclusion, however. What is criticisable is not that this person sticks pins into a wax dummy, but that she tries to kill someone. And trying to kill someone is plausibly something that she has *actual* reasons not to do. If she does not have such reasons, it is difficult to see how she could still be blameworthy. In contrast, the assumption that she ought subjectively to refrain from

---

<sup>19</sup> See Parfit (2011a, 152).

sticking pins into a wax dummy is irrelevant for the criticism, for whether she sticks pins into a wax dummy is beside the point. The example thus supports the above conclusion that having decisive apparent reason against a response is neither necessary nor sufficient for the criticisability of this response. The normative judgement that grounds criticism therefore cannot be the judgement that the agent has violated a subjective 'ought'.

## 2.4 Criticisability and well-functioning

I shall next turn to the proposal that criticisability is grounded not in the violation of an 'ought' at all, but in the violation of an evaluative standard. This option has some initial plausibility, since evaluative standards can indeed ground evaluative attitudes such as approval and disapproval, which seem to be at least in the same ballpark as criticism. We need to be cautious, however. There is a loose sense of 'praise', by which we simply mean approval of an agent's properties such as their beauty. This praise needs to be distinguished from what we might call *agent praise*, which can be appropriate only if it is directed at something that an agent is accountable for. Similarly, not every kind of disapproval of an agent's properties amounts to agent criticism. We can disapprove of a person's smell, but we can criticise this person only for an action or some other agential response, such as their refusal to take showers.

The claim that criticism can be grounded in the violation of evaluative standards can be interpreted in two different ways. One idea is that such evaluative standards are provided by facts about the well-functioning of human beings or their rational capacities. The other idea is that they are set by virtuous character traits such as courage, generosity, or temperance. These ideas are also sometimes combined, for according to some authors, to function well as a human being is to be virtuous.<sup>20</sup> I shall discuss them separately nevertheless, starting with standards of proper functioning in this section and turning to standards of virtue in the next.

---

<sup>20</sup> This idea seems to be at the core of the theories proposed by Foot (2001), Thompson (2008) and Thomson (2008).

A natural place to look for a conception of criticisability grounded in standards of proper functioning is Joseph Raz's *well-functioning approach* of rationality. Raz agrees that irrationality ascriptions involve criticism; he states for example: "People who fail to pursue the means to their ends display or manifest a form of malfunctioning criticisable as a form of irrationality."<sup>21</sup> But Raz also thinks it is a "fallacy" to believe that "irrationality [...] occurs only if one fails to conform to a reason."<sup>22</sup> So Raz seems to embrace premise (1) of the criticism argument – the criticisability of irrationality – but deny its conclusion (3) – the normativity of rationality. He must be denying premise (2), then, maintaining that one can be criticisable even if one conforms to all of one's decisive reasons.

What makes an agent criticisable when he behaves irrationally, if in doing so the agent may not violate any reason? According to Raz, the agent violates a "standard of deliberative rationality".<sup>23</sup> So much is common ground: Irrationality involves the violation of some kind of rational standard. This, however, cannot by itself ground legitimate criticism if we follow Raz in thinking that these standards are not to be understood as normative or reason-implying. Think of the standards of etiquette: Sometimes we have reasons to conform to them, and sometimes we do not. Often, for example when we are visiting a foreign country, violating the standards of etiquette would be disrespectful, which gives us reasons to conform to them. In such cases, we can also be criticisable if we violate them. In other cases, we may have no reasons to conform to etiquette standards: You (a woman) may on occasion have no reason to refrain from asking a man to dance tango with you, even though tango etiquette requires women to wait until they are asked to dance. But in such a case, it is also clear

---

<sup>21</sup> Raz (2005a, 18), see also Raz (2005b, 10). Simon Gaus suggested to me the possibility of interpreting Raz in such a way that the sense of criticism involved in irrationality attributions is not the sense involved in what I have called *agent criticism*, but rather the sense in which we can criticise a knife as blunt or a piano performance that was intentionally played incorrectly as faulty. Now, I find it unnatural in such cases to say that the knife or the performance was *criticisable*, as Raz says that irrationality is criticisable; I thus take him to maintain that irrationality attributions do involve genuine agent criticism. Be that as it may, it seems to me worth considering whether agent criticism can be grounded in malfunctioning, regardless of whether Raz is committed to thinking this or not. And most importantly, it is evidently plausible that ordinary irrationality attributions involve agent criticism; the charge of irrationality is a form of accusation that differs significantly from judgements such as the one that an intentionally incorrect piano performance was faulty. See also Section 2.6.

<sup>22</sup> Raz (2005a, 15).

<sup>23</sup> Raz (2005a, 15).

that the violation of the standard in question does not provide legitimate ground for criticism. Hence, as long as we do not understand rational standards as reason-implicating, the appeal to such standards cannot answer our question of what makes irrationality criticisable.

Raz, however, has a particular view of the standards of rationality, which sets his conception apart from that of John Broome, Niko Kolodny, and others who prefer to call these standards “requirements”. As we have seen, Broome construes rational requirements along the model of the law: as a code or a system of rules. It seems clear that the violation of such rules can be criticisable only if we assume that we have reason or ought to follow these rules in the first place. Raz, in contrast, invokes mainly evaluative vocabulary to explicate what he has in mind; he thinks that the standards in question are related to an “ideal of rational agency”<sup>24</sup>; they are “standards that determine well-functioning deliberative processes”<sup>25</sup>. As he summarises his view in an article that responds to criticisms:

Sometimes people [...] display or manifest a form of malfunctioning criticizable as a form of irrationality. Broadly speaking, that much is common ground. My argument was that what makes such failures irrational is not that the agent’s irrational beliefs or actions are undesirable, nor that he has any reason to avoid them, or their combination. It is that he is not functioning properly.<sup>26</sup>

According to Raz, then, rationality and irrationality ascriptions amount to, or are at least grounded on, judgements about the proper functioning of an agent or his rational capacities.<sup>27</sup> The standards in question are standards of well-functioning, perhaps comparable to the standards that we are applying when we call a knife a “good knife”,

---

<sup>24</sup> Raz (2005a, 18).

<sup>25</sup> Raz (2005a, 17).

<sup>26</sup> Raz (2005b, 10).

<sup>27</sup> I take Raz’s somewhat obscure talk of the malfunctioning of an *agent* (as it occurs in the quote given above) to be short for talk of the malfunctioning of this agent’s *capacities* (as it occurs elsewhere in his article, see e.g. 2005a, 15). Raz himself seems to use both locutions interchangeably.

or say of someone that he has “good eyes”.<sup>28</sup> This might explain why Raz does not think of rational standards as reason-implying: neither does it seem that we necessarily have reasons to do whatever would be the result of the well-functioning of our capacities, nor does this well-functioning itself qualify as a response that a reason could possibly favour. However, such standards seem at least potentially capable of providing us with grounds for legitimate criticism, for even though they are not normative or reason-implying, they are often taken to be evaluative.

Now, the first problem with this suggestion is that there are good reasons to be sceptical that standards of proper functioning really are evaluative standards. The main examples of things with functions are artefacts and biological organisms. In the former case, the standards in question seem likely to be reducible to the intentions of the producers or the users of the artefact; in the latter case, they seem to be reducible to mechanisms that have been selected by evolution because they contributed to the fitness of the species.<sup>29</sup> These seem to be merely descriptive statements that need to be complemented with more substantial evaluative assumptions in order to be regarded as evaluative standards. Raz appears to respond to this concern by adding that the rational capacity the well-functioning of which is in question is a valuable capacity.<sup>30</sup> It is unclear, however, what this additional assumption accomplishes. The value of some X may influence what impact truths about the proper functioning of this X have on some evaluative questions, but it does not seem to affect whether such truths are themselves evaluative or merely descriptive. The statement “the engine of your car is broken” does not become an evaluative statement simply because it turns out that your car is valuable. It thus remains unclear whether truths about the proper functioning of our rational capacities have any normative or evaluative content, and consequently it remains unclear whether they are the kind of truths which could at least in principle ground criticism.

For the sake of the argument, I shall ignore this difficulty here. I shall assume that the content of such truths involves something that would make them at least in

---

<sup>28</sup> These are the kind of standards that according to Thomson (2008, esp. ch. 2) are set by “goodness-fixing kinds” – things that have some kind of function and can therefore be assessed as functioning better or worse.

<sup>29</sup> As Kolodny (2005, 552), among others, points out.

<sup>30</sup> Raz (2005a, 17–18; 2005b, 10–11).

principle suitable for grounding criticism. So let us tackle the question directly: Do facts about the faulty functioning of an agent's rational capacity provide grounds for appropriate criticism of that agent? The answer to this question is, I think, that they do not. In some cases, for example when such malfunctioning is the result of negligence on our part, we might be criticisable for our failure to maintain, practice, or exercise our capacities in a certain way, but these are cases in which we fail to respond to our reasons to maintain, practice, or exercise our capacities. The malfunctioning of someone's capacities itself provides no ground for criticism.

Indeed, the contrary is true. Inasmuch as a person's rational capacity ceases to function properly, this person's criticisability decreases. This is what partners of the elderly whose rational capacities cease to function well sometimes do not recognise when they go on criticising behaviour which results from such malfunctioning. Criticisability thus cannot be grounded in the malfunctioning of the rational capacity; rather, it presupposes its absence: We are criticisable only if our rational capacity functions properly.

This point, I think, is rather obvious. We can criticise a person for something only insofar as this person is accountable for it. But we are not accountable for the functioning of our capacities; at most we are indirectly accountable, because we are accountable for actions that have influence on the functioning of our capacities. So the idea that irrationality consists in faulty functioning, far from providing an understanding of the criticisability of irrationality, is in fact incompatible with it. Insofar as irrationality is thought of as criticisable, as Raz agrees, it cannot be understood in terms of malfunctioning, for malfunctioning undermines rather than grounds criticisability.

What went wrong with Raz's conception of rationality? I think that Raz's mistake is to construe irrationality too closely to what I have called the capacity sense of rationality. For him, irrationality is a degree of arationality – a breakdown of the rational capacity, though plausibly a temporal and local one. But irrationality is different from arationality, even a temporal and local one. If irrationality is criticisable, it must be a failure of the agent rather than a failure of one of his capacities.

## 2.5 Criticisability and virtue

We are still looking for a conception of criticisability which provides a viable alternative to the criticisability claim. In this section, I shall consider the possibility that criticism can be grounded in the fact that an agent acted in a less-than-virtuous way and thereby violated the standards of virtue. This suggestion is less problematic than the preceding one, for in contrast to statements about proper functioning, judgements about virtue are uncontroversially evaluative judgements. They are thus more clearly suitable starting points for a conception of criticisability. The suggestion is also of interest for our discussion because some philosophers try to capture the normativity of rationality in terms of virtue rather than reasons in order to avoid commitment to the claim that we have reason to conform to rational requirements.<sup>31</sup> Proponents of this view seem likely to counter the criticism argument with a virtue-related conception of criticisability.

Now, some moral philosophers claim that we ought always to act in accordance with virtue. On this view, the claim that criticisability can be grounded in the violation of the standards of virtue does not speak against the criticisability claim, of course. Hence, the thesis we have to consider is that *even though* we sometimes ought *not* to act in accordance with virtue, we are still criticisable if we do not.

The claim that we sometimes ought not to act in accordance with virtue is plausible. A famous example from the literature is the defeated squash player who plausibly ought not to shake hands with his opponent because he is so intemperate that he would then smash his racquet into his face.<sup>32</sup> Let us call this squash player Bert. We are assuming that, even though it would be virtuous of Bert to shake hands, he ought not to do it. The question is whether Bert can still be criticisable for not shaking hands. Now, a tendency to agree with this claim might stem from a source that we have to

---

<sup>31</sup> Dancy (2009) suggests this position as a response to the sceptical challenge to the normativity of rationality raised by Kolodny (2005). Kolodny (2005, 553–555) anticipates and rejects this view, though for different reasons than the ones I am bringing forward here. See also Svarvarsdóttir (2003) for a virtue-related account of rationality.

<sup>32</sup> Watson (1975, 210) originally brought up the example in a different context. Smith (1995, 163) discusses it as a challenge for his ideal rationality account of reasons. Williams (1995, 190) expresses the same kind of worries about McDowell's (1995) virtue-based conception of reasons.

disregard here: Bert's behaviour might reveal that he has not acted as he ought to have in the past. Perhaps he has without good reasons refused an offer to participate in psychological training, or more generally, he has not worked on his intemperate dispositions even if he could and ought to have done that. In this case, the example could not provide an argument against the criticisability claim, so we need to exclude this possibility.

We suppose instead that Bert has always been a conscientious follower of his reasons. He was born with a strong intemperate disposition from which he suffers, and he has worked and improved on this disposition all his life. But character traits are not so easily changed, and through the advice of his therapist and his friends Bert became convinced that the only way to deal with his disposition was to avoid situations in which he could react intemperately, situations like shaking hands after losing a match. We can add that the people around Bert are warned, and that he only plays squash with good friends who know him well. With these assumptions in the background, we can ask again: Is Bert criticisable for not shaking hands? Again, I think it is clear that the answer to this question is that he is not.

I am not saying that we cannot disapprove of his behaviour. Being a sore loser is a bad character trait, which gives Bert reasons to work on his dispositions, and might give others reasons for attitudes of dismissal or dislike towards him. But we can disapprove of many things that an agent is actually not accountable for, and certain unfortunate character traits may be among these things.

We have seen no reason to suppose that a violation of the standards of virtue can ground criticism independently of whether an agent acts as he ought to. If we accept that we sometimes ought not to act in accordance with virtue, we need to conclude that we are sometimes not criticisable for acting in less-than-virtuous ways.

## 2.6 Rationality and criticism

The criticisability claim states that a person is criticisable only if she violated a decisive reason. I have defended this claim in the last four sections, by giving an independent argument for it and showing that alternative conceptions of criticisability fail. If this is

correct, then the sceptic about the normativity of rationality ultimately needs to reject the first premise of the criticism argument; he has to deny that irrationality is criticisable. This section considers this possibility in more detail.

Let me start by clearing up a possible misunderstanding. I do not wish to deny that the word “irrational” is sometimes used in ways that do not involve a critical attitude. Some people use the word “rational” to refer to explicitly negative properties such as lack of sympathy and emotion, or cold and calculating behaviour, and they use the word “irrational” e.g. to refer to the capacities for spontaneous decision-making or intuitive judgement. Obviously, they need not think that what they call “irrationality” is criticisable. But this is just because they use “irrationality” in a sense different from the one used here and in most other philosophical treatments of rationality. Irrationality is here understood as the violation of a standard that applies to beings in virtue of their rational capacities. I believe that this is also the dominant meaning of “irrational” as it occurs in ordinary discourse (which does not, of course, mean that everyone using the word in this sense must be capable of explicating its meaning in the way suggested). But someone who uses the word “irrational” such that showing emotions or making spontaneous decisions turns out to be in itself irrational is probably better interpreted not as claiming that showing emotions or making spontaneous decisions by itself violates standards that apply to us in virtue of our rational capacities, but as using the word in a different sense. In any case, what is at issue here is whether irrationality ascriptions involve criticism when they are taken to involve the violation of a rational standard.

Nadeem Hussain proposes the following interesting case.<sup>33</sup> He reports that he often finds himself in discussions with his well-educated Muslim friends, who believe in God while freely admitting that they do not have sufficient reason to do so. He imagines himself claiming that this is irrational, while his friends would reject this assertion. On a natural interpretation of this disagreement, Hussain takes himself to utter a form of criticism when he calls his friends irrational, and his friends, too, understand it in this way. It seems that their disagreement is partly about whether there is something

---

<sup>33</sup> Compare Hussain (unpublished, §3). I would like to thank Nadeem Hussain for the permission to quote his unpublished paper.

wrong, something criticisable about believing in God while believing there is insufficient reason to believe in God. But then Hussain invites us to imagine him saying something like this: “I think you are irrational, but this is not meant as a criticism. I don’t think there is anything *wrong* with being irrational.” As Hussain indicates, this statement would undermine the point of their disagreement. It must seem to his friends that he has changed the topic or that they have been talking past each other all along. In any case, it seems appropriate for them to reply: “If you don’t think there’s anything wrong with us, then stop calling us irrational.”

As this example emphasises, ordinary attributions of irrationality are commonly understood as criticism. Moreover, the criticism involved seems to be agent criticism: when agents get called irrational, they do not merely understand this to mean that they fell short of some evaluative standard; they feel personally criticised for their responses. As far as I can see, this is pretty uncontroversial. The point has often been conceded and, to my knowledge, never been explicitly denied by any of the sceptics of the normativity of rationality.<sup>34</sup> So how can these sceptics avoid the criticism argument? We have already dismissed the first option, namely the denial of the criticisability claim. To avoid the conclusion of the criticism argument, then, the sceptics have to explain this feature of ordinary discourse in a way that is compatible with denying the first premise of the argument, according to which the violation of a rational requirement is criticisable. John Broome finds himself exactly in this position: He seems to accept both the criticisability claim as well as the observation that irrationality attributions involve criticism, but being agnostic about the normativity of rationality, he is unwilling to accept the conclusion of the criticism argument. He suggests an explanation that can

---

<sup>34</sup> The claim is widely shared by both sceptics and non-sceptics about the normativity of rationality. Compare: “When we accuse someone of irrationality, we are surely criticizing her” (Broome 2005a, 336; 2007a, 177). “We use the word ‘irrational’ to express the kind of criticism that we express with words like ‘senseless’, ‘stupid’, ‘idiotic’, and ‘crazy’” (Parfit 2011a, 33). “People who fail to pursue the means to their ends display or manifest a form of malfunctioning criticisable as a form of irrationality” (Raz 2005a, 18). “Instrumental irrationality is something that can be criticized” (Rippon 2011, 21). “There is a [...] pressure to hold others to account in respect of the rational requirements to which they are subject, and to regard those who violate them as appropriate objects of criticism” (Southwood 2008, 12). “Evaluations of rationality have a credible [...] critical force” (Svavarsdóttir 2003, 133). “The charge of irrationality is, in and of itself, a serious criticism” (Way 2009, 1).

capture the spirit of ordinary attributions of irrationality, without being committed to the criticisability of irrationality *itself*:

When we accuse someone of irrationality, we are surely criticizing her. How could we be entitled to do so if there is no reason for her to satisfy the requirements of rationality in the first place? [...] There is another possible explanation of why a charge of irrationality is a criticism. [...] Plausibly, you ought to have the rational faculty. If you do not satisfy some particular rational requirement, that is evidence that you do not have this faculty, or at least that you do not have it to the highest degree. So it is evidence that you are failing to achieve something that, plausibly, you ought to achieve. That makes it grounds for justified criticism.<sup>35</sup>

According to Broome, then, we are generally justified in criticising an irrational person. But this is not because irrationality is criticisable in itself, but because it is *evidence* that a person violates an ‘ought’ or decisive reason. And it is evidence for this because (i) persons ought to have the rational faculty, and (ii) if a person is irrational, this is evidence that she does not have the rational faculty.

This explanation has many flaws. First, it is not plausibly true that (i) persons ought to have the rational faculty. For one, having the rational faculty is neither an action nor a judgement-sensitive attitude, nor anything else that could possibly be a response to the recognition of reasons, and it is therefore dubious to say that we have reasons or ought to “have” the rational faculty. We can have reasons to maintain, exercise, improve, or protect the rational faculty, but it is unclear what it is supposed to mean that in addition to these reasons, we also have reasons to *have* the rational faculty. For another, it seems very plausible that our having the rational faculty is a precondition for our being subject to normative requirements in the first place, and a demand that only applies when it is satisfied does not make sense. We should therefore reject the claim that we ought to have the rational faculty.

---

<sup>35</sup> Broome (2007a, 177). Cf. also Broome (2005a, 336).

Second, it is also false that (ii) the irrationality of a person is evidence that this person does not have the rational faculty. Quite the contrary, having the rational faculty is a necessary condition for being irrational. Irrationality is the failure to meet the standards of rationality, but if you do not have the rational faculty, you are not *failing* to meet these standards, you are not subject to them. You are arational, not irrational.

Third, even if both of these assumptions were true, they could not explain why irrationality ascriptions involve criticism. For one, even if it were true that we ought to have the rational faculty, it would be clear that a violation of this very requirement could not generally ground any criticism. Apart from some radical cases (say, people who voluntarily wreck their brains by continually taking highly damaging substances), we are normally not accountable for whether and to what degree we have the rational faculty, and so we cannot be criticisable for it. For another, we can perfectly well imagine someone who *knows* that the person he believes to be irrational is in possession of the rational faculty. For this person, the irrationality cannot be evidence for the absence of the rational faculty, but this does not change the fact that we would understand a charge of irrationality as criticism.

Hence, Broome's explanation fails on all counts. It does not follow that there cannot be an explanation of the kind that Broome envisages, but it seems likely that other explanations meet similar difficulties, especially the last one mentioned: Whatever the irrationality is supposed to be evidence for, it seems possible to imagine someone who knows in a particular case that the irrationality is no evidence for it, but whose irrationality ascription would still be understood as a form of criticism. At least as long as nobody brings forward a convincing example of such a property, we have to conclude that the only available plausible vindicating explanation of our practice of irrationality ascriptions is that irrationality is indeed criticisable. And this means that someone who denies the conclusion of the criticism argument is committed to a quite radical error theory about ordinary irrationality ascriptions. He is committed to holding that we are all deeply misled in believing that there is anything criticisable

about irrationality. And he owes us an explanation of why we all went wrong in thinking this.<sup>36</sup>

This, I think, puts a lot of pressure on what I have called *the mixed view* – the view that there are structural requirements of rationality that are not normative. First, an error theory about a whole range of judgements and attitudes is quite generally a very strong and unattractive philosophical position. It might be true, of course, but as a theoretical option regarding a given subject matter, it is generally considered the last resort.

Second, it seems questionable to me that our rationality and irrationality ascriptions could remain stable after we give up the idea that irrationality is criticisable. Given the deep-rootedness of our inclination to understand irrationality as worthy of criticism, the error theory of this part of our rationality ascriptions might well reveal that our talk of rationality and irrationality is incoherent on the whole. This is suggested by Hussain's case, if we assume that his friends have a point in saying "If you do not think there is something wrong with us, then do not call us irrational." In this respect, "irrational" seems like "unchaste": If you do not mean it in a bad way, you do not mean it at all. If, however, the denial of the normativity of rationality undermines the standard-relative sense of rationality as such, then the mixed view – being based on the idea of non-normative requirements of rationality – does not make sense to begin with. Whether this is so depends, of course, on how persistent the connection between judgements of irrationality and criticisability is. Though I have not given a conclusive argument for this claim, the preceding remarks at least suggest that the assumption of criticisability is part of the very content of irrationality attributions.<sup>37</sup> If this is true, then scepticism about the normativity of rationality collapses into scepticism about rationality as such.

---

<sup>36</sup> Note that Kolodny's (2007a; 2008a) error theory aims to explain something entirely different: it does not aim to explain why it seems to us that irrationality is criticisable, but why it seems to us that first-order coherence is rationally required. It is thus concerned with the first thesis of the myth view (that there are no first-order rational requirements), not with the second (that rational requirements are not necessarily normative).

<sup>37</sup> This also seems to follow from some of the above-mentioned characterisations of irrationality, for example from Parfit's claim that "we *use* the word 'irrational' to express [...] criticism" (2011a, 33, my emphasis), or Way's contention that "the charge of irrationality is, *in and of itself*, a serious criticism" (2009, 1, my emphasis).

The upshot of the criticism argument, then, is this. We have seen no reason to doubt the claim that persons are criticisable only if they violate a decisive reason. Therefore, sceptics of the normativity of rationality have to deny that the violation of a rational requirement is as such criticisable. This is a significant challenge, for it seems clear that ordinary attributions of irrationality presuppose such criticism. The sceptic thus either needs a sophisticated vindicating explanation of this fact which is compatible with the denial of the criticisability of irrationality – a task that no one has so far accomplished. Or he is committed to a radical error theory about ordinary attributions of irrationality. But such an error theory, I have suggested, might well undermine the practice of rationality and irrationality ascriptions on the whole, to the effect that scepticism about the normativity of rationality collapses into scepticism about rationality in general.

## 2.7 The obscurity of non-normative requirements of rationality

This concludes my discussion of the criticism argument. In this final section, I offer a further reason to think that scepticism about the normativity of rationality is at risk of collapsing into scepticism about rationality in general. The reason is that the very notion of a non-normative requirement of rationality, on which the possibility of a mixed view – which is sceptical about the normativity of rationality but not of rationality itself – relies, is obscure.

If we could describe the properties of rationality and irrationality in purely descriptive terms and then derive from this a notion of a rational requirement in terms of necessary conditions for having the property of rationality, we would have an unproblematic non-normative notion of a rational requirement. But as I have argued in the last chapter, this would not be the relevant notion of a rationality-related standard of correctness, which involves the idea of a code or system of rules that prescribes certain responses or omissions. In John Broome's terminology, we are concerned with requirements in the code sense rather than the property sense. And the question is whether we can understand *this* notion of a 'rational requirement' – the notion of a rational code that prescribes certain responses – in a non-normative way.

Above, I have argued myself that the existence of a requirement in the code sense does not entail its normativity. We are happy to accept the existence of certain rules of etiquette, for example, without supposing that there is necessarily any reason for us to follow them. Nevertheless, it seems that the code sense of ‘requirement’ is in need of explanation in a way that the property sense is not: It is clear what it means to say that something is a necessary condition for having a certain property, but to say that something is prescribed by a certain code calls for explanation. What does it mean to say that there is a code or system of rules that requires us to do something?

Here are two possible answers to this question. First, it might mean that we have certain kinds of reasons, or ought, to do it. I find this answer plausible in the case of requirements of morality or prudence.<sup>38</sup> According to this view, to say that morality requires us to do certain things is to say that there are reasons of a certain kind (or reasons that weigh in a particular way) in favour of doing it. Another possible answer is, very roughly speaking, that there is a convention or a practice that sets up this standard. This seems to me the correct answer in the case of requirements of etiquette or grammar. In some cases, it is controversial whether a requirement is to be explained in one way rather than the other. For example, legal positivists and proponents of natural law theory disagree about whether the requirements of law are to be understood in terms of a convention or practice or (at least in part) in normative terms.

It thus seems that talk of requirements can be understood, very broadly, in two different ways. Consider Broome’s own examples. Besides rational requirements, Broome mentions requirements of morality, prudence, evidence, convention, the law, and Catholicism.<sup>39</sup> Morality, prudence, and evidence requirements can be understood as statements about reasons. Requirements of etiquette, religion, and the law are plausibly understood as conventions. Proponents of natural law theory insist that requirements of the law have to be understood (at least partly) in normative terms, and some Catholics are likely to maintain the same for the requirements of Catholicism. But however these disagreements resolve, Broome’s examples do not give us any reason

---

<sup>38</sup> Way (2010b, 1065) also finds it plausible.

<sup>39</sup> Broome (2007b, 2–3).

to believe that there are requirements that could *neither* be explained in normative *nor* in conventional terms.

The trouble for the proponent of the mixed view is that he cannot give either of these explanations of rational requirements. He obviously cannot give the normative explanation, for he thinks that rational requirements are not necessarily normative. He thus seems committed to the conventional explanation. But the conventional explanation is obviously false in the case of rational requirements. Whether it is irrational to believe contradictions or fail to take the believed means to one's ends is not a matter of contingent fact that we detect empirically through the observation of certain practices. In contrast, rational requirements seem necessary truths that are detectable through non-empirical modes of understanding such as reflective equilibrium.<sup>40</sup> If there are any requirements which are not to be understood as conventions, then certainly rational requirements will be among them.

The proponent of the mixed view is thus committed to quite peculiar entities. Being necessary demand-like truths that are recognisable through reflective equilibrium, rational requirements have all the properties that normative truths have, except that the proponent of the mixed view insists that they are not normative. But what are they then?

There have been attempts to reduce the notion of a rational requirement to statements of psychological facts or counterfactual statements about reasons. Both of these proposals bring along their own problems, which are discussed in later chapters.<sup>41</sup> My point here is merely that proponents of the mixed view face the challenge of clarifying the notion of a non-normative requirement of rationality. As long as it is not clear that this notion is coherent, it is not clear whether there is the theoretical space for the very view that most philosophers currently participating in the debate seem to embrace.

---

<sup>40</sup> Way (2010b, 1065) makes similar points against Broome's conception of conventional and normative requirements as a unitary phenomenon.

<sup>41</sup> See Chapters 5.5–5.6.

## 2.8 Summary and prospects

Let me take stock. The aim of this chapter was to motivate the normativity of rationality. I have done so by exploring the connection between irrationality and criticisability. I have defended a conception of criticisability according to which criticism is appropriate only if a decisive reason has been violated. For the sceptic of the normativity of rationality, this means that he has huge difficulties explaining the fact that ordinary irrationality ascriptions involve criticism. At the end of the day, he seems committed to a highly revisionary error theory about an essential part of ordinary ascriptions of rationality and irrationality.

I have also argued that because of the close connection between irrationality attributions and criticism, there is reason to think that scepticism about the normativity of rationality ultimately collapses into scepticism about rationality in general. This is also suggested by doubts about whether the notion of a non-normative requirement of rationality is coherent. Whether or not this stronger conclusion can be maintained, however, I hope at least to have shown that abandoning the normativity of rationality comes with high theoretical costs and should be considered our last resort.

After emphasising these problems for the mixed view, I shall now turn to the structuralist view, by considering more closely the idea of structural requirements of rationality (Chapter 3), and the arguments to the effect that such requirements cannot be normative (Chapters 4 and 5).



## Structural requirements of rationality

---

This chapter provides an examination of some elementary questions about structural requirements of rationality, which will serve as a basis for the discussion in forthcoming chapters. I discuss these questions, so far as this is possible, in abstraction from the normative question about rationality. I begin by introducing the main disagreement about requirements of structural rationality, namely whether they take wide or narrow scope (3.1). I go on to discuss several questions about the form of such requirements: whether their objects are propositions or responses (3.2); whether they are conditional or unconditional (3.3); whether they are synchronic or diachronic (3.4); and whether they govern states or processes (3.5). I argue that on both the narrow- and the wide-scope readings, structural requirements of rationality are best understood as conditional and diachronic state-requirements that take responses as their object.

### 3.1 Wide- or narrow-scope requirements?

Let me briefly call to mind the two claims that, according to what I have argued in Chapter 1, motivate the idea of structural requirements of rationality. The first is the hypothesis that having certain combinations of attitudes is necessarily irrational, as is reflected by the claims (A)-(D). For the purpose of this discussion, it will suffice to repeat only the first of these claims here:

- (A) *Weakness of will*: If A believes that she ought to  $\phi$ , and A does not intend to  $\phi$ , then A is irrational.

The second claim is that irrationality involves the violation of a standard or requirement of rationality. Taken together these claims suggest that structural irrationalities,

such as the one described by (A), correspond to structural *requirements* of rationality. Above I formulated the (A)-corresponding requirement as follows:

(A)<sub>1</sub> *Self-governance requirement*: Rationality requires A to intend to  $\phi$  if A believes that she ought to  $\phi$ .

I have chosen this formulation deliberately because it lends itself to two different ways of understanding the requirements in question, an ambiguity that we can now dismantle. The kinds of structural irrationality I focus on involve at least two attitude-states that are in tension with one another.<sup>1</sup> On a natural interpretation, structural requirements of rationality are stated in the following general form: “If A is in attitude-state  $a_1$ , then A is required to be in attitude-state  $a_2$ ”. Such requirements are conditional on one attitude-state and have the other attitude-state as their object. We can call the first attitude-state the *antecedent state* and the second the *consequent state*. Because only the consequent state is the object and thus part of the *scope* of the requirement, such requirements are often called ‘narrow-scope requirements’.<sup>2</sup> According to the narrow-scope account, then, the structural requirements that correspond to the forms of irrationality described by (A)-(D) are:

(A)<sub>2</sub> *Self-governance requirement, narrow-scope*: If A believes that she ought to  $\phi$ , then rationality requires A to intend to  $\phi$ .

(B)<sub>2</sub> *Instrumental requirement, narrow-scope*: If A intends to  $\phi$ , and if A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, then rationality requires A to intend to  $\psi$ .

(C)<sub>2</sub> *Evidence-belief requirement, narrow-scope*: If A believes that she has sufficient evidence that  $p$ , then rationality requires A to believe that  $p$ .

<sup>1</sup> Recall that I use ‘attitude-state’ as the generic term for both attitudes and *lack* of attitudes, which allows me to say, e.g., that in cases of weakness of will, there is a tension between the attitude-state of believing one ought to  $\phi$  and the attitude-state of not-intending to  $\phi$ .

<sup>2</sup> Technically speaking, the scope of a term is the smallest part of a proposition in which the term occurs that it affects logically. Since according to the interpretation at hand, the smallest part of a proposition about rational requirements that “rationality requires” affects logically is concerned with the consequent state only, this is a *narrow-scope* interpretation. In contrast, an interpretation according to which the smallest part of a proposition about rational requirements that “rationality requires” affects logically is concerned with both the antecedent and the consequent state, is a *wide-scope* interpretation.

(D)<sub>2</sub> *Modus ponens requirement, narrow-scope*: If A believes that p, and if A believes that  $p \rightarrow q$ , then rationality requires A to believe that q.

As several authors have emphasised, however, structural requirements of rationality can also be understood in a different way.<sup>3</sup> According to this other interpretation, the requirements in question are not conditional on the antecedent attitude. Instead, both the antecedent and the consequent attitude are part of the scope of the requirement. Therefore, such requirements are often called ‘wide-scope requirements’:

(A)<sub>3</sub> *Self-governance requirement, wide-scope*: Rationality requires of A that [if A believes that she ought to  $\phi$ , then A intends to  $\phi$ ].

(B)<sub>3</sub> *Instrumental requirement, wide-scope*: Rationality requires of A that [if A intends to  $\phi$ , and if A believes that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then A intends to  $\psi$ ].

(C)<sub>3</sub> *Evidence-belief requirement, wide-scope*: Rationality requires of A that [if A believes that she has sufficient evidence that p, then A believes that p].

(D)<sub>3</sub> *Modus ponens requirement, wide-scope*: Rationality requires of A that [if A believes that p, and if A believes that  $p \rightarrow q$ , then A believes that q].

#### *Differences between wide- and narrow-scope requirements*

Both the narrow-scope requirements (A)<sub>2</sub>-(D)<sub>2</sub> and the wide-scope requirements (A)<sub>3</sub>-(D)<sub>3</sub> could explain the structural irrationality claims (A)-(D), but they differ in some other important respects. Let me briefly recall two such differences that will not be new for those familiar with the literature. The first difference concerns the *satisfiability*

---

<sup>3</sup> For the distinction between a wide- and a narrow-scope understanding of rational requirements, see esp. Broome (1999), (2007b), (2007c). Note that within the period between these essays, Broome has changed his vocabulary to make room for the question of whether rational requirements are normative. In essence, the distinction has been emphasised by a number of other authors before. Philosophers advocating wide-scope versions of structural requirements include Bittner (1980, 220), Dancy (1977), Darwall (1983, 16), Gensler (1985), Greenspan (1975), and Hill (1973). More recently, the wide-scope account has been defended (besides Broome) by Brunero (2010a) and (2012), Dancy (2000a, ch. 5, esp. 70-76), Wallace (2001, 16–17), as well as Way (2010a) and (2011). See Schroeder (2005a) for an interesting discussion of whether Kant can be interpreted as holding a wide-scope view about hypothetical imperatives.

*conditions* of the requirements. Suppose you believe that you ought to  $\phi$ , but you do not intend to  $\phi$ . You violate both the narrow-scope requirement (A)<sub>2</sub> and the wide-scope requirement (A)<sub>3</sub>. But you can satisfy (A)<sub>2</sub> only by intending to  $\phi$ , while you can satisfy (A)<sub>3</sub> either by intending to  $\phi$  or by giving up the belief that you ought to  $\phi$ , for in both cases you make it true that *if* you believe that you ought to  $\phi$ , then you intend to  $\phi$ .<sup>4</sup> So by ceasing to believe that you ought to  $\phi$ , you satisfy (A)<sub>3</sub> but not (A)<sub>2</sub>. Further, you do not violate (A)<sub>2</sub>, since if you do not believe that you ought to  $\phi$ , you are not subject to (A)<sub>2</sub>. Recall that (A)<sub>2</sub> is a conditional requirement that applies to you only if you have this belief. So if you stop believing that you ought to  $\phi$ , you neither satisfy nor violate (A)<sub>2</sub> – you avoid it entirely.<sup>5</sup> Consequently, we have to distinguish between satisfying, violating, and avoiding a requirement. And the question of whether the narrow-scope interpretation or the wide-scope interpretation of rational requirements is more plausible, depends in part on which responses we would count as satisfying rather than avoiding a requirement.

The second important difference between narrow- and wide-scope requirements is that the former allow for *detachment* of a requirement to have a particular attitude, while the latter do not generally allow such detachment in the first place. For the purpose of illustration, compare the two formalised interpretations of the self-governance requirement:

---

<sup>4</sup> This presupposes that the conditional in question is the material conditional ‘ $p \rightarrow q$ ’ which is equivalent to the disjunction ‘ $\neg p \vee q$ ’ and is thus true if  $p$  is false – a standard assumption in the literature, especially by advocates of the wide-scope interpretation. Since there are reasons to doubt that conditionals in natural language have the same truth conditions as the material conditional, this may seem problematic if the wide-scope account is understood as a logical interpretation of certain conditionals in natural language (cf. Evers 2011). However, as I shall argue in the next section, the wide-scope account need not be understood in this way; indeed there are independent reasons to defend wide-scope requirements directly as requirements on disjunctive responses rather than requirements on conditionals. Since this will be my preferred understanding of wide-scope requirements, the assumption that the conditional expresses a material conditional seems to me innocuous in the present context.

<sup>5</sup> For the difference between satisfying, violating, and avoiding, see Vranas (2008). Broome (2007b, 38) similarly distinguishes between complying, infringing and avoiding requirements, while Lord (2011) distinguishes complying and violating from *exiting*, and Schroeder (2004, 352–353) distinguishes satisfying and violating from *escaping*. Exiting and escaping seem to be special cases of avoiding, namely cases in which an agent avoids a requirement that applied to him or her before.

(A)<sub>2</sub>  $BO\phi \rightarrow R(I\phi)$

(A)<sub>3</sub>  $R(BO\phi \rightarrow I\phi)$

I take the term ‘ $BO\phi$ ’ to refer to the proposition that A believes that she ought to  $\phi$ , the term ‘ $I\phi$ ’ to refer to the proposition that A intends to  $\phi$ , and the term ‘ $R$ ’ as the operator ‘rationality requires’. The narrow-scope requirement (A)<sub>2</sub> allows us to derive, by *modus ponens*, a requirement to intend to  $\phi$ , in the case that A believes that she ought to  $\phi$ :

*Valid Detachment:* (1a)  $BO\phi \rightarrow R(I\phi)$

(2a)  $BO\phi$

(3a)  $R(I\phi)$

The wide-scope requirement, in contrast, does not generally allow for such a detachment, for the following inference is not valid:

*Invalid Detachment:* (1b)  $R(BO\phi \rightarrow I\phi)$

(2b)  $BO\phi$

(3b)  $R(I\phi)$

So according to the wide-scope reading, we cannot conclude that A is rationally required to intend to  $\phi$  from the fact that A believes that she ought to  $\phi$ .

Why is the inference from (1b) to (3b) invalid? It is clear that, in contrast to the inference from (1a) to (3a), it is not validated by propositional calculus. So in order for it to be valid, the specific logic of requirements would need to contain the following axiom:

*Factual detachment:* If rationality requires of A that  $p \rightarrow q$ , and p, then rationality requires of A that q.<sup>6</sup>

It would follow that if rationality requires of A that [if she  $\phi$ -s, then she  $\psi$ -s], and if A does in fact  $\phi$ , then rationality requires of A that she  $\psi$ -s. But this is implausible. Even if A in fact  $\phi$ -s, this does not undermine the fact that she *can* satisfy the requirement that [if she  $\phi$ -s, then she  $\psi$ -s] either by  $\psi$ -ing or by not  $\phi$ -ing. That is, even if A in fact  $\phi$ -s, she *can* satisfy the relevant requirement without  $\psi$ -ing. Therefore, a requirement to  $\psi$  cannot plausibly be detached.<sup>7</sup>

That wide-scope requirements, in contrast to narrow-scope requirements, do not allow for detachment, is probably the most significant difference between the two interpretations, not least because of the considerable consequences for the normative question about rationality. For suppose that A has a crazy belief about what she ought to do – perhaps she thinks she ought to extinguish all human life. And suppose that the narrow-scope requirement (A)<sub>2</sub> is true. By valid detachment, we have to conclude that A is rationally required to intend to extinguish all human life. And if we suppose that rationality is strictly normative, we have to conclude that A *ought* to intend to extinguish all human life. This is the ‘bootstrapping’ worry that I shall discuss in more detail in later chapters. The problem is that if we assume narrow-scope requirements of structural rationality, and if we suppose that rationality is normative, it follows that we can make it the case that we ought to do horrible things, simply by adopting some crazy attitude. This is generally seen as an unacceptable result, and most authors therefore agree that narrow-scope requirements of structural rationality are incompatible with

---

<sup>6</sup> This formulation of *factual detachment* presupposes, as does the whole discussion in this section, that rational requirements operate over propositions – an assumption that I will question in the next section. However, the principle can also be formulated in a way that is independent of this assumption. Accordingly, if rationality requires A to [ $\phi$  or  $\psi$ ], and A does not  $\phi$ , then rationality requires A to  $\psi$ .

<sup>7</sup> See esp. Broome (1999; 2013, 123–24). As far as I can see, this is not disputed in the literature on rational requirements. For further discussion of detachment, see also Chapters 3.2 and 4.4–4.5.

the normativity of rationality.<sup>8</sup> So this is one (if not the main) reason why it matters whether rational requirements take wide or narrow scope.

*Logical relations between wide- and narrow-scope requirements*

I shall conclude this section with a brief discussion of the logical relations between wide- and narrow-scope requirements. Since wide-scope requirements do not generally allow for detachment, while narrow-scope requirements do, it is clear that wide-scope requirements do not entail their narrow-scope counterparts. It is sometimes argued, however, that narrow-scope requirements entail their wide-scope counterparts and are, in this respect, logically stronger.<sup>9</sup> This is a dubious assumption, however. Consider again the formalised versions of the self-governance requirement:

(A)<sub>2</sub>  $BO\phi \rightarrow R(I\phi)$

(A)<sub>3</sub>  $R(BO\phi \rightarrow I\phi)$

It is clear that the supposed entailment  $(A)_2 \Rightarrow (A)_3$  is not validated by the propositional calculus; assuming that the entailment holds thus presupposes the endorsement of an axiom of the specific logic of rational requirements or ‘oughts’ that validates it. But now consider the two requirements that appear in these claims. While both claims (A)<sub>2</sub> and (A)<sub>3</sub> are supposed to be true in every possible world, the requirement occurring in (A)<sub>2</sub> is conditional on a contingent attitude of A’s. That is, the requirement itself holds only in those worlds in which particular agents have particular attitudes, and it is a requirement only for those particular agents, and not for all. In contrast, the requirement occurring in (A)<sub>3</sub> holds in every possible world, and it is a requirement for every agent, independently of any contingent facts about this agent’s attitudes.

So the claim that wide-scope requirements such as (A)<sub>3</sub> can be derived from narrow-scope requirements such as (A)<sub>2</sub> amounts to the claim that conditional agent-

---

<sup>8</sup> Korsgaard (2009a) seems to be an exception, though she also seems to have a more restrictive conception of the antecedent attitudes that appear as the condition of narrow-scope requirements (cf. Korsgaard 1997). See Wallace (2001, 3–10) for an interesting discussion of Korsgaard’s view.

<sup>9</sup> See e.g. Korsgaard (2009a, 42): “You can derive a wide-scope requirement from a narrow-scope one, but you can’t derive a narrow-scope requirement from a wide-scope one.”

relative requirements presuppose the existence of unconditional agent-neutral requirements. As especially Mark Schroeder has emphasised, this rules out substantial philosophical views according to which *all* requirements are conditional agent-relative requirements, including many desire-based conceptions of rationality and normativity.<sup>10</sup> Some arguments in the literature could be understood as claiming that the philosophical views in question are incoherent for this reason.<sup>11</sup> But one cannot refute such views simply by adopting an axiom according to which they are false. So the view that narrow-scope requirements entail wide-scope requirements is a substantial philosophical thesis that needs to be supported by independent argument: it cannot be taken for granted. Here I shall not assume that any such entailment holds.

Perhaps the view that narrow-scope requirements entail wide-scope requirements can be given a less controversial interpretation. According to this suggestion, (A)<sub>2</sub> does not entail the unconditional wide-scope requirement (A)<sub>3</sub>, but it does entail the *conditional* wide-scope requirement (A)<sub>4</sub>:

(A)<sub>4</sub> *Self-governance requirement, conditional wide-scope*: If A believes that she ought to  $\phi$ , then rationality requires of A that [if A believes that she ought to  $\phi$ , then A intends to  $\phi$ ].

Put formally:

(A)<sub>4</sub>  $BO\phi \rightarrow R(BO\phi \rightarrow I\phi)$

The claim that narrow-scope requirements such as (A)<sub>2</sub> entail wide-scope requirements such as (A)<sub>4</sub> avoids the complaint that this would allow a derivation of an unconditional

---

<sup>10</sup> See Schroeder (2004, 340–341) and (2007a, ch. 3 and 5.2) for discussion. As far as I can see, Schroeder never explicitly discusses the possibility that narrow-scope requirements straightforwardly *entail* wide-scope requirements, but only the claim that they need to be *explained* by wide-scope requirements (because of what Schroeder calls the “standard model theory”) and the claim that wide-scope requirements have to be accepted as uncontroversial by everyone.

<sup>11</sup> Hampton’s argument to the effect that “the force of hypothetical imperatives is dependent on, and is at least in part constituted by, the force of some antecedent categorical imperative that is in part definitive of instrumental rationality” (1998, 166) *could* be interpreted in this way (cf. Schroeder 2007a, 46–47 and 91).

agent-neutral requirement from a conditional agent-relative requirement, since both (A)<sub>2</sub> and (A)<sub>4</sub> are conditional and agent-relative. But on what grounds should it be accepted? It seems that the underlying idea must be that the content of the requirement that occurs in (A)<sub>2</sub> entails the content of the requirement that occurs in (A)<sub>4</sub>. So the underlying principle that is supposed to validate the inference in question seems to be this:

*Inheritance*: If rationality requires of A that p, and p entails q, then rationality requires of A that q.<sup>12</sup>

However, I have already rejected this principle when arguing against the view that rational requirements can be understood as necessary conditions for having the property of rationality. *Inheritance* allows, among other things, the introduction of a disjunction within the scope of a rational requirement. This implies that every rational requirement entails infinitely many disjunctive requirements, and, consequently, that we satisfy a rational requirement literally every time we adopt some belief or other attitude, no matter how stupid this attitude may be. I think that this conclusion is not acceptable.<sup>13</sup> So I will not assume that (A)<sub>2</sub> entails (A)<sub>4</sub> either. I shall proceed on the assumption that narrow-scope and wide-scope requirements are logically independent approaches to the explanation of structural rationality.

### 3.2 Proposition- or response-requirements?

Up to this point of the exposition, I have neglected some interesting distinctions and made some implicit assumptions that I now want to consider explicitly.

First, some of the formulations I gave imply that ‘rationality requires’ is an *operator* that ranges over *propositions*. I have been speaking of rationality requiring *that* A

---

<sup>12</sup> Cf. van Fraassen (1973, 12) who endorses a similar principle in deontic logic: “logical consequences of what ought to be, ought to be”. See also Broome (2007b, 19), who rejects the principle. As in the case of *factual detachment*, the above formulation of *inheritance* presupposes that rational requirements operate over propositions. The following formulation would do the same job, however: if rationality requires A to  $\phi$ , and if A’s  $\phi$ -ing entails A’s  $\psi$ -ing, then rationality requires A to  $\psi$ .

<sup>13</sup> Here I agree with Broome (2007b, 19–22).

believes something, for example. On an alternative conception, rational requirements are *relations* between an *agent* and a *response* such as an action, a believing, or an intending. On this reading, we should say that rationality requires A *to* believe something rather than *that* A believes it. Second, the formulations I gave suggest that structural requirements of rationality are *synchronic* rather than *diachronic*. They do not refer to different points of time, such that rationality would require you to show a certain response at a certain time, conditional on your believing or intending something at an *earlier* time. Third, I have been assuming that rationality requires certain *states* (or combinations of states), such as beliefs and intentions, while some authors suggest that rationality actually requires certain *processes*, such as reasoning or the *forming* of certain mental states.

All of these assumptions have been questioned. Unfortunately, they are often lumped together. Consider, for example, the first sentences of Niko Kolodny's paper "State or Process Requirements?":

State requirements require that you be a certain way at a given time. Process requirements require you to do something over time, where 'do' is understood broadly, so as to include forming and revising beliefs.<sup>14</sup>

Kolodny seems to think that the above-mentioned distinctions fall together: Either rational requirements require (i) that (ii) you be in a certain state (iii) at a certain point in time, i.e. they take propositions, require states, and are synchronic. Or they require you (i) to (ii) do something (iii) over time, i.e. they relate agents to responses rather than operating over propositions; they require doings rather than states; and they are diachronic.

I believe that we are well advised to keep these questions apart. I shall argue that there are good reasons to take rational requirements to be response-requirements rather than proposition-requirements, and that we should think of them as diachronic rather than synchronic. But I doubt that the considerations that Kolodny brings

---

<sup>14</sup> Kolodny (2007b, 371).

forward really count in favour of understanding the structural requirements that correspond to (A)-(D) as process- rather than state-requirements.

*The propositional and the relational account*

I shall begin by considering the question of whether rational requirements operate over propositions or relate agents to responses. Let us call the view that a given requirement operates over propositions a ‘propositional account’ of this requirement and the respective requirement a ‘proposition-requirement’, and let us call the view that a given requirement relates agents to responses a ‘relational account’ of this requirement and the respective requirement a ‘response-requirement’. There is an extensive debate about the question of whether the propositional or the relational account is correct for ‘ought’.<sup>15</sup> I cannot discuss this question here in a satisfactory way, but merely indicate why it seems to me plausible that ‘ought’ (in the sense in which it is used in deliberation) relates agents to responses, and is not well captured by the propositional account. The propositional account treats all ‘ought’-statements on the model of the ‘ought’ of “there ought to be world peace” – as stating, in effect, that a certain state of affairs would be desirable. To cite an old example given by Gilbert Harman, in this evaluative sense of ‘ought’ there is no difference between the statements “Jones ought to examine Smith” and “Smith ought to be examined by Jones”; both merely state that a certain state of affairs – namely one in which Jones examines Smith – is desirable.<sup>16</sup> Read in this way, nothing has been said about whether *Jones* in particular is accountable for examining Smith; whether it would be appropriate for us to advise or demand of Jones to examine Smith; or whether “I ought to examine Smith” would be an appropriate conclusion for Jones’s first-personal deliberation about what to do. However, it seems that there is a natural sense in which the former statement, but not the latter, is used in a way that carries exactly these implications. In treating both statements as equivalent, the propositional account thus fails to capture the deliberative sense of

---

<sup>15</sup> See Schroeder (2011b) for a recent survey of the debate, an excellent discussion of the question, and a convincing argument in favour of the relational account.

<sup>16</sup> See Harman (1973, 235–236).

‘ought’; this ‘ought’ cannot be understood as a sentential operator, it needs to be interpreted as a relation that has a separate argument place for an agent.<sup>17</sup>

Now, if the relational account is correct for the deliberative ‘ought’, it seems plausible to assume that it is also correct for rational requirements. It *must* be correct if rational requirements are to be understood in terms of the (factual or counterfactual) deliberative ‘ought’. But even if the notion is independent, the step seems plausible. For recall that we do not understand rational requirements as necessary conditions for having the property of rationality, but as a code or rule that in some way or other tells us what to do. It seems part of the point of the concept of a requirement in this sense that it could at least *possibly* figure as an advice, demand, or deliberative conclusion, and should therefore share at least the logical form of the deliberative ‘ought’. If it is merely conceptually possible that we sometimes ought to do what we are rationally required to do (i.e. if the content of a rational requirement could possibly be the content of a deliberative ‘ought’), and if ‘oughts’ are response-requirements rather than proposition-requirements, then rational requirements must also be response-requirements.

#### *A problem for the wide-scope account?*

It has been argued that this is a problem for the view that requirements of structural rationality take wide scope.<sup>18</sup> It may have been no coincidence that I formulated the wide-scope requirements as proposition-requirements, and the narrow-scope requirements as response-requirements. To see this, it is helpful to have a look at the context

---

<sup>17</sup> Broome (2013, 12–25) and Wedgwood (2006, 131–137) also argue that the deliberative ‘ought’ needs an argument place for an agent, but maintain that ‘ought’ relates agents to propositions rather than responses. According to the two of them, ‘ought’ is neither a sentential operator, nor a relation between an agent and response, but a relation between an agent and a *proposition*. This solves the problem of accountability, for it allows us to say that it is Jones in particular who ‘ought that Jones examines Smith’. As Schroeder (2011b, 23–33) argues, however, it also creates new problems. In particular, it entails that statements are meaningful which are very hard to make sense of, e.g. “Smith ought that Jones examines Smith”. I agree with Schroeder that this is a good reason to prefer an account according to which ‘ought’ relates agents directly to response types rather than propositions (where the response types might in turn be understood as properties of the agent). Though I will proceed on this assumption, I do not think that anything substantial in the arguments to come hinge on it.

<sup>18</sup> See Schroeder (2004, 342–344) and (2011b, 35–36).

in which the wide-scope strategy was originally deployed, namely the so-called modal fallacy.<sup>19</sup> Consider the following argument:

- (1a) If today is Monday, then tomorrow must be Tuesday.
- (2a) Today is Monday.
- (3a) Tomorrow must be Tuesday.

The problem with this argument is that even though (1a) seems to be true, and (2a) can be stipulated, (3a) is false: It is not necessarily the case that tomorrow is Tuesday. Even if in the actual world today is Monday, there is a possible world in which tomorrow is another day than Tuesday, namely a possible world in which today is not Monday. So something must be wrong with the argument. The wide-scope strategy to solve this problem is to say that modal conditionals of the form

*Modal conditional:* If  $p$ , then it is necessary that  $q$

are ambiguous between a narrow-scope and a wide-scope interpretation:

*Narrow-scope modal conditional:*  $p \rightarrow \Box q$

*Wide-scope modal conditional:*  $\Box(p \rightarrow q)$

Though on the surface it seems that natural expressions of modal conditionals such as (1a) have the narrow-scope structure, wide-scope advocates argue that they sometimes should be interpreted as having the wide-scope structure, according to which (1a) reads (roughly): “It must be the case that if today is Monday, then tomorrow is Tuesday”. Since the wide-scope interpretation does not allow for detachment, it blocks the inference from (1a) and (2a) to (3a), and thus avoids the modal fallacy.<sup>20</sup>

<sup>19</sup> See, e.g., Bradley and Swartz (1979, 330–332).

<sup>20</sup> The wide-scope interpretation is often regarded as uncontroversial, especially in the case of alethic modals. See Dreier (2009), however, for general scepticism about wide-scope interpretations with respect to both alethic and deontic modals.

This strategy has then been applied to normative conditionals to avoid the detachment of implausible normative conclusions. It has been applied to for example, the case of promises or hypothetical imperatives. Consider the following arguments:

- (1b) If A has promised to  $\phi$ , then A ought to  $\phi$ .  
 (2b) A has promised to do everything she can to extinguish mankind.  
 (3b) A ought to do everything she can to extinguish mankind.
- (1c) If A wants to  $\phi$ , then A ought to take the necessary means to  $\phi$ -ing.  
 (2c) A wants to extinguish mankind.  
 (3c) A ought to take the necessary means to extinguishing mankind.

Again, the premises (1b) and (1c) seem *prima facie* plausible, (2b) and (2c) can be stipulated, but the conclusions (3b) and (3c) are hard to accept. The wide-scope solution to this problem is to say that normative conditionals of the form

*Normative conditional:* If A  $\phi$ -s, then A ought to  $\psi$

are ambiguous between a narrow-scope and a wide-scope interpretation:

*Narrow-scope normative conditional:*  $A \phi\text{-s} \rightarrow O(A \psi\text{-s})$

*Wide-scope normative conditional:*  $O(A \phi\text{-s} \rightarrow A \psi\text{-s})$

And even though natural expressions of normative conditionals such as (1b) and (1c) seem to have narrow-scope structure on the surface, according to wide-scope advocates they should at least sometimes be understood as having wide-scope structure. Since the wide-scope formula does not allow for detachment, the conclusions (3b) and (3c) cannot be inferred.

As Mark Schroeder has pointed out, however, the application of the wide-scope strategy to normative conditionals presupposes that ‘oughts’ relate to propositions

rather than responses, for it presupposes that an ‘ought’ can govern a whole conditional, and a conditional is a proposition and not an action or some other response.<sup>21</sup> It thus seems that if we understand ‘oughts’ and rational requirements as relating agents to responses rather than operating over propositions, it is not open to us to interpret normative conditionals as taking wide scope.

Do we have to conclude that structural requirements of rationality cannot take wide scope, given that we want to understand them as response-requirements rather than proposition-requirements? We do not. It is true that the wide-scope account is often motivated in the way sketched above, namely as a logical interpretation of certain normative conditionals, which seem *prima facie* plausible but yield unacceptable conclusions when their apparent logical structure is taken at face value.<sup>22</sup> One has to agree with Schroeder that this way of arguing presumes that the object of an ‘ought’ can be a conditional and thus a proposition, and is therefore not available to someone who asserts that the object of an ‘ought’ can only be an agential response. But there are other reasons in favour of wide-scope interpretations of certain requirements and wide-scope advocates might well understand themselves as providing a genuine alternative to (rather than a logical interpretation of) claims of the form expressed in *normative conditional*.<sup>23</sup> One alternative is this:

*Normative disjunction*: A ought to [not  $\phi$  or  $\psi$ ]

According to the propositional account, *normative disjunction* is simply equivalent to *wide-scope normative conditional*. The reason is that the proposition ‘A does not  $\phi$  or A  $\psi$ -s’ is equivalent to the proposition ‘if A  $\phi$ -s, then A  $\psi$ -s’. According to the relational account, however, *wide-scope normative conditional* cannot be the logical form of any meaningful sentence, because it requires that ‘ought’ takes propositions. *Normative*

<sup>21</sup> See Schroeder (2004, 342–344) and (2011b, 35–36).

<sup>22</sup> For this kind of argument, see e.g. Dancy (1977).

<sup>23</sup> Schroeder (2004, 344) and (2011b, 36) concedes this possibility. Note that I did not introduce the wide-scope reading of structural requirements of rationality as an interpretation of certain conditionals at all; rather I introduced them as an alternative explanation for the idea that (A)-(D) are cases of irrationality and thus involve the violation of a rational standard.

*disjunction*, in contrast, can be understood as relating to a response directly, because we can interpret it as an ‘ought’ that relates to a disjunctive response (rather than a disjunctive proposition about responses, as the propositional account would have it). And that ‘oughts’ can relate to disjunctive responses seems to me uncontroversial. For example, consider Frank, who ought to take the bus or take the cab in order to be at the station in time. It is not true that Frank ought to take bus, nor that he ought to take the cab. Instead he ought to perform what we can call the ‘disjunctive action’ of taking-the-bus-or-the-cab.

### *Seeing to it that*

There is a further alternative that is sometimes used to express wide-scope normative conditionals:

*Normative STIT-statement*: A ought to see to it that [if A  $\phi$ -s, then A  $\psi$ -s].<sup>24</sup>

This *seeing-to-it-that* account of wide-scope requirements is also compatible with the view that ‘oughts’ relate to certain kinds of responses rather than propositions. The conditional proposition ‘if A  $\phi$ -s, then A  $\psi$ -s’ does not directly relate to the ‘ought’, but figures as part of the description of the response, namely the response of seeing to it that this proposition is true. But what does it mean to say that a person “sees to it that a proposition p is true”? As I understand it, it roughly means that the person acts or intends in ways that help to preserve that p (in case that p is the case), or bring it about that p (in case that p is not the case).

Two things should be kept in mind when *normative STIT-statement* is proposed as an alternative to *normative conditional* or *wide-scope normative conditional*. The first is that a person’s seeing to it that she  $\phi$ -s is not identical with this person’s  $\phi$ -ing. This is clearest in the case that ‘ $\phi$ -ing’ designates believing something or taking some other

---

<sup>24</sup> Broome (1999, 399) uses this expression “as a mere grammatical padding” that allows him to relate ‘oughts’ to propositions. He describes his doing so as an act of “desperation”. Since Broome intends this expression to be entirely content-free, he should not be understood as holding a seeing-to-it-that account of wide-scope requirements.

attitude. A person's believing  $p$  and a person's doing something that helps to bring it about or to preserve that she believes  $p$  are two very different matters of fact, and consequently a requirement to believe something is different from a requirement to see to it that one believes something. This means that *normative STIT-statement* is concerned with a genuine response-type that *normative conditional* (and all its interpretations or substitutions considered so far) does not refer to at all. *Normative conditional*, *wide-scope normative conditional*, and *normative disjunction* are concerned with  $\phi$ -ing and  $\psi$ -ing only, but *normative STIT-statement* brings into play a third kind of response, namely seeing to it that a certain proposition is true. Indeed, it is this kind of response which is fundamental according to the *normative STIT-statement*, since  $\phi$ -ing and  $\psi$ -ing are relevant only because they figure in the content of the proposition that specifies the required response. So we should keep in mind that if we substitute *normative conditional* with *normative STIT-statement*, we substantially change parts of the content that are preserved by *normative disjunction*.

The second, related point that we should keep an eye on is that a requirement to see to it that  $p$  is essentially *practical*. It calls for actions or intentions to bring about, or preserve, or make it the case that  $p$ , independently of whether the responses mentioned in  $p$  are themselves actions or intentions.

In some cases, *normative STIT-statement* might be an appropriate reinterpretation of a *prima facie* plausible normative conditional. Consider the paradox of the gentle murder.<sup>25</sup> "If you kill him, you ought to kill him gently" has *prima facie* plausibility, but allows for the detachment of an 'ought' that seems unacceptable. Given that 'oughts' do not take propositions, our alternatives are "You ought to not-kill-him-or-kill-him-gently" (the disjunctive account) or "You ought to see to it that if you kill him, you kill him gently" (the STIT-account). Though I do not want to insist on this, it seems to me that in this case the STIT-account more plausibly captures the intuitive appeal of the kind of advice expressed in the initial statement.

I believe, however, that in the case of structural requirements of rationality, the STIT-account is not really an option. To begin with, it seems to me clear that theo-

---

<sup>25</sup> See Forrester (1984).

retical requirements of rationality should be understood as calling for certain cognitive states, not as calling for actions or intentions to bring about or preserve these states. And if the theoretical requirements follow the disjunctive model, it would be surprising if the practical requirements followed the STIT-model. In any case, we would at least need an explanation for this asymmetry. More generally speaking, however, I believe that even in the practical cases, the point of structural rationality is not well captured by the idea of being practically effective in bringing about or preserving coherence. Rationality partly consists in a coherent mental life; *prima facie*, it requires us to believe and intend coherently, but in doing so is connected to these mental attitudes directly, not in virtue of a volitional act of attitude arrangement.

### *Summary*

Let me briefly sum up this discussion. Wide-scope ‘oughts’ (or, more generally, requirements) are often presented and motivated in a way that presupposes the questionable assumption that ‘oughts’ operate over propositions rather than relating directly to responses. On a response-relational account, wide-scope requirements can be understood either as requirements to show a disjunctive response, or as requirements to *see to it that* a conditional is true. As I have argued, the latter understanding is not suitable for rational requirements. The lesson to be drawn is that the wide-scope counterparts to narrow-scope structural requirements of rationality should be formulated as disjunctive requirements. The self-governance requirement, for example, will read:

(A)<sub>5</sub> *Self-governance requirement, disjunctive*: Rationality requires A to [intend to  $\phi$  or not believe that she ought to  $\phi$ ].

This formulation is neutral with respect to the question of whether ‘oughts’ and rational requirements take propositions – it allows us to understand them as relating agents to responses, which is, I think, the more plausible account. It also avoids the problems of the STIT-account. In addition, I think that the formulation provides a

more intuitive grasp of one of the major differences between wide- and narrow-scope requirements, namely that the former can be satisfied in more than one way.

In light of the discussion of this section one might raise the question of whether the narrow/wide-scope terminology is still appropriate. Perhaps we should talk of disjunctive and non-disjunctive requirements instead, as the scope terminology does indeed invite the questionable assumption that the relevant distinction is concerned with the logical interpretation of certain expressions that are subject to an ambiguity in scope. Since this terminology is very much established, however, I shall continue calling disjunctive requirements such as  $(A)_5$  ‘wide-scope’ and non-disjunctive requirements such as  $(A)_2$  ‘narrow-scope’ requirements. Nothing in the arguments to come hinges on this terminological question.

### 3.3 Conditional or unconditional requirements?

Before I get on to the possible diachronicity and processuality of structural requirements of rationality, I will take a closer look at the distinction between conditional and unconditional requirements. As I have introduced them, narrow-scope requirements such as  $(A)_2$ - $(D)_2$  are conditional on an antecedent state, while wide-scope requirements such as  $(A)_3$ - $(D)_3$ , and  $(A)_5$  as well, are unconditional. It has been argued that this counts against the wide-scope account. Mark Schroeder, for example, holds that rational requirements amount to certain kinds of ‘oughts’ or reasons, and he believes that ‘oughts’ and reasons have to be explained by, and are thus conditional on, certain mental states. This is part of the version of the Humean theory of reasons that Schroeder favours.<sup>26</sup> He puts his complaint as follows:

Like the other Wide-Scope accounts,  $[(A)_3]$  works by positing a basic, eternal, agent-neutral requirement rationally binding on every agent, no matter what they are like. As with all of the others, it offers no explanation of this requirement. Narrow-Scope accounts can explain the obligations or reasons that they

---

<sup>26</sup> He calls it “hypotheticalism”. See esp. Schroeder (2007a). For a brief discussion of how the question of scope relates to the Humean theory, see Schroeder (2007c, 209–214).

postulate. After all, these obligations or reasons only exist given a certain condition – so we can use that condition to explain them. But not so for the Wide-Scopers.<sup>27</sup>

So according to Schroeder, the wide-scope account is committed to the controversial view that there are unconditional agent-neutral requirements valid anytime for anyone, requirements that in contrast to their narrow-scope counterparts cannot be further explained. Part of his complaint seems to be guided by a request for a fundamental explanation of normativity in descriptive terms – a request whose legitimacy is open to dispute. But Schroeder points to an important feature of a wide-scope account along the lines of (A)<sub>3</sub> or (A)<sub>5</sub>, which is of interest independently of this request. Such a view really is committed to the claim that *all of us* are *constantly* subject to these requirements, an implication that the narrow-scope account avoids. Note the consequence that such requirements *always apply*; one cannot avoid or exit unconditional requirements – either one satisfies them or one violates them. This has implications that may well sound implausible to some. For recall that the statements of requirements discussed here merely provide a general formula for an infinite number of requirements, not only for every agent, but also for every possible action. As Errol Lord puts it, the wide-scope account seems to entail that

[...] I am always rationally required to [intend to hop up and down on my left foot, if I believe that there is conclusive reason to hop up and down on my left foot]. Moreover, I am *complying* with that requirement right now, since I do not believe I have decisive reason to hop up and down on my left foot. The narrow-scope requirements are not like this. For them to apply to you, you actually have to have specific attitudes. I would have to believe that I have conclusive reason to hop up and down on my left foot in order to be required by [(A)<sub>2</sub>] to do anything.<sup>28</sup>

---

<sup>27</sup> Schroeder (2004, 349).

<sup>28</sup> Lord (2011, 398). The second pair of square brackets is Lord's own.

Call this the *weird requirement problem*. It is not obvious how serious this problem is. Think of requirements on omissions, such as the moral requirement not to lie without good reason. We might think that it follows from this requirement that I am now required not to tell you that I am doing a headstand in Kuala Lumpur, and that, moreover, I am complying with this requirement right now, simply because I am not telling you that I am doing a headstand in Kuala Lumpur. But this would not lead us to give up the idea that we are required not to lie without good reason. So perhaps we simply need to accept that the approval of general requirements commits us to a lot of particular instances of this requirement that are by themselves not obviously plausible. Their lack of plausibility might be explained with reference to their redundancy rather than their incorrectness.

However, one could also reply that this objection rests on the questionable assumption that a general requirement to refrain from telling lies without good reason entails a variety of requirements, for each possible lie, not to tell this lie without good reason. Certainly, telling any lie without good reason constitutes the violation of the general requirement, but it is not obvious that the general requirement must be understood in such a way that it entails an infinite number of particular requirements to the effect that every case of not telling a particular lie counts as satisfying such a requirement.<sup>29</sup> In this respect, there seems to be a relevant disanalogy to structural requirements of rationality, which are commonly conceptualised as being concerned with local conflicts between particular response-types. It is a standard assumption of wide-scope advocates that giving up the antecedent state of a structural requirement of rationality constitutes a satisfaction of it, quite independently of whether one holds other particular attitude-states that instantiate the antecedent state as well. For example, giving up your belief that you ought to brush your shoes is a way of satisfying the wide-scope self-governance requirement even if at the same time you violate this requirement because you believe you ought – but you do not intend – to brush your teeth. This can be so only if there really is a valid requirement claim for every possible ought-belief. The

---

<sup>29</sup> Such an entailment would hold according to a version of *inheritance*, which states that necessary consequences of something that is required must themselves also be required. However, I have rejected a similar claim above and see no reason to accept this one.

weird requirement problem thus seems to be a particular problem for the wide-scope account of structural requirements of rationality; it is at least not obvious that it occurs in other cases of independently plausible requirements as well.

Perhaps the implausibility of such requirements can be explained in other ways, but I think it should be admitted that it is a disadvantage of an account that it entails that I am now required to [intend to hop up and down on my left foot, if I believe that I ought to hop up and down on my left foot]. It runs counter to intuition to say that all of us are constantly satisfying an infinite number of rational requirements, each for every possible wayward ‘ought’-belief that we do *not* have. However, as I will argue, this counts only against an *unconditional* wide-scope account as is present in (A)<sub>3</sub> or (A)<sub>5</sub>, not against the wide-scope account as such.

It is important to note that the question of when a requirement applies (the question of conditionality) and the question of what it requires (the question of scope) are logically independent. While it is true that *narrow-scope* structural requirements of rationality can be plausibly interpreted only as conditional requirements, there is no particular reason why the corresponding wide-scope requirements must be unconditional.<sup>30</sup> “Wide scope” does not have to mean “widest scope”: As we have seen in the case of (A)<sub>4</sub> already, there is also a conditional interpretation of such requirements. If we combine this with the result of the last section, namely that wide-scope requirements are disjunctive, we get the following version of the self-governance requirement:

(A)<sub>6</sub> *Self-governance requirement, conditional + disjunctive*: If A believes that she ought to  $\phi$ , then rationality requires A to [intend to  $\phi$  *or* not believe that she ought to  $\phi$ ].

---

<sup>30</sup> Cf. Bratman (2009a, 432) who wants to allow that the normative reason to satisfy wide-scope requirements of rationality is conditional on an agent’s actual concern about self-governance. Whether Bratman thinks the requirements themselves to be conditional in this way is nevertheless unclear. Note also that this kind of conditionality would not solve the weird requirement problem: It still seems implausible that all agents concerned about self-governance constantly satisfy infinitely many rational requirements, each for every possible ‘ought’-belief (or intention, or means/end-belief, etc.) that they do not have.

(A)<sub>6</sub> does not entail that I am now satisfying an infinite number of requirements for every wayward ought-belief that I do not have, which I think is the right result. It also allows us to say that we *avoid* rather than satisfy rational requirements in certain circumstances where this seems the plausible thing to say (e.g. when we fall asleep). Finally, (A)<sub>6</sub> is not committed to the existence of unconditional, agent-neutral requirements. So if requirements are to be explained by mental states, then the wide-scope requirement (A)<sub>6</sub> is just as easy to explain as the narrow-scope requirement (A)<sub>2</sub>.

Conditional wide-scope requirements do not allow for detachment, just as unconditional ones do not: From the fact that A believes that she ought to  $\phi$ , we cannot conclude that A is required to intend to  $\phi$ , only that she is required to intend-to- $\phi$ -or-not-believe-she-ought-to- $\phi$ . The step from unconditional to conditional requirements thus preserves the most important difference between narrow- and wide-scope requirements. There is, however, a difference that it does not preserve: conditional wide-scope requirements share the satisfiability conditions with their narrow-scope counterparts rather than with their unconditional wide-scope counterparts. If you do not believe that you ought to  $\phi$ , then the requirement does not apply to you. The plausible consequence of this is that we avoid the weird requirement problem. The less plausible consequence is that even though the requirement calls for the disjunctive response of intending-to- $\phi$ -or-not-believing-one-ought-to- $\phi$ , you can satisfy the requirement only by intending to  $\phi$ . If you respond by dropping your belief, you *escape* the requirement rather than satisfy it. This point generalises: Conditional versions of wide-scope structural requirements of rationality cannot be satisfied by dropping the antecedent state, only by adopting (or remaining in) the consequent state. In this respect, they are like their narrow-scope counterparts and unlike their unconditional wide-scope counterparts.

On the one hand, it sounds odd to say that the requirement is disjunctive, but can only be satisfied by giving one of the responses that are part of the disjunctive response. On the other hand, it does not follow that one violates the requirement unless one adopts the consequent state, as it is still possible to *escape* the requirement by dropping the antecedent state. Moreover, the disjunctive structure still blocks detach-

ment, so there is no danger that such requirements collapse into narrow-scope requirements. Taking all of this into account, it seems to me that the weird requirement problem still provides sufficient reasons for wide-scope advocates to conditionalise their requirements.

Having said this, let me now point to a possible way for wide-scope advocates to avoid both the ‘weird requirement problem’ *and* the odd consequence that the disjunctive requirement can only be satisfied in one way. They can do so by giving their requirements a *diachronic* rather than *synchronic* interpretation. Consider the following general pattern for diachronic conditional wide-scope requirements:

*Pattern for diachronic wide-scope requirements:* If A is in the antecedent state at t, and if there is a later time t' such that t' is appropriately related to t, then rationality requires A to [be in the consequent state at t' *or* not-be in the antecedent state at t'].

Evidently, this pattern leaves open a number of questions, particularly the question of what it means to say that t' is “appropriately related” to t. I shall say more about this in the next section. For now, it is important to note that wide-scope requirements of this form *can* be satisfied by dropping the antecedent state, even though they are conditional in a way that avoids the weird requirement problem. The reason is that, by dropping the antecedent state at t', one does not escape a requirement that is conditional on one's being in the antecedent state at t.

In the next section, I shall give an independent argument to the effect that structural requirements of rationality should be understood diachronically. If this argument is sound, we can conclude that the most plausible versions of wide-scope requirements would be conditional *and* diachronic, avoiding both the weird requirement problem and the odd consequence that they can be satisfied only in one way.

### 3.4 Synchronic or diachronic requirements?

At first appearance, it may seem that structural requirements of rationality are most naturally understood as being synchronic. The starting point of our discussion was,

after all, that we can detect that someone is irrational just by looking at a certain combination of his attitudes at a certain time. I shall argue, however, that this way of describing the situation is not quite accurate.

We have to take into account the fact that rational adjustment of one's attitudes often takes time. Perhaps this is not the case for all kinds of rational adjustments; some possibly work instantaneously. Suppose, for example, that rationality requires you not to believe  $p$  and believe  $\neg p$ . Suppose you believe  $\neg p$ , and then on the basis of a perception of  $p$ , you come to believe  $p$ . Plausibly, if you are rational, then in the moment of your acquiring this new belief in  $p$ , you drop your belief in  $\neg p$  instantaneously.<sup>31</sup>

### *The problem*

But things are more complicated when we consider cases in which the contents of one's attitudes are not as closely related as  $p$  and  $\neg p$ , and if one's attitudes are involved in different kinds of rational conflicts at the same time. Consider the case of Michael. After thinking about some question, weighing the evidence for and against a possible answer to it, and discussing it with friends, Michael ends up acquiring a new belief in  $p$ , and he holds this belief for good reasons. Michael also believed for a long time that if  $p$ , then he has sufficient evidence to believe that  $q$ ; that if  $q$ , then he ought to  $\phi$ ; and also that  $\psi$ -ing is a necessary means for  $\phi$ -ing. According to a synchronic narrow-scope account, Michael becomes irrational in the very moment that he comes to believe  $p$ , unless he does not *at the same time* also adopt the following other attitudes: a belief that he has sufficient evidence to believe that  $q$ ; a belief that  $q$ ; a belief that he ought to  $\phi$ ; an intention to  $\phi$ ; and an intention to  $\psi$ . According to a synchronic wide-scope account, Michael becomes irrational in the very moment he comes to believe  $p$ , unless he does not *at the same time* adopt the aforementioned attitudes *or* give up at least one of his other attitudes. In each case, however, this verdict seems too harsh. We can suppose that Michael is a very good and quick thinker, that he runs through the process of reconsidering and adjusting his attitudes in the fastest rational way available – but still

---

<sup>31</sup> Broome made this point in chapter 6 of the manuscript of his book *Rationality Through Reasoning* (2013), but apparently it was omitted in the published version of this chapter (which is chapter 9 in the book).

this process will at least take *some* time. Synchronic requirements have the implausible implication that Michael is irrational during this process. And Michael's case seems likely to be the rule rather than the exception; rarely will it be the case that we are psychologically capable to rationally adjust all of our attitudes in the very moment in which we acquire a new attitude. So if requirements of structural rationality are synchronic, we have to conclude that we are very often irrational when acquiring a new attitude.

Broome, who thinks that almost all rational requirements are synchronic, does not seem to worry about this conclusion: "I know I fail to be rational in many respects anyway, and I am happy to recognise this temporary respect."<sup>32</sup> But Broome also acknowledges that there is an important connection between irrationality ascriptions and criticism: "When we accuse someone of irrationality, we are surely criticizing her."<sup>33</sup> Michael, however, does not seem to be criticisable at all. So we should not be so happy to accept this extra respect of irrationality.

In response, it might be argued that insufficient time for attitude revision provides an excuse for irrationality, so that Michael's case would be an example of uncriticisable irrationality. Whether or not there is such a thing as uncriticisable irrationality, it seems that the conclusion that we are constantly irrational when we acquire new attitudes is in itself implausible, even if we abstract from the connection between irrationality and criticisability. There is simply nothing irrational in acquiring a new attitude for good reasons and adjusting other attitudes in the fastest rational way.

In order to avoid this problem, we could restrict the content of rational requirements to those cases in which the agent does not currently revise her attitudes. While this would avoid the undesirable conclusion that Michael is irrational during the process of revision, it would also yield the conclusion that Michael is not under the relevant requirement at all during this process. But this involves a distorted picture of what is going on in his case. According to this approach, Michael violates a rational requirement in the moment he acquires his new belief, he then exits this requirement by starting to reconsider other conflicting beliefs, and finally he satisfies the require-

---

<sup>32</sup> Broome (2013, 153).

<sup>33</sup> Broome (2005a, 336) and (2007a, 177).

ment when he finishes his adjustments. It seems, however, that in revising his attitudes, Michael responds to a rational requirement that applies to him, not to one that only applied to him before and will again once he has finished his revisions. In addition, it seems implausible that we can avoid the kinds of irrationality in question simply by continually revising our attitudes.

Note that there is a similar problem with reasons. Suppose that Marina perceives  $p$ , and suppose that  $p$  is conclusive evidence that  $q$ . It seems, then, that in the moment of her perception, Marina acquires a decisive reason to believe that  $q$ , and thus ought to believe that  $q$ . But even if Marina has very well-functioning rational capacities, she will need some time at least to realise the relevance of  $p$  for  $q$ , and thus to conclude that  $q$ . We seem to be forced to conclude that in the meantime, Marina violates a normative requirement, for she does not believe what she has decisive reason to believe. But since 'ought' implies 'can', this is an implausible thing to say. And again, it would be strange to hold that during the time of adjustment, Marina does not really have this reason and exits the normative requirement to believe that  $q$ .

### *Diachronic requirements*

So what should we say about these cases? The right thing to say about Marina's case seems to be this: In the moment of her perception, at  $t$ , Marina acquires a reason to believe that  $q$ . But this reason might not be a reason to believe *at  $t$*  that  $q$ ; it could be a reason to believe at some later time  $t'$  that  $q$ , or a reason to come to believe  $q$  within a period of time starting at  $t$ . The reason, as I said, is present from  $t$  on. So while Marina adjusts her attitudes she can count as responding to a normative requirement that applies to her without violating this requirement. Similarly, I believe that in Michael's case we should say that the rational requirement holds from the very moment in which he acquires his new belief, but the coherence with his other attitudes is required for a later time. This allows him to rationally adjust his attitudes without violating this requirement.

The upshot is that structural requirements of rationality must be diachronic after all. If we had the impression that we could tell someone is irrational just by

looking at his attitudes at a singular moment, we were ignoring that this person might have just acquired new attitudes and might not have had sufficient time for revision. This can happen without any failure of rationality on the agent's part.<sup>34</sup>

But how should diachronic requirements of rationality be formulated? I have already suggested above a general formula for diachronic wide-scope requirements. Let me here use a version that is neutral concerning wide- and narrow-scope interpretations:

*General formula for diachronic requirements:* If A is in the antecedent state at  $t$ , and if there is a later time  $t'$  such that  $t'$  is appropriately related to  $t$ , then rationality requires A to be in the consequent state at  $t'$  (resp. be in the consequent state at  $t'$  or not be in the antecedent state at  $t'$ ).

Let me elaborate on this formula a bit further, and consider how we should move on from it to more specific formulations of requirements. To begin with, recall the earlier remark that some rational adjustments might take place instantaneously. This poses the question of whether rationality does not sometimes require instantaneous adjustment. We can incorporate this by defining  $t'$  as a point of time that is *possibly* later, but not earlier than  $t$ , thereby allowing  $t$  and  $t'$  to be the same point of time.

If  $t$  and  $t'$  are the same point of time, then we again encounter the odd idea of a disjunctive requirement that can only be satisfied by adopting the consequent state, because if A is not in the antecedent state at  $t$ , the requirement does not apply. This follows from my assumptions. In order for it to be the case that  $t$  and  $t'$  are the same point in time, we have to assume that rationality requires instantaneous adjustment at  $t$ . In order for it to be the case that the requirement applies, we have to assume that A

---

<sup>34</sup> It follows that not only the rational requirements discussed so far, but also the structural irrationality claims (A)-(D) are strictly speaking inaccurate and have to be substituted by diachronic versions. For instance, (A) has to be substituted by a claim such as this: If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , and if A believes at  $t'$  that she ought to  $\phi$ , and A does not intend at  $t'$  to  $\phi$ , then A is irrational at  $t'$ .

is in the antecedent state at  $t$ .<sup>35</sup> Since  $A$  cannot both be and not-be in the antecedent state at  $t$ , it follows that the only way to satisfy the requirement (be it disjunctive or not) is to adopt the consequent state at  $t$  as well. So either we have to accept the odd consequence, or we have to deny that disjunctive requirements of rationality can call for instantaneous adjustment. I will leave this question open.

Some authors include the fact that  $A$  is not in the consequent state at  $t$  as a condition of a diachronic requirement.<sup>36</sup> However, this conception entails that structural requirements of rationality only apply to us if we have incoherent attitude-states in the first place. As a result, it would be impossible for us to satisfy any such requirement at  $t'$ , unless we displayed an incoherent mind-set at  $t$ . This strikes me as an implausible and unnecessarily strong constraint.

A further question is whether we should include the fact that  $A$  *remains* in the antecedent state until  $t'$  as a condition of a diachronic requirement. Note that in the wide-scope case, we would again end up with the odd result of disjunctive requirements that can be satisfied only in one way. This is because dropping the antecedent state at  $t'$  would effect that the requirement – which would then be conditional on  $A$ 's being in the antecedent state at  $t'$  – no longer applies.<sup>37</sup> But there is also a more general consideration that counts against including any such condition, and this consideration is independent of the scope of the requirements in question.

We turned to diachronic requirements in order to account for processes of rational adjustment and the fact that such processes take time. It seems clear that it should at least be *possible* that the processes in question could themselves be guided by the very diachronic requirements that demand the states that make the rational adjustment necessary. In other words, it should be possible that agents adjust their

---

<sup>35</sup> We have to do so if we follow the argument of the last section that we should understand both wide- and narrow-scope structural requirements of rationality conditionally. I will briefly consider the possibility of unconditional diachronic requirements below.

<sup>36</sup> See the diachronic requirements that Reisner (2009, 248–249) discusses, and the process-requirements Kolodny (2007b, 373) suggests. Note that process-requirements that call for the *formation* of the consequent state plausibly need to presuppose the absence of that consequent state (and thus include the condition at issue). This is a good reason to doubt Kolodny's thesis that process requirements exhaust the claims of rationality (more on this below).

<sup>37</sup> Perhaps the result could be avoided if we included only the condition that  $A$  remains in the antecedent state until some point *right before*  $t'$ . However, I think this condition is vulnerable to the same objection that I go on to discuss in the main text.

attitudes *in response* to such requirements. But in order for this to be possible, the requirements have to apply at the time at which the processes of rational adjustment are supposed to *begin*, not only at the time at which they are supposed to be finished, or nearly finished. However, if we include as a condition of the requirement that A remains in the antecedent attitude until some time after  $t$ , then the requirement does not apply until this later time, and consequently, it cannot possibly guide the rational processes it was supposed to make room for in the first place.<sup>38</sup> So I shall not include any such condition.<sup>39</sup>

Finally, we have to consider the relation between  $t$  and  $t'$ . It seems to be clear that the exact period between  $t$  and  $t'$  (or indeed whether  $t$  is different from  $t'$  at all), partly depends on contingent features of the particular situation. It depends, for example, on the number of rational conflicts in which a person's antecedent attitude is involved, and this in turn depends on which other attitudes she holds at  $t$ . It is also possible that it depends on A's particular rational capacities; it could be that the better they function, the closer we have to conceive of  $t$  and  $t'$ . I shall not discuss these questions here. My aim is just to provide a general outline of diachronic requirements.

I do, however, need to say something on a general level about what I have called the "appropriate relation" between  $t$  and  $t'$ . I used this expression as a placeholder, because some qualification of  $t'$  beyond its being not earlier than  $t$  is certainly necessary: Agents are certainly not required to give the responses in question at *every* time that is possibly later than  $t$ . I believe it is an advantage of the formula that it is

---

<sup>38</sup> It might be argued that requirements following the above formula are vulnerable to the same objection, simply because they are conditional on there being a time  $t'$  such that  $t'$  is appropriately related to  $t$ . However, I am using the expression "there is a time  $t$ " in a way that is independent of whether  $t$  is in the past, present or future. The mentioned condition can therefore already be fulfilled at  $t$ , even if  $t$  is before  $t'$ .

<sup>39</sup> This opens up the possibility of an objection to the narrow-scope account. Diachronic narrow-scope accounts entail that if the above-mentioned conditions are satisfied, and the agent is not in the consequent state at  $t'$ , then the agent is irrational. But suppose that the agent drops the antecedent attitude between  $t$  and  $t'$ . Then, at  $t'$  this agent might not be in the consequent state and still be entirely rational. This is at least what Broome (2007c, 366–368) argues. In contrast, Kolodny (2007b, 381–382) maintains that given  $t$  and  $t'$  are sufficiently close to one another, such an agent could not be entirely rational. I shall not discuss this dispute here. I have argued on independent grounds that structural requirements of rationality are diachronic, and that they are not conditional on the agent's remaining in the antecedent state until  $t'$ . If it turns out that there is no plausible narrow-scope requirement that meets these criteria, then all the worse for the narrow-scope account. As I will argue in Chapter 6, there are sufficient reasons to reject the narrow-scope account quite independently of the diachronicity of structural requirements of rationality.

neutral with regards to different controversial conceptions of the appropriate relation between  $t$  and  $t'$ , and I will go on to use that expression as a placeholder for whatever this relation is. In what follows, I shall only suggest one possible conception of this relation.

One thing we might want to say is that it must be psychologically possible for  $A$  to rationally adjust her relevant attitude-states between  $t$  and  $t'$ . By the “relevant” attitude-states, I mean the attitude-states whose rational status is (directly or due to their relation to other attitude-states) affected by the antecedent attitude. It could be objected that we would thereby exclude the possibility that it is psychologically impossible for a person to be rational. I am willing to accept that conclusion, however. I agree that we are sometimes inclined to call a person irrational even if we regard it as an open question as to whether this person is psychologically capable of avoiding this kind of irrationality. Yet on reflection it seems to me more plausible to say that if a person really is incapable of conforming to a putative requirement of rationality, then the requirement does not apply to her. This is a view that gains support from the connection between irrationality and criticisability. But we can also characterise the relation between  $t$  and  $t'$  in a more neutral way. Instead of assuming that it is psychologically possible for  $A$  to rationally adjust her attitudes between  $t$  and  $t'$ , we could say that it is either psychologically possible for  $A$  or would be psychologically possible for a human being with standard capacities in  $A$ 's circumstances to rationally adjust her attitudes between  $t$  and  $t'$ .

More is needed in order to spell out the appropriate relation between  $t$  and  $t'$ , since there are too many points in time at which this condition can be fulfilled. Even if it is psychologically possible to rationally adjust one's attitudes within 50 years, pensioners are normally not rationally required to intend something only because they believed as teenagers that they ought to do it.

This problem could be solved if we said that  $t'$  must be the point in time closest to  $t$  for which it is true that it is psychologically possible for  $A$  (or  $A$ 's counterpart) to rationally adjust her relevant attitude-states between  $t$  and  $t'$ . But as a result of this view, there would only be one single point in time that could meet this criterion. This raises the question what happens at points in times right after this single point. Is  $A$  no

longer required to give the response at  $t'+x$ , just because  $t'+x$  is slightly later than the *first* possible time at which she could have adjusted her attitudes? This would of course be absurd. Fortunately, it also does not follow from the above amendment. A is required to give the response at  $t'+x$  only if there is a time  $t+x$  at which A is in the antecedent state, and  $t'+x$  is the time closest to  $t+x$  for which it is true that it is psychologically possible for A (or A's counterpart) to rationally adjust her attitudes between  $t+x$  and  $t'+x$ . Unless A abandons the antecedent attitude immediately after  $t$ , this condition will be fulfilled. And *if* A abandons the antecedent attitude immediately after  $t$ , it seems acceptable to say that she is no longer required to give the response at  $t'+x$ .

Perhaps, then, the current characterisation is viable. Accordingly, the relation between  $t$  and  $t'$  is appropriate if, and only if,  $t'$  is the point of time closest to  $t$  for which it is true that it is psychologically possible for A (or a human being with standard capacities in A's circumstances) to rationally adjust her relevant attitude-states between  $t$  and  $t'$ . This seems to me a viable option, but I shall not presuppose this conception in what follows. For all I know there might be other, perhaps more satisfying, characterisations of the appropriate relation.

#### *Conditional and unconditional diachronic requirements*

For the purpose of illustration, let us consider what a diachronic version of the self-governance requirement looks like, in both the narrow- and the wide-scope interpretations:

- (A)<sub>7</sub> *Diachronic self-governance requirement, narrow-scope*: If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , then rationality requires A to intend at  $t'$  to  $\phi$ .
- (A)<sub>8</sub> *Diachronic self-governance requirement, disjunctive*: If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , then rationality requires A to [intend at  $t'$  to  $\phi$  *or* not believe at  $t'$  that she ought to  $\phi$ ].

My interpretation of the diachronic disjunctive (or wide-scope) requirement proceeded on the assumption argued for in the last section, namely that wide-scope requirements are, just like narrow-scope requirements, best understood conditionally. Despite this it is worth taking a look at an unconditional version of the diachronic wide-scope requirement. John Broome formulates the diachronic wide-scope version of the self-governance requirement as follows:

Rationality requires of you that, if at one time you believe you ought to F, there is a later time at which you intend to F.<sup>40</sup>

This requirement is unconditional, it applies to you independently of your states. I think that it is not a charitable reading of a diachronic wide-scope requirement, since it does not capture the spirit of the wide-scope account. As Broome himself notes, this requirement plausibly allows for detachment. Let  $t$  be the time at which you believe you ought to F, and let  $t'$  be the later time. Then, it is temporally necessary at  $t'$  that you believed at  $t$  that you ought to F; the only possible way for you not to violate this requirement is to intend to F. In this case it is very plausible to assume that we can derive a rational requirement for you to intend to F.<sup>41</sup> As a result, Broome's diachronic wide-scope requirement *entails* the diachronic narrow-scope requirement that "If at one time you believe you ought to F, there is a later time at which rationality requires you to intend to F". Consequently, Broome rejects this requirement. He seems to conclude that there is no diachronic requirement of self-governance, and that there are no diachronic versions of any of the other requirements at issue here.

Broome simply ignores the possibility that the requirement in question could be a conditional requirement like (A)<sub>8</sub>.<sup>42</sup> In fact, Broome is committed to this diachronic

---

<sup>40</sup> Broome (2013, 144).

<sup>41</sup> Broome (2013, 123) calls the detachment rule that validates this derivation *necessary detachment*: " $((\text{Necessarily } p) \ \& \ (S \text{ requires of } N \text{ that } p \rightarrow q)) \rightarrow (S \text{ requires of } N \text{ that } q)$ ". Or, as I would prefer to put it: If A is required to  $[\phi \text{ or } \psi]$ , and it is necessary that A does not  $\phi$ , then A is required to  $\psi$ . Reisner (2009, 250) accepts this detachment rule as well, and draws the same conclusion.

<sup>42</sup> He also ignores the possibility of an unconditional requirement of the following form UC-(A)<sub>8</sub>: "Rationality requires A to [not believe at  $t$  that she ought to  $\phi$  or [not believe at  $t'$  that she ought to  $\phi$  or intend at  $t'$  to  $\phi$ ]]" (cf. the requirement E5 in Reisner 2009, 249, which is roughly equivalent). Note, however, that for the

requirement, because (time-independent details of formulation aside) his favoured unconditional synchronic version of the self-governance requirement entails it. Recall:

(A)<sub>5</sub> Rationality requires A to [intend to  $\phi$  or not believe that she ought to  $\phi$ ].

Understanding this requirement synchronically means to understand it as the claim that for every time  $t$ , rationality requires A to [intend at  $t$  to  $\phi$  or not believe at  $t$  that she ought to  $\phi$ ]. This claim implies that the same requirement applies at all times that are appropriately related to a (possibly) earlier time at which A believes that she ought to  $\phi$ , which is the content of (A)<sub>8</sub>. Hence, (A)<sub>5</sub> implies (A)<sub>8</sub>. Consequently, the only question is whether wide-scope advocates should hold, in addition to (A)<sub>8</sub>, the stronger (A)<sub>5</sub>. Michael's case is a counterexample to (A)<sub>5</sub>, but not to (A)<sub>8</sub>, so I think that wide-scope advocates are well advised to maintain only the weaker claim.

### *Conclusion*

As we have seen, diachronic requirements pose an abundance of questions concerning matters of detail, which go far beyond the scope of this chapter. Though much of the details of a diachronic account of rationality remain to be filled in, I have argued that we have good reasons to consider the rational requirements at issue here as being diachronic rather than synchronic. For the sake of simplicity, I will nevertheless sometimes use the simpler time-insensitive versions in what follows. For many of the questions that I shall go on to discuss the difference does not matter, and introducing time indices and further conditions to all requirements complicates some questions in an unnecessary way. I shall pay attention to time whenever it is relevant for the arguments at issue rather than consistently using diachronic requirements.

---

same reasons that Broome's diachronic requirement entails the corresponding narrow-scope requirement, UC-(A)<sub>8</sub> implies (A)<sub>8</sub>. Once  $t$  has passed, we can treat it as necessary and derive (A)<sub>8</sub> through necessary detachment. Reisner (2009, 249) expresses doubts that the diachronic self-governance requirement UC-(A)<sub>8</sub> "differs in any meaningful way" from the synchronic version (A)<sub>5</sub>, and his reasoning indicates that he would say the same about (A)<sub>8</sub>. Since (A)<sub>8</sub> is implied by (A)<sub>5</sub> (as I go on to argue in the main text), it is clear that it adds nothing to it, but both (A)<sub>8</sub> and UC-(A)<sub>8</sub> differ in a meaningful way from (A)<sub>5</sub> in requiring considerably less, and thus avoiding counterintuitive charges of irrationality. Michael's case provides a counterexample to (A)<sub>5</sub>, but not to (A)<sub>8</sub> nor to UC-(A)<sub>8</sub>.

### 3.5 State- or process-requirements?

Finally, I will discuss the question of whether the requirements of rationality that correspond to structural irrationalities like (A)-(D) are state- or process-requirements. As I understand it, this distinction concerned with the kind of response that a requirement relates to: State-requirements require agents to be in certain states (e.g. the state of believing something), while process-requirements require agents to run through certain processes (e.g. the process of reasoning in a certain way). I shall treat all kinds of agential responses that are temporally extended as possible relatives of process-requirements. This will include *forming* or *dropping* certain attitudes, but also *retaining* attitudes or *remaining* in a state.<sup>43</sup> Since processes are temporally extended unities, all process-requirements are also diachronic: In referring to processes, they refer (at least implicitly) to at least two points in time. Yet not all diachronic requirements are process-requirements. The diachronic requirements discussed above did not include mention of any temporally extended unity; they referred only to the occurrence of certain states at different points in time.<sup>44</sup>

Niko Kolodny argues that all rational requirements, including the requirements corresponding to structural irrationality claims like (A)-(D), are process- rather than state-requirements. As far as I can see, he gives three arguments for this view. I shall first examine his most recent argument, according to which “the distinction between

---

<sup>43</sup> I follow Kolodny (2005, 517) here, who describes requirements to retain attitudes as process-requirements, but not Fink (2012, 118), who treats *change* as a necessary aspect of a process. Fink argues that “by thinking of any requirement to retain an attitude as a process-requirement, we would undermine any meaningful distinction between *state-* and *process-*requirements” (119). But terminological questions aside, there is an interesting distinction between requirements to be in a certain state at a certain time and requirements to run through temporally extended processes, even though the latter may include requirements to be in a certain state over a period of time. Put differently, there is a difference between requirements to believe or intend, and requirements to *form*, *drop*, or *retain* certain beliefs and intentions in certain ways. This is the distinction that I, following Kolodny, refer to with the terms ‘state-’ and ‘process-requirements’. In any case I do not think that any of my arguments depends on this terminological question.

<sup>44</sup> This claim might be challenged in the following way. The state-requirements discussed above call for states such as beliefs or intentions as responses, i.e. as *believings* or *intendings*. Such responses are events, and events are necessarily temporally extended. These assumptions are not uncontroversial, but let us grant them for the sake of argument. It follows that there is a sense in which all rational requirements call for temporally extended responses and thus for processes. It does not follow that the requirements in question really are requirements to *form*, *drop*, or *retain* beliefs or intentions in certain ways, rather than requirements to believe or intend. It is this latter claim, however, that Kolodny defends (and needs for his arguments against the wide-scope account), and which I go on to reject.

wide and narrow scope has little interest” unless requirements that govern structural irrationalities like (A)-(D) are process-requirements.<sup>45</sup>

*First argument: only process-requirements give point to the debate over scope*

Kolodny here seizes on a point made by John Broome, who shows that the narrow- and wide-scope versions of synchronic rational requirements do not differ in their implications for the question of whether a person is rational or irrational at any point in time.<sup>46</sup> He provides a helpful illustration of Broome’s formal proof. Consider again the narrow- and the wide-scope versions of the synchronic self-governance requirement:

(A)<sub>2</sub> *Narrow*: If A believes that she ought to  $\phi$ , then rationality requires A to intend to  $\phi$ .

(A)<sub>5</sub> *Wide*: Rationality requires A to [intend to  $\phi$  *or* not believe that she ought to  $\phi$ ].

Next, consider a certain point in time  $t$ . Either you believe at  $t$  that you ought to  $\phi$ , or you do not; and either you intend at  $t$  to  $\phi$ , or you do not. That is, necessarily, at  $t$  you are in one of the following states:

- (i) You believe that you ought to  $\phi$  and you do not intend to  $\phi$ .
- (ii) You believe that you ought to  $\phi$  and you intend to  $\phi$ .
- (iii) You do not believe that you ought to  $\phi$ .

If you are in state (i), you violate both (A)<sub>2</sub> and (A)<sub>5</sub>. Hence both requirements imply that you are irrational at  $t$ . If you are in state (ii), you violate neither (A)<sub>2</sub> nor (A)<sub>5</sub>. Hence neither requirement implies that you are irrational at  $t$ . If you are in state (iii), you also violate neither (A)<sub>2</sub> nor (A)<sub>5</sub>, and again neither requirement implies that you are irrational at  $t$ . Hence, (A)<sub>2</sub> and (A)<sub>5</sub> have the very same implications for the question of whether you are rational or irrational at a certain point in time.

---

<sup>45</sup> Kolodny (2007b, 373).

<sup>46</sup> See Broome (2007c, 361–364 and 369–370).

There is but one difference. While in (i) you violate both requirements, and in (ii) your non-violation constitutes a satisfaction of both requirements, in (iii) your non-violation constitutes a satisfaction only of the wide-scope requirement. Since the narrow-scope requirement applies to you only if you believe that you ought to  $\phi$ , you do not satisfy this requirement in (iii); instead you avoid it. This, however, does not make a difference concerning the question whether you are irrational, for avoiding a rational requirement is not irrational, only violating it is.

The conclusion that there is no difference in violation between narrow- and wide-scope requirements is indeed surprising. It seems that one of the main motivations for the wide-scope account could be expressed as follows: “It is irrational not to intend what one believes one ought to do, or not to intend to take the necessary means to our ends – that much is clear. But it is not true that we can become rational only in one way, namely by adopting the intention in question. Instead, we might just as well revise our belief or abandon our end!” If this thought is supposed to capture an important idea of the difference between wide and narrow scope, however, we do indeed have good reason to assume that the disagreement was not about synchronic requirements in the first place. The reason is that, as we have seen, wide- and narrow-scope versions of such requirements do not differ in violation. It follows that neither of them entails that you are irrational when you revise your belief or give up your end! The only difference is that in doing so, you *satisfy* the wide-scope requirement, while you *avoid* or *exit*, rather than satisfy, the narrow-scope requirement. But this difference does not effect the question of whether you are rational or irrational at some time.

Kolodny’s conclusion that if this were true, the difference between wide and narrow scope would be “almost negligible”<sup>47</sup>, strikes me as overhasty, for an important difference remains: While wide-scope versions of synchronic requirements do not generally allow for detachment, narrow-scope versions do. And this difference is important regarding the major argument against the normativity of rationality, namely the ‘bootstrapping’ objection, which I shall discuss below. I nevertheless agree with Kolodny that it is an awkward feature of synchronic requirements that the difference

---

<sup>47</sup> Kolodny (2007b, 375).

between wide and narrow scope comes down to the rather technical question whether a requirement is satisfied or merely not violated.

According to Kolodny, this indicates that wide- and narrow-scope advocates actually disagree about the scope of *process*- rather than *state*-requirements, for wide- and narrow-scope versions of *process*-requirements do differ in violation. But this conclusion is based on a misguided contrast. The problem is not that the requirements are concerned with states rather than processes, but that they are synchronic rather than diachronic. For the above-mentioned diachronic *state*-requirements also differ in violation, and can thus fully account for the intuition that there is an important difference between wide- and narrow-scope concerning the question of whether we are rational or not in certain cases. For purposes of illustration, consider again the two versions of the diachronic self-governance requirement:

- (A)<sub>7</sub> *Narrow*. If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , then rationality requires A to intend at  $t'$  to  $\phi$ .
- (A)<sub>8</sub> *Wide*. If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , then rationality requires A to [intend at  $t'$  to  $\phi$  *or* not believe at  $t'$  that she ought to  $\phi$ ].

Both requirements apply to you only if you believe at  $t$  that you ought to  $\phi$ , so let us suppose this is the case. The question is what happens at  $t'$ . Necessarily, at  $t'$  you are in one of the following states:

- (i) You believe that you ought to  $\phi$  and you do not intend to  $\phi$ .
- (ii) You believe that you ought to  $\phi$  and you intend to  $\phi$ .
- (iii) You do not believe that you ought to  $\phi$  and you intend to  $\phi$ .
- (iv) You do not believe that you ought to  $\phi$  and you do not intend to  $\phi$ .

If you are in state (i), you violate both (A)<sub>7</sub> and (A)<sub>8</sub>. If you are in state (ii), you satisfy both (A)<sub>7</sub> and (A)<sub>8</sub>, as you do when in state (iii). But now consider state (iv). Since you do not intend to  $\phi$ , you violate (A)<sub>7</sub>. But you do not violate (A)<sub>8</sub>, indeed you satisfy it, for one way to satisfy (A)<sub>8</sub> is to give up your belief that you ought to  $\phi$ . So here we have a real difference in violation.

We can draw two conclusions. First, Kolodny's considerations about the significance of the contrast between wide and narrow scope lend further and independent support to my earlier conclusion that rational requirements are best to be understood diachronically rather than synchronically. Second, these considerations do nothing to support the conclusion that rationality requires processes rather than states.

*Second argument: only process-requirements are response-guiding*

Let us turn to Kolodny's second argument against state-requirements. According to this argument, rational requirements are process-requirements because only process-requirements can be normative in a way that makes them capable of guiding our responses, while state-requirements can at most be evaluative standards of appraisal. This argument is somewhat surprising given that Kolodny's main purpose is to show that rationality is not, after all, normative. We may understand him as claiming that rational requirements *seem* to be normative, and only process-requirements exhibit this *apparent* normativity. In his own words:

[...] the very idea of a state-requirement is questionable. If rational requirements are normative, deontic, or response guiding, then they call for the subject to *respond* in a certain way.<sup>48</sup> [...] they can function as advice or guide your deliberation. Process requirements can be normative in this sense, since they tell you to do something. But state requirements cannot be normative in this sense, since they do not tell you to do anything. At most, state requirements

---

<sup>48</sup> Kolodny (2005, 517).

might be evaluative requirements: that is, necessary conditions for qualifying for a certain kind of appraisal.<sup>49</sup>

It is difficult to evaluate this argument. As we have seen, we have good reasons to understand rational requirements as relating agents to responses rather than to propositions. In a very broad sense of ‘doing’, then, rationality requires us to *do* certain things, as Kolodny insists. But this does not seem to support the conclusion that rationality requires processes rather than states. Beliefs and intentions, e.g., are states, but believing and intending are also things that we can *do* in a broad sense of ‘doing’, and they can naturally be understood as responses. Kolodny does not seem to have any problem with the assumption that there are *reasons* for beliefs and intentions, or the assumption that we *ought* to believe or intend certain things. And surely, Kolodny thinks of such reasons and ‘oughts’ as normative, deontic, and response guiding.

Unless there is a relevant difference between believing *p* and being in the state of believing *p*, a reason or rational requirement to believe *p* will thus require you to be in a state. Otherwise, if believing *p* is supposed to qualify as a response but being in the state of believing *p* not, then this difference is sufficiently captured by the distinction between proposition and response requirements.

But perhaps Kolodny does not take ‘doing’ as broadly as I have understood it here. As he puts it, “process requirements require you to *do* something [...], where ‘do’ is understood broadly, so as to include *forming* and *revising* beliefs.”<sup>50</sup> Accordingly, Kolodny’s claim might be that we are never rationally required to *believe* or *intend* anything, but only to *form* or *revise* our beliefs and intentions. All rational requirements then turn out to be essentially practical: Strictly speaking, we are required to perform certain mental actions of forming and revising certain attitudes, not to have the attitudes themselves.

But why should we accept this view? The first thing to note, again, is that the claim is unmotivated, given that Kolodny does not object to the idea that there are *reasons* to believe and intend. He never claims that these reasons are actually reasons for

---

<sup>49</sup> Kolodny (2007b, 371–372).

<sup>50</sup> Kolodny (2007b, 371), my emphasis on “forming” and “revising”.

the mental actions of forming or revising attitudes. It seems uncontroversial that *reasons* for attitudes can guide our responses, so why should it not be possible that rational requirements for attitudes do so, too?

Of course, Kolodny might now claim that reasons for attitudes are also to be understood as reasons for the mental actions of forming and revising attitudes.<sup>51</sup> As far as I can see, however, there is not much to be said in favour of this view. For one, the stress on mental actions of forming and revising attitudes strikes me as an artificial constraint that is often phenomenologically inadequate. The cases in which we respond to our reasons for attitudes by any recognisable act of forming or revising are rare: We often respond to reasons in a much more immediate way. For another, it seems that we can be required by reasons or rationality to perform an action only insofar as this action is intentional.<sup>52</sup> So if reasons to intend are to be understood as reasons to *form* an intention, it seems that every reason for an intention requires a further intention to form this intention, *ad infinitum*. The contrary view, according to which all reasons are actually reasons for judgement-sensitive attitudes, and reasons for action are best understood as reasons for intentions themselves, strikes me as much more plausible.<sup>53</sup> In any case, we should allow for the existence of reasons for attitudes.

To sum up, the view that rational requirements should be able to guide our responses does not give us any reason to accept that rational requirements are actually requiring processes rather than states. If there are reasons for attitudes, then there is no ground for doubting that rational requirements on attitudes could be guiding our responses. And the revisionary view that all reasons for attitudes are actually practical reasons to *form* or *revise* attitudes is not very plausible; it seems, to the contrary, deeply problematic.

---

<sup>51</sup> See Kornblith (2002, ch. 5) for the view that epistemic norms are actually practical norms, namely hypothetical imperatives. Kornblith presents this thesis in the course of a general defence of naturalism in epistemology.

<sup>52</sup> See Scanlon (1998, 21).

<sup>53</sup> This seems to be Scanlon's view when he maintains that "reason for action" is not to be contrasted with "reason for intending" in defence of his claim that "judgment-sensitive attitudes constitute the class of things for which reasons [...] can sensibly be asked for or offered" (1998, 21). See also Smith (2013, 60), who follows Scanlon on this point. I mention this only as a plausible possibility. All that is needed for the argument above is the much weaker claim that not all reasons for attitudes can be reduced to reasons for action.

*Third argument: state-requirements are redundant*

The third and last argument that Kolodny brings forth is an argument from exhaustion. He begins by pointing out that we ordinarily “judge that a person is rational or irrational not only in virtue of the state he is in at a given time, but also in virtue of how he transitions from one state to another over time.”<sup>54</sup> Again, this claim evokes a false dichotomy. If the argument of the last section is correct, we can (almost) never judge a person’s rationality by considering the state she is in at only one point in time; we need to consider at least two different points in time. But Kolodny’s point can be reconstructed to the effect that ordinary attributions of rationality are concerned not only with the states in which we are in at different points in time, but also with how we transition from one state to another.

This claim seems correct. It does not strictly follow that there is a code of rationality that requires certain transitions and bans others, but the existence of such process-requirements is no less plausible than the existence of structural state-requirements of rationality. So far this does not give us any reason to suppose that the requirements which are supposed to explain structural irrationalities such as (A)-(D) should be understood as process-requirements; it only makes a case for the claim that there plausibly are some process-requirements of rationality.<sup>55</sup>

Yet Kolodny also claims that “once we have specified the process-requirements of rationality, state-requirements become superfluous. By complying with process-requirements, one thereby complies with any state-requirement that might apply to one.”<sup>56</sup> But this is a strong thesis, for which Kolodny does not give any argument. To evaluate this claim, we need to know which process-requirements really are supported

---

<sup>54</sup> Kolodny (2005, 517), see also Kolodny (2007b, 371).

<sup>55</sup> Fink (2011) argues that there are no process-requirements of rationality because it is impossible to satisfy such requirements. His argument is that one violates such requirements as long as one is not in the consequent state, and one is not subject to the requirement as soon as one is in the consequent state. However, he revises the first of these assumptions in a later paper, and does so for good reasons. Fink (2012) agrees that there are process-requirements of rationality and indeed presents an interesting argument for their existence. According to him, process-requirements are necessary for determining the *degree* to which a person is rational. I find it difficult to assess this argument, however, because Fink never mentions a particular process-requirement that he embraces. He might therefore be subject to the same complaint that he himself brings forth against Kolodny, namely that Kolodny “talks the process talk”, but “does not walk the process walk” (2011, 478).

<sup>56</sup> Kolodny (2005, 517).

by the fact that ordinary attributions of rationality are partly concerned with how we transition from one state to another, and whether it is plausible to assume that these requirements exhaust the claims of rationality. I will focus on the (A)-(D)-corresponding requirements. One way to argue for Kolodny's exhaustion claim would be this: The (A)-(D)-corresponding *state*-requirements require of us to *be* in certain attitude-states (or others). Ordinary rationality attributions suggest that for each such requirement, there is also a requirement to *form* or *retain* the attitude-state in question. But if we satisfy these latter process-requirements, we will automatically satisfy the state-requirements too. Hence the latter are redundant.

The problem with this argument is that ordinary rationality attributions do not support general requirements to form or retain intentions or beliefs rather than, or in addition to, requirements to intend or believe. We regularly regard intentions and beliefs as rational or irrational, but the cases in which we attribute these properties to the mental actions of *forming* attitudes are rare, if at all existent. So the claim that ordinary rationality attributions are regularly concerned with processes cannot plausibly be understood in this way. On a more plausible interpretation, this claim is concerned with *reasoning*, or more generally, *basing* attitudes on one another. It seems correct to say that ordinary rationality attributions are regularly concerned with how we base attitudes on one another, and since basing an attitude on another is a process, this claim does support the existence of process-requirements.<sup>57</sup> To give an example, rationality seems to require us not to disbelieve that we ought to  $\phi$  on the basis of a lack of intention to  $\phi$  or a belief that we lack such an intention.<sup>58</sup>

The question is thus whether there are general basing requirements that are necessarily violated in (A)-(D), to the effect that the postulation of state-requirements that are violated in (A)-(D) becomes redundant. This will not be the case if, as Broome argues, all basing-requirements are negative, i.e. prohibitions to base certain attitudes

---

<sup>57</sup> It might be said that basing an attitude on another is a way of forming this attitude and that, consequently, basing requirements are cases of requirements to form an attitude, which I excluded as untypical above. This is compatible with my argument, however. I assume only that ordinary rationality attributions do not support widespread general requirements to form attitudes insofar as these requirements are independent of basing.

<sup>58</sup> Broome (2013, 141) embraces this requirement. See also Chapter 6.4 for more on this.

on others.<sup>59</sup> It is certainly possible to be in conditions like those described by (A)-(D) without violating a basing prohibition. One may, for example, believe that one ought to  $\phi$  on the basis of evidence, and not intend to  $\phi$  without any basis at all. In such a case, (A) implies that one is irrational, but one has not violated a basing prohibition.

So in order for Kolodny's thesis to be true, there need to be positive basing principles that require us to base certain attitudes on one another, and these principles have to entail that (A)-(D) are cases of irrationality. Kolodny's own approach is to say that agents that are in the antecedent state are always rationally required to base the consequent state on the antecedent state. The version of the self-governance requirement he favours thus reads:

Necessarily, if you believe at  $t$  that you ought to  $X$ , but you do not intend at  $t$  to  $X$ , then rationality requires you to form going forward from  $t$ , on the basis of the content [of] your belief, the intention to  $X$ .<sup>60</sup>

Note that this requirement has narrow-scope. Part of Kolodny's purpose is to argue for the narrow-scope account by showing that no plausible wide-scope counterpart of this basing-requirement exists. Kolodny's self-governance requirement, and analogous requirements of this form, could indeed explain why (A)-(D) are cases of irrationality; we would need no state-requirements to do this job. But as I shall argue, there are weighty reasons to reject requirements of this form, quite independently of the controversy over the scope of requirements.

First, the requirement is too weak to fully account for ordinary rationality attributions concerning (A). The requirement is conditional not only on your believing at  $t$  that you ought to  $\phi$ , but also on your not intending at  $t$  to  $\phi$ . It must include this condition, because a requirement to form an intention that one already has does not make sense. Consequently, the requirement only applies to you if you have incoherent attitudes at  $t$ . So while the requirement suffices to show that an agent who is in the

---

<sup>59</sup> See Broome (2013, 141). Broome also embraces basing permissions, but these are just negations of basing prohibitions (see Broome 2013, 190).

<sup>60</sup> Kolodny (2007b, 373).

condition described in (A) *violates* a rational requirement, it cannot explain why agents who intend in coherence with their ought-beliefs *satisfy* a rational requirement, even if they have not been incoherent before (see the discussion on this in the last section). And this indicates that even if Kolodny's requirement were true, it would need to be complemented by a state-requirement.<sup>61</sup>

More importantly perhaps, the requirement is also too strong. Suppose you believe that you ought to visit your friend in the hospital, but you do not yet intend to do so.<sup>62</sup> Say you then form the intention to visit your friend at the hospital on some basis other than your belief that you ought to do so. Perhaps you base your intention on a belief that you also take to support the ought-claim, e.g. the belief that your visit would cheer up your friend. Or you form the intention simply on the basis of your desire to see your friend. According to Kolodny's requirement, you are irrational. But this seems clearly false. There need be no irrationality involved in forming the intention in these ways, even if you believe you ought to visit your friend.

This case does not only provide a counterexample to Kolodny's requirement, I think it also shows that, quite generally, the prospects of finding plausible basing-requirements that could explain the structural irrationalities described by (A)-(D) are very dim. For such an explanation to work, we would need a very restrictive account of rational basing-requirements: it would have to be true that for every rational requirement to be in a certain attitude-state (or disjunctive attitude-state), there is a requirement to *base* that attitude-state (or one of the attitude-states part of the disjunctive attitude-state) in a particular way on some other present attitude. It seems, however, that rationality normally permits us to base our attitudes – at least our intentions – on

---

<sup>61</sup> Perhaps this objection could be avoided if Kolodny, instead of his suggestion, accepted the disjunctive requirement claim "If you believe at  $t$  that you ought to  $\phi$ , then rationality requires you to [form ... the intention to  $\phi$  or retain your intention to  $\phi$ ]" (not to be confused with the relevant wide-scope requirement which would allow satisfaction by dropping the antecedent state). Note, however, that this requirement would not be a basing-requirement anymore, and that it would also be vulnerable to the objection I go on to raise in the main text.

<sup>62</sup> This example is inspired by a case presented by Stocker (1976, 462), which has been much discussed in the context of debates about the value of acting from duty.

a variety of considerations; perhaps it permits us to acquire intentions without any basis at all.<sup>63</sup>

In any case, as long as Kolodny does not give an alternative account of process-requirements that plausibly renders state-requirements redundant, the fact that ordinary attributions of irrationality are partly concerned with processes does not give us any reason to conclude that all rational requirements are process-requirements. In particular, it does not provide any reason to think that the requirements that are violated in the conditions described in (A)-(D) are basing- or any other process-requirements. To the contrary, the idea that they are is fraught with problems.

### *Conclusion*

The purpose of this section has been mainly negative: I have argued that the requirements of rationality that correspond to structural irrationalities as described in (A)-(D) are plausibly not process-requirements. I do think, however, that there is an important idea behind Kolodny's emphasis on processes, so let me conclude this section with a more constructive remark. It is uncontroversial that certain transitions from one state to another are rational and others irrational. If that is so, then it is plausible to assume that state-requirements of rationality have to meet a certain constraint, namely that (unless one does not already satisfy such a requirement) there must be a rational process available to an agent, which leads to its satisfaction. Let us call this the

*Rational process constraint:* For every rational requirement R, R requires A to be in state S, only if [either A satisfies R or there is a rational process available to A which leads to S].

Some remarks on this. First, by "rational processes", I do not mean rationally *required* processes, but processes that are (1.) governed by rational process requirements, and (2.) do not violate any of these requirements. Dreaming, for example, is not a rational process because it is not governed by rational requirements, and fallacious reasoning is

---

<sup>63</sup> As Broome (2013, 189) maintains.

not a rational process because it violates the requirements by which it is governed. Correct reasoning, in contrast, is a rational process, but it may not be a rationally *required* process. To see why, suppose you are required to be in a certain state, and there is a correct course of reasoning available to you that would lead you to be in this state, but there also is another course of correct reasoning available, that leads to the same conclusion. You are then not required to reason in the first way, even though this would be a rational process. Note that not all rational processes need to be processes of reasoning: Believing something on the basis of a perception might not count as reasoning, but still count as a rational process.

Second, by saying that the process must be “available” to A, we need not suppose something as strong as that A needs to be psychologically capable of running through the process; it suffices to say that A would be capable of running through this process if her rational capacities were in ideal conditions, or something of that sort. We thereby leave open the controversial question of to what extent A’s incompetence to run certain rational processes might itself count as a form of irrationality. The point of the rational process constraint is merely that a possible rational process exists.

Read in this way, the rational process constraint seems hard to deny. A rational state-requirement to the effect that A is required to be in a state S at a certain time entails that A is irrational if A is not in this state at that time. Irrationality, in turn, must involve some kind of rational failure. But now suppose that there is no rational process available to A which could lead A to S. This might be because there is no process available at all, or because the only processes available are irrational processes (e.g. fallacious reasoning) or arational processes, i.e. processes not governed by rationality (e.g. meditation). If there is no process available at all, A’s not being in S cannot be the result of any failure of A’s. If the only process available is irrational, then A’s not being in S might possibly result from some kind of failure, but certainly not a failure of rationality. And the same is true if the only process available is arational. If you can come to be in a state only by some arational process such as meditation, then a failure not to come to be in this state may be failure of your meditation skills or the failure of other non-rational capacities of yours, but not a failure of rationality. So any rational

state-requirement seems to involve, by way of its implications about irrationality, a commitment to the rational process constraint.

Moreover, the rational process constraint seems to me to capture the intuitive ideas of Kolodny's main concerns about state-requirements. It accounts for the fact that ordinary attributions of rationality are concerned not only with states, but also with transitions between states, without presupposing that process-requirements exhaust all state-requirements. More importantly, it assures us that state-requirements can be guiding, for it guarantees that there is a rational process available that leads to the satisfaction of a state-requirement, without presupposing that this process is itself rationally required. This seems to me to be the true kernel of Kolodny's concern with processes. Instead of saying that rationality can require only processes, we should say that rationality can require states only if there are rational processes leading to that state. I shall come back to this idea in Chapter 6.

### 3.6 Summary

In this chapter, I have discussed some basic questions concerning the form of structural requirements of rationality. I have argued that both narrow- and wide-scope versions of (A)-(D)-corresponding requirements are best understood as relating to responses rather than propositions, as being conditional on the presence of the antecedent state, and as being diachronic rather than synchronic, though they are plausibly state- rather than process-requirements. These results are important in themselves, insofar as they contribute to an understanding of structural requirements of rationality. But they also provide an important background against which we can tackle the major disagreements about such requirements: the disagreement about whether these requirements have wide or narrow scope, and about whether they are normative or not.

## Bootstrapping and other detachment problems

---

In this chapter, I consider a family of difficulties that arise when we assume, in line with the last two chapters, that rationality is normative and that there are structural requirements of rationality. The general problem is that taken together, these two assumptions seem to allow for the detachment of unacceptable normative conclusions about what we ought or have reason to do. I begin by illustrating the *bootstrapping problem* that occurs when we take the relevant requirements to have narrow scope (4.1), and then discuss and reject two strategies to solve this problem: the reasons strategy (4.2), and the subjective ‘ought’ strategy (4.3). I go on to present a third, and more promising, strategy, which blocks bootstrapping by taking structural requirements of rationality to have wide scope (4.4). In the remainder of the chapter, I discuss further detachment problems that arise when the wide-scope account is combined with independent principles about the transmission of reasons and ‘oughts’: *normative detachment* (4.5), *liberal transmission* (4.6), and *necessary means transmission* (4.7). My conclusion is that the wide-scope account ultimately fails to block detachment of unacceptable normative conclusions (4.8).

### 4.1 The bootstrapping problem

Let us suppose, for the sake of the argument, that rationality is normative in a strict sense:

*The normativity of rationality, strict:* If A is rationally required to  $\phi$ , then A ought to  $\phi$ .

Further, let us suppose that structural requirements of rationality exist, and that they are to be understood along the narrow-scope interpretation (A)<sub>2</sub>-(D)<sub>2</sub>. What we get are these four *normative* requirements of rationality:

- (A)<sub>9</sub> *Normative self-governance requirement, narrow-scope*: If A believes that she ought to  $\phi$ , then A ought to intend to  $\phi$ .
- (B)<sub>9</sub> *Normative instrumental requirement, narrow-scope*: If A intends to  $\phi$ , and if A believes that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then A ought to intend to  $\psi$ .
- (C)<sub>9</sub> *Normative evidence-belief requirement, narrow-scope*: If A believes that she has sufficient evidence that  $p$ , then A ought to believe that  $p$ .
- (D)<sub>9</sub> *Normative modus ponens requirement, narrow-scope*: If A believes that  $p$ , and if A believes that  $p \rightarrow q$ , then A ought to believe that  $q$ .

Since these requirements have narrow scope, they allow for detachment. Whenever the antecedent is satisfied, we can detach a normative conclusion about what we ought to do. The antecedent is satisfied whenever we adopt the attitudes mentioned in it. As a consequence, (A)<sub>9</sub>-(D)<sub>9</sub> entail that just by adopting an intention or a belief, we can make it the case that we *ought* to adopt some other attitude, independently of any reasons that we had to adopt the former attitude in the first place. This, however, seems untenable.

Michael Bratman first posed this problem under the name “bootstrapping” in discussing instrumental rationality.<sup>1</sup> He describes a case similar to this one: Suppose you are weighing your reasons for and against two incompatible courses of actions, say getting some work done at home and watching a football match with your friends. Your reasons, we can suppose, together require you to stay home: you have to hand in important work tomorrow, and the match is not supposed to be very promising, after all. In deliberating, you are reaching the correct conclusion that you ought to stay home. But then, out of weakness of will, you decide to go watch the football match nevertheless. You also recognise that in order to see the football match with your friends, you need to call them and ask where they are meeting. So now you intend to watch the football match, and you believe that in order to do so, you have to call your friends. According to the normative interpretation of the *instrumental requirement* (B)<sub>9</sub>, which is

---

<sup>1</sup> In Bratman (1981). See also Bratman (1987, 24–27). Note that Bratman had changed his mind by the time of the second publication. While he defends the claim that intentions are reasons against the bootstrapping objection in the former, he rejects such bootstrapping in the latter.

under discussion here, we can now detach the conclusion that you ought to call your friends. But this seems an absurd conclusion, given that we have just said that you ought *not* to meet your friends, but rather stay home. The fact that you intend to meet your friends contrary to your own recognised reasons cannot change the fact that you ought not to meet them, and thus cannot make it the case that you ought to take steps to meeting them. You cannot ‘bootstrap’ a decisive reason to take some means into existence, simply by intending an end you have decisive reason not to intend.

A.C. Ewing encounters a similar problem when he discusses the *self-governance requirement*:

It is a recognized principle of ethics that it is always our duty to do what we when considering it sincerely think we ought to do, but suppose we are mistaken, then we by this principle ought to do something which is wrong and which therefore we ought not to do. Is not this a contradiction?<sup>2</sup>

Suppose, with Ewing, that you sincerely but mistakenly believe that you ought to do something. You could believe all sorts of crazy things, for example that you ought to put all your efforts into building a machine which enables you to travel through time and meet Elvis Presley. So let us suppose that you ought not to put all your efforts into this project, and consequently ought not to intend to do so, but you sincerely believe you ought to do it. By the normative interpretation of the *self-governance requirement* (A)<sub>9</sub>, we can now detach the conclusion that you ought to intend to put all your efforts into this crazy project. However, as Ewing points out, this is a contradiction, for we have just said that you ought not to do that, and consequently not intend to do that. We thus seem forced by this argument to conclude that you cannot have a false belief about what you ought to do. But your ought-beliefs cannot make themselves true. You cannot, just by adopting some crazy belief about what you ought to do, make it the

---

<sup>2</sup> Ewing (1953, 144–145). The problem he describes has come to be called “Ewing’s problem”; see Piller (2007) for an extensive discussion. Ewing’s own solution is to distinguish between different senses of ‘ought’. I consider this *subjective ‘ought’ strategy* below.

case that you really ought to do it – you cannot bootstrap an ‘ought’ into existence simply by believing in it.

Similar considerations apply to the theoretical requirements as well. Suppose that due to brainwashing by a religious sect, you believe that the world is not more than 1,000 years old. You have no reason for having this belief, but you have it anyway, and you further believe that if the world is not more than 1,000 years old, then the theory of evolution is false. By the normative interpretation of the *modus ponens requirement* (D)<sub>9</sub>, you now ought to believe that the theory of evolution is false. But this, again, is unacceptable bootstrapping. Similarly, suppose that you believe you have sufficient evidence that the world is not more than 1,000 years old, even though in fact you have no evidence believe this. On the version of the *evidence-belief requirement* under discussion (C)<sub>9</sub>, you now ought to believe that the world is not more than 1,000 years old. But this is an absurd conclusion.

The lesson to draw is that rationality cannot be strictly normative if structural requirements of rationality take narrow scope, since (A)<sub>9</sub>-(D)<sub>9</sub> all involve unacceptable bootstrapping. We have no choice but either to reject (A)<sub>2</sub>-(D)<sub>2</sub> or give up the strict normativity of rationality.

## 4.2 The reasons strategy

Before I discuss the wide-scope strategy to deal with the bootstrapping problem, I will briefly consider two alternative strategies that hold on to the narrow-scope interpretation and weaken the normativity of rationality instead. The first of these strategies suggests that instead of claiming that we always *ought* to be rational, we should only assume that we have a *reason* to do so:<sup>3</sup>

---

<sup>3</sup> The reasons strategy is part of Schroeder’s (2004) account of the instrumental requirement. Schroeder is predominantly concerned with desires rather than intentions, but he invites the reader to “substitute ‘end’ or ‘intention’” for ‘desire’ (2004, 362, n. 6). His view is nevertheless more complex than the one discussed here, since for Schroeder means *believed* necessary to desired ends give rise only to what he calls “subjective reasons”. In a later article, Schroeder renounces the reasons strategy with respect to intentions and adopts a version of the subjective ‘ought’ strategy instead (see Schroeder 2009), which will be discussed in the next section. Cheng-Guajardo (forthcoming) also makes use of the reasons strategy.

*The normativity of rationality, contributory:* If A is rationally required to  $\phi$ , then A has a reason to  $\phi$ .

If the normativity of rationality is understood in this way, (A)<sub>2</sub>-(D)<sub>2</sub> do not entail the unacceptable claims (A)<sub>9</sub>-(D)<sub>9</sub> anymore. Instead, they only entail the corresponding claims about reasons. To avoid unnecessary reiteration, I will confine the discussion to the practical requirements:

(A)<sub>10</sub> *Pro tanto normative self-governance requirement, narrow-scope:* If A believes that she ought to  $\phi$ , then A has a reason to intend to  $\phi$ .

(B)<sub>10</sub> *Pro tanto normative instrumental requirement, narrow-scope:* If A intends to  $\phi$ , and if A believes that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then A has a reason to intend to  $\psi$ .

These claims seem less vulnerable to counterexamples than (A)<sub>9</sub>-(D)<sub>9</sub>. Since reasons can be outweighed, no implausible ‘ought’-claims can be directly detached from these requirements. Nevertheless, there are severe problems with the reasons strategy.

The first problem is that (A)<sub>10</sub> and (B)<sub>10</sub>, while they do not allow for bootstrapping of ‘oughts’, still allow for bootstrapping of reasons. Bootstrapping of reasons is *less* implausible, because it does not necessarily affect the normative overall level. But it is still implausible that simply by adopting an attitude, we can make it the case that we have a reason to adopt some other attitude, independently of any reason we have for the former attitude in the first place.<sup>4</sup>

---

<sup>4</sup> There is a close relationship between the bootstrapping objection concerning reasons and what Schroeder (2007a, ch. 5) calls the “too many reasons problem” for desire-based accounts of reasons. Both find fault with theories that make it very easy to come by reasons with *prima facie* implausible content. The difference is that the too many reasons problem is primarily concerned with the extension of reasons, whereas the bootstrapping objection is primarily concerned with the genesis of reasons. The too many reasons problem is that desire-based theories entail the existence of reasons that, intuitively, are no reasons at all. Schroeder’s strategy to deal with this problem is to argue that negative existential intuitions about reasons are systematically misleading. In contrast to the too many reasons problem, the bootstrapping problem is not (at least not primarily) that the reasons implied by a theory have implausible content, but that it is implausible that one can generate such reasons simply by adopting an intention or belief. Examples of reasons with implausible content generated by bootstrapping merely reinforce this point about genesis. For this reason, I do not think that Schroeder’s response to the too many reasons problem automatically works as a response to the boot-

When Bratman first considered bootstrapping in the case of instrumental rationality, he was concerned about the bootstrapping of reasons, not of ‘oughts’.<sup>5</sup> Let me recall the case that was inspired by his observations: You know that your reasons weigh in favour of working at home, but akratically you decide to watch football nevertheless. According to (B)<sub>10</sub>, you now have a new reason with independent force that counts in favour of taking the means to watching football. This reason might well tip the scale and thus make it the case that, all things considered, you now *ought* to take these means. But this is an implausible consequence. Plausibly, you cannot bootstrap such a reason into existence by akratically adopting an intention.

Questions concerning reasons, motivating states like intentions, and instrumental rationality are notoriously contentious, however. It is part of the conception of many desire-based accounts of reasons that every desire provides a reason to promote its satisfaction.<sup>6</sup> And according to some theories in the philosophy of mind and action, intentions just are a certain kind of desire, or at least involve a desire.<sup>7</sup> Proponents of such theories might therefore approve of the bootstrapping described in the football case. Moreover, proponents of desire-based theories have presented sophisticated explanations of why counterintuitive reasons, such as the putative reason based on your irrational intention to watch football, are of particularly low weight and almost never affect what we ought to do, and why it therefore seems to us that they do not exist.<sup>8</sup>

Even if we buy into such an account of desire-based reasons, however, the bootstrapping objection still holds. The first thing to note is that (B)<sub>10</sub> not only allows for the bootstrapping of a reason by adopting an *intention* – which might be permitted according to certain desire-based accounts of reasons – but also by adopting a *belief* about necessary means. (B)<sub>10</sub> entails, for example, that if you intend to make a good impression on your first date, and if you believe (without any reason) that in order to do so, you need to dress as a parrot and tell a lot of sexist jokes, then you have a reason

---

strapping objection as well. Even if our intuitions about negative reasons existentials are systematically misguided, this does not show that our intuitions about the genesis of reasons are likewise misguided.

<sup>5</sup> Bratman (1981; 1987, 24–27).

<sup>6</sup> See e.g. Schroeder (2007a).

<sup>7</sup> See e.g. Ridge (1998), Sinhababu (forthcoming), Smith (1994, 117).

<sup>8</sup> See esp. Schroeder (2005b, 6–11; 2007a ch. 5 and ch. 7; 2007d).

to intend do so. But this is implausible bootstrapping, which does not seem to gain support from a desire-based account of reasons.

Secondly, even if (B)<sub>10</sub> can be defended in connection with a desire-based account of reasons, the bootstrapping objection still applies to the other requirements that follow from the reasons strategy. It is implausible, for example, that you can make it the case that you have a reason to intend to put all your efforts into building a time machine, simply by believing that you ought to do so, or that you have a reason to believe that the earth is flat, simply by believing that you have evidence for it. But if only one of these consequences is unacceptable, this suffices to show that the solution to our problem cannot generally consist in weakening the normativity of rationality from 'oughts' to reasons.

A further problem for the reasons strategy is that it cannot account for the criticism that is commonly associated with irrationality.<sup>9</sup> As I have argued above, criticism cannot be grounded in the fact that one had *some* reason or other to conform to the rational requirement; it presupposes that these reasons were decisive. It is not appropriate to criticise people for not conforming to a reason if this reason was outweighed or counterbalanced. Our practice of irrationality ascriptions thus presupposes strict normativity, not merely contributory normativity. Note that even if we admitted the possibility of uncriticisable irrationality, this could not sustain the reasons strategy, since even the proponent of this strategy would not want to claim that all cases of irrationality in which the bootstrapped reason was outweighed are cases of uncriticisable irrationality. One case of a person who is criticisably irrational even though her bootstrapped reason was outweighed suffices to show that the reasons strategy is incapable of accounting for criticisability.

To sum up, the reasons strategy fails on all grounds: first, because it still allows for unacceptable bootstrapping, and second, because the connection between rationality and normativity it proposes cannot account for the practice of rational criticism.

---

<sup>9</sup> Schroeder (2009, 232) concedes that the reasons strategy fails to account for what he calls the "stringency" of rational requirements, which seems to be a related point. Compare also Broome's "strictness test" (1999, 405).

### 4.3 The subjective ‘ought’ strategy

Another strategy to deal with the bootstrapping problem is to introduce a subjective notion of ‘ought’.<sup>10</sup> The idea is that rationality connects with subjective ‘oughts’ only, and therefore does not lead to the bootstrapping of any other ‘oughts’:

*Subjective normativity of rationality:* If rationality requires you to  $\phi$ , then you subjectively ought to  $\phi$ .

This version of the normativity of rationality, together with the narrow-scope interpretation of rational requirements, yields the following claims (I focus again on the practical requirements):

(A)<sub>11</sub> *Subjectively normative self-governance requirement, narrow-scope:* If A believes that she ought to  $\phi$ , then A subjectively ought to intend to  $\phi$ .

(B)<sub>11</sub> *Subjectively normative instrumental requirement, narrow-scope:* If A intends to  $\phi$ , and if A believes that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then A subjectively ought to intend to  $\psi$ .

It is unclear, however, what the so-called subjective ‘ought’ really amounts to. On one possible view, to say that A subjectively ought to  $\phi$  just *means* that A believes that she ought to  $\phi$ , or that A would have decisive reason to  $\phi$  if her beliefs were true, and thus does not entail anything about A’s actual normative reasons to  $\phi$ . On this view, (A)<sub>11</sub> and (B)<sub>11</sub> may seem unproblematic – indeed, (A)<sub>11</sub> turns out to be almost trivially true. However, it also seems to follow that subjective ‘oughts’ are not normative require-

---

<sup>10</sup> See e.g. Ewing (1947, 120–21); Parfit (2011a, 33–34); Schroeder (2009); Way (2009). Setiya (2007a, 657–59) adopts the subjective ‘ought’ strategy for the self-governance requirement, but not for the instrumental requirement (Setiya’s cognitivist view about instrumental rationality is discussed in Chapter 10.3).

ments at all.<sup>11</sup> We can set this view aside, here, for we are looking for a vindication of the normativity of rationality.

On another view, subjective ‘oughts’ are genuinely normative requirements.<sup>12</sup> Proponents of this view either have to deny the reasons claim – i.e. the thesis that A is normatively required to  $\phi$  iff A has decisive reason to  $\phi$  – or maintain that subjective ‘oughts’ are supported by subjective reasons, understood as genuinely normative reasons.<sup>13</sup> In either of these cases, it looks as if the subjective ‘ought’ strategy could in principle yield a vindication of the normativity of rationality. However, it is then unclear how it avoids the charge of unacceptable bootstrapping. If subjective ‘oughts’ are genuinely normative requirements, then (A)<sub>11</sub> and (B)<sub>11</sub> entail that just by adopting an attitude for no reason, we can make it the case that we stand under a genuinely normative requirement to adopt some other attitude. But this seems doubtful. It seems doubtful, for example, that just by believing, without any reason, that you could impress your date by dressing as a parrot and telling a lot of sexist jokes, you can make it the case that you ought, in a genuinely normative sense, to intend to do so. It seems also doubtful that someone who believes that he ought to kill all homosexuals thereby makes it the case that he ought, in a genuinely normative sense, to intend to do so. To the extent that the subjective ‘ought’ strategy is capable of vindicating the normativity of rationality, it seems subject to the bootstrapping objection as well.

There is a further problem for the subjective ‘ought’ strategy, understood as a vindication of the normativity of rationality. Proponents of this view have to clarify

---

<sup>11</sup> Finlay (2010, 72), too, raises the worry that subjective ‘oughts’, interpreted as believed or apparent ‘oughts’, are not normative. The question is considered in more detail in Chapter 5.6.

<sup>12</sup> This seems to be Schroeder’s view. According to it, subjective ‘oughts’ can be analysed in terms of subjective reasons, which Schroeder takes to be genuinely normative. Schroeder also analyses subjective reasons as apparent objective reasons, i.e. objective reasons that one would have if one’s beliefs were true, but importantly he calls this a “substantive theory” (2007a, 14). It thus appears that according to his view, by saying that (1) [A ought subjectively to  $\phi$ ] one makes a substantive normative claim that goes beyond the assertion that (2) [A has a collection of beliefs such that it *would* be the case that A ought to  $\phi$  if these beliefs were true], even though it is also part of his view (at least roughly) that for (1) to be the case just is for (2) to be the case.

<sup>13</sup> Parfit (2011a, 34–35) seems to take the former option and Schroeder (2007a, ch. 1.3; 2009, §5) the latter. However, as I understand the reasons claim, it is about reasons in the sense in which reasons matter in deliberation, and I am sceptical that what Schroeder calls “subjective reasons” are reasons in this sense (I am sceptical for the same reasons that I am pressing here against the idea that the subjective ‘ought’ is the ‘ought’ that matters in deliberation). This suggests that the postulation of a genuinely normative subjective ‘ought’ is incompatible with the reasons claim (as I understand it), regardless of whether this subjective ‘ought’ is supposed to be supported by subjective reasons.

how the subjective ‘ought’ relates to what I have called the *deliberative ‘ought’* in Chapter 1 – the ‘ought’ that figures in the content of those beliefs that are rationally connected to intentions, i.e. the beliefs mentioned in the antecedent of (A)<sub>11</sub>. They can either say that the subjective ‘ought’ *is* the deliberative ‘ought’ or that it is different from it. Now, if the subjective ‘ought’ were the deliberative ‘ought’, then (A)<sub>11</sub> would entail a particularly objectionable form of bootstrapping; it would tell us that whenever we conclude in deliberation that we ought to do something, then we ought, *in the same sense*, to intend to do it. But this is absurd. Surely it is possible for us to conclude in deliberation that we ought to do something, even if it is not the case that we ought to do it or ought to intend to do it. Deliberative conclusions are not generally infallible. So the proponent of the subjective ‘ought’ strategy must be understood as invoking a sense of ‘ought’ that is different from the one that matters in deliberation.

But when we were asking for the normativity of rationality, we were looking for an answer that *does* matter in deliberation. Further, the normativity presupposed in rational criticism is a kind of normativity that does matter in deliberation, for it seems incoherent to criticise a person for a response while accepting that it would have been correct for her to conclude in deliberation that she ought to give this response.<sup>14</sup> The subjective ‘ought’ strategy thus misses the point. Even if it could establish that we subjectively ought to be rational in a genuinely normative sense, this would just not be the relevant sense. What we want to know is whether we ought to be rational in the sense of ‘ought’ that matters for deliberation.

#### 4.4 The wide-scope strategy

The bootstrapping problem appears when we combine the narrow-scope interpretation of structural requirements of rationality with the claim that rationality is strictly normative. In the last two sections, I have discussed two strategies to deal with the bootstrapping problem by giving a different interpretation of the normativity of rationality. Both of these strategies failed: for one, because they still were susceptible to the boot-

---

<sup>14</sup> See Chapter 2.2 for a defence of this claim. See Chapter 2.3 for an argument for the thesis that the violation of a subjective ‘ought’ is in itself no ground for criticism.

strapping problem; for another, because the alternative interpretations of the normativity of rationality could not plausibly account for our practice of rational criticism.

We can conclude that the narrow-scope interpretation of structural requirements of rationality is incompatible with any plausible account of the normativity of rationality. This leaves us with the wide-scope strategy to deal with the bootstrapping problem. As we have already seen, the wide-scope account rejects the narrow-scope requirements (A)<sub>2</sub>-(D)<sub>2</sub> and replaces them with wide-scope requirements such as (A)<sub>3</sub>-(D)<sub>3</sub>, or rather (A)<sub>5</sub>-(D)<sub>5</sub> (if we take into account some of the discussion of Chapter 3, while putting aside – for the sake of simplicity – matters of conditionality and diachronicity). Together with the strict normativity of rationality, we get normative requirements of the following form:

- (A)<sub>12</sub> *Normative self-governance requirement, wide-scope*: A ought to [not believe that she ought to  $\phi$  or intend to  $\phi$ ].
- (B)<sub>12</sub> *Normative instrumental requirement, wide-scope*: A ought to [not intend to  $\phi$  or not believe that  $\psi$ -ing is a necessary means for  $\phi$ -ing or intend to  $\psi$ ].
- (C)<sub>12</sub> *Normative evidence-belief requirement, wide-scope*: A ought to [not believe that A has sufficient evidence that p or believe that p].
- (D)<sub>12</sub> *Normative modus ponens requirement, wide-scope*: A ought to [not believe that p or not believe that  $p \rightarrow q$  or believe that q].

Since wide-scope requirements do not allow for factual detachment, they seem to avoid the bootstrapping problem. The wide-scope account does not entail that just by adopting the antecedent attitude, you can make it the case that you ought to adopt the consequent attitude, for you can just as well satisfy the requirement by dropping the antecedent attitude. In contrast to the narrow-scope account, the wide scope account seems therefore compatible with the strict normativity of rationality.

It has been argued, however, that the wide-scope account has unacceptable normative implications as well. While wide-scope requirements do not generally allow for factual detachment, they do allow for detachment in particular cases, via certain

bridge principles. In the next three sections, I shall discuss three such bridge principles and consider their implications for the wide-scope account.

#### 4.5 Wide-scope and *normative detachment*

Recall that the wide-scope account escapes the bootstrapping objection only because the following principle is false:

*Factual detachment*: If A ought to  $[\phi$  or  $\psi]$ , and A does not  $\phi$ , then A ought to  $\psi$ .

If either wide-scope requirements or wide-scope ‘oughts’ allowed for *factual detachment*, the wide-scope account would, in conjunction with the normativity of rationality, entail the same bootstrapping as the narrow-scope account. I argued against *factual detachment* with respect to requirements rather than ‘oughts’ in the last chapter. The fact that A does not  $\phi$  does nothing to undermine the fact that A can  $[\phi$  or  $\psi]$  – and thus satisfy a disjunctive requirement to  $[\phi$  or  $\psi]$  – either by  $\phi$ -ing or by  $\psi$ -ing. That is, even if A does not in fact  $\phi$ , A *can* satisfy this disjunctive requirement by  $[\phi$ -ing and not  $\psi$ -ing], without thereby violating any requirement. Therefore, we cannot detach a requirement to  $\psi$ . The same point applies to ‘oughts’.

Consider, however, the following, more plausible detachment principle:

*Normative detachment*: If A ought to  $[\phi$  or  $\psi]$ , and A ought not to  $\phi$ , then A ought to  $\psi$ .

*Normative detachment* has intuitive appeal, and it avoids the objection against *factual detachment*. Surely, the fact that A *ought* not to  $\phi$ , also does not change the fact that A can  $[\phi$  or  $\psi]$  either by  $\phi$ -ing or  $\psi$ -ing. However, satisfying a disjunctive requirement to  $[\phi$  or  $\psi]$  by  $\phi$ -ing in this case *does* constitute a violation of an ‘ought’, and the only way to satisfy the disjunctive requirement without violating an ‘ought’ is by  $\psi$ -ing. Hence,

the detachment of a requirement to  $\psi$  seems plausible in this case. As John Gibbons puts it, “if there’s only one permissible option, that’s what you ought to do”.<sup>15</sup>

*Normative detachment* is also entailed by two other plausible principles:

*Agglomeration*: If A ought to  $\phi$ , and A ought to  $\psi$ , then A ought to [ $\phi$  and  $\psi$ ].

*Necessary means transmission*: If A ought to  $\phi$ , and  $\psi$ -ing is a necessary means for A to  $\phi$ , then A ought to  $\psi$ .

For according to *agglomeration*, if A ought to [ $\phi$  or  $\psi$ ], and A ought not to  $\phi$ , then A ought to [( $\phi$  or  $\psi$ ) and not  $\phi$ ]. And a necessary means to doing this is to  $\psi$ .

*An objection to normative detachment*

John Broome rejects *normative detachment* by way of a counterexample.<sup>16</sup> Assuming that you have entered a marathon, we might suppose:

- (1) You ought to exercise hard every day.
- (2) You ought to make sure that if you exercise hard every day, then you eat heartily.

The second assumption is supposed to be a wide-scope ‘ought’, which is equivalent to the claim that you ought to [not exercise hard every day *or* eat heartily].<sup>17</sup> Hence, according to *normative detachment*, it follows that

- (3) You ought to eat heartily.

---

<sup>15</sup> Gibbons (2009, 172). See also Finlay (2010, 71) and Greenspan (1975, 260).

<sup>16</sup> See Broome (2013, 120–21); see also the structurally equivalent example in Broome (2007b, 17–19). In both cases, Broome discusses a detachment principle for requirements rather than ‘oughts’, and he presents his counterexample in terms of requirements of prudence. I can restate his example in terms of ‘oughts’ because Broome is explicit that his rejection of *normative detachment* (Axiom K, as he calls it) “goes for ought as well as for requirement” (2013, 128).

<sup>17</sup> I employ the ‘making sure that’ construction of wide-scope ‘oughts’ merely for reasons of convenience here. Even though ‘A makes sure that [if A  $\phi$ -s, then A  $\psi$ -s]’ is not strictly speaking equivalent to ‘A does not  $\phi$  or A  $\psi$ -s’ (as was argued in Chapter 3.2), I here *use* the former expression to express the same content. The same points could be made in terms of disjunctive ‘oughts’.

Broome objects that it is possible that (1) and (2) are true, while (3) is false. For (1) and (2) seem compatible with:

(4) You do not exercise hard every day.

(5) If you do not exercise hard every day, then it is not the case that you ought to eat heartily.

And (4) and (5) together entail the negation of (3):

(6) It is not the case that you ought to eat heartily.

So if (5) is compatible with (1) and (2), then (3) cannot follow from (1) and (2), and *normative detachment* needs to be rejected.

This counterexample seems plausible at first sight, but turns out to rely on very problematic assumptions. I shall argue that we should reject the claim that (5) is compatible with (1) and (2) in favour of the view that (1) and (2) entail (3). To begin with, it seems independently plausible that given (1) and (2), we may also assume:

(7) You ought to [exercise hard every day *and* eat heartily].<sup>18</sup>

This means that if we deny that (3) follows from (1) and (2), we violate not only *normative detachment* (and thus either *agglomeration* or *necessary means transmission*), but also:

*Distribution*: If A ought to [ $\phi$  and  $\psi$ ], then A ought to  $\phi$ .

---

<sup>18</sup> This also follows from *agglomeration* and *necessary means transmission*: ‘You ought to exercise’ and ‘you ought to [not exercise *or* eat]’ agglomerate to ‘you ought to [exercise and (not exercise *or* eat)]’, and a necessary means to doing this is to [exercise *and* eat].

For given that *distribution* holds, (3) follows from (7). I take this to be independent support for the claim that (3) follows from (1) and (2), since *distribution* is intuitively quite compelling.<sup>19</sup>

Further, and more importantly, it seems that the same kind of considerations that lead us to think that (1) and (2) are compatible with (5) can also be employed in arguing that (1) and (2) are compatible with:

(8) If you do not exercise hard every day, then you ought not to eat heartily.

In both cases, the point is that (1) and (2) tell us nothing about how good or bad it is for you to eat heartily if you do not in fact exercise at all. Depending on how bad it is, intuition favours not only the conclusion that *it is not the case* that you ought to eat heartily, but also the stronger one that *you ought not* to do it. In other words, if despite (1) and (2) it can be permissible not to eat heartily, then it can also be required not to eat heartily. So anyone who accepts that (5) is compatible with (1) and (2) should likewise be willing to accept that (8) is compatible with (1) and (2).

If this is so, however, Broome's assumptions violate:

*Joint satisfiability*: If A ought to  $\phi_1$ , ..., and A ought to  $\phi_n$ , then it is possible for A to  $\phi_1$ , ..., and  $\phi_n$ .

For (4) and (8) entail:

(9) You ought not to eat heartily.

And (1), (2), and (9) cannot possibly be satisfied together. *Joint satisfiability*, however, is confirmed by a number of considerations. First, as Bernard Williams points out, it follows from two other eminently plausible principles, namely:

---

<sup>19</sup> Given that  $\phi$ -ing is a necessary means to [ $\phi$ -ing and  $\psi$ -ing], *distribution* also follows from *necessary means transmission*. *Distribution* is famously denied by actualists, such as Jackson and Pargetter (1986). I reject actualism as a theory about the deliberative 'ought' and discuss these issues in more detail in my paper "Instrumental normativity" (unpublished).

*Agglomeration*: If A ought to  $\phi$ , and A ought to  $\psi$ , then A ought to [ $\phi$  and  $\psi$ ].

'Ought' implies 'can': If A ought to  $\phi$ , then it is possible for A to  $\phi$ .<sup>20</sup>

Second, recall that we are concerned with the sense of 'ought' in which believing one ought to  $\phi$  rationally commits one to intending to  $\phi$ . Now, if *joint satisfiability* were false, agents could truly believe that they ought to perform actions that they believe to be incompatible. Consequently, they could be rationally committed to intending actions that they truly believe to be incompatible – simply by believing the truth about what they ought to do. Moreover, since it is itself irrational to intend actions one believes to be incompatible, it follows that agents in such situations would be necessarily irrational just because they believe the truth. This implication, however, seems unacceptable.

Third, recall that the 'ought' in question is supposed to *conclude* deliberation. *Joint satisfiability* tells us that if you ought to, all things considered, do one thing, and you ought to, all things considered, do some other thing, then it must be possible for you to do both of these things. As I see it, the deliberative 'ought' is that conception of 'ought' of which we may say with certainty that it renders this claim true. If the 'oughts' in question were to require practically incompatible responses, then they could not conclude deliberation, for we would still need an answer to the question of what to do. Note that this argument does not presume that there is always an answer to the question of what to do. The point is merely that an 'ought'-statement that allows for the violation of *joint satisfiability* could *not* count as an answer to that question, and thus could not be a statement about the deliberative 'ought'.

---

<sup>20</sup> See Williams (1965, 117–123). Williams famously rejects *agglomeration* for the *moral* 'ought' in order to make room for the possibility of tragic moral dilemmas. However, we are concerned here with the overall 'ought' of deliberation, and with respect to this sense of 'ought', Williams clearly accepts *joint satisfiability*: "The practical *ought* is to be taken to be equivalent to the 'all-in' or 'conclusive' answer to the question 'What ought I to do?', and an *ought* which has that role will have a property which we might call that of being 'exclusive': if I ought to do *X* and I ought to do *Y*, then it must be possible that I can do both *X* and *Y*" (Williams 1980, 119; see also Williams 1965, 123–24). (By a *tragic* moral dilemma I mean a dilemma that involves moral obligations to perform incompatible actions. Moral dilemmas need not be tragic in this sense, they might involve only equally strong or incommensurable moral reasons for incompatible actions. Only tragic dilemmas are incompatible with the joint satisfiability of moral obligations.)

I conclude that a commitment to denying *joint satisfiability* constitutes a *reductio* of any argument about the deliberative ‘ought’. Since taking (5) to be compatible with (1) and (2) without taking (8) to be compatible with (1) and (2) is *ad hoc*, and taking (8) to be compatible with (1) and (2) violates *joint satisfiability*, we should reject Broome’s claim that (5) is compatible with (1) and (2).

Fortunately, there are ways to reconcile the basic intuitions about Broome’s example while maintaining that (5) is incompatible with (1) and (2). First, we could reject (2) as a plausible interpretation of the ‘ought’ that it is supposed to capture. There seems indeed no reason to prefer the wide-scope reading suggested in (2) to the following narrow-scope claim:

(2)\* If you exercise hard every day, then you ought to eat heartily.

So understood, (1) and (2)\* are compatible both with (5) and with (8), but neither (3) nor (7) are entailed, via *normative detachment* or any of the other principles mentioned. Alternatively, if the wide-scope interpretation of (2) is simply stipulated as part of the example, we should insist that (5) and (8) should likewise be interpreted in a wide-scope fashion:

(5)\* You are permitted (it is not the case that you ought not) to make sure that if you do not exercise hard every day, then you do not eat heartily.

(8)\* You ought to make sure that if you do not exercise hard every day, then you do not eat heartily.

That is, we should understand the claims at issue as wide-scope contrary-to-duty imperatives (or permissions).<sup>21</sup> This captures the intuition that there is something right about *not* eating heartily if you do not exercise, even if (1) and (2) are assumed. But since such wide-scope claims do not allow for *factual detachment*, neither (6) nor (9)

---

<sup>21</sup> Chisholm (1963) argues that contrary-to-duty imperatives take narrow rather than wide scope. But his argument is based on *inheritance*, which I rejected in Chapter 3.1 for requirements, and reject for ‘oughts’ as well. Broome (2013, 121–22) rejects *inheritance* as well.

follow, and the conflict with the inference validated by *normative detachment* (1)-(3) disappears.

To sum up, Broome's case works as a counterexample against *normative detachment* only if we interpret (2) as having wide scope and (5) as having narrow scope. But since significant independent problems can be avoided by interpreting either (2) as having narrow scope or (5) as having wide scope, such an alternative interpretation is clearly preferable. Consequently, the example provides no reason for rejecting *normative detachment*.

#### *Implications for the wide-scope account*

I shall now turn to the question of whether *normative detachment* provides problems for the normative wide-scope interpretation of structural requirements of rationality. I shall begin with the evidence-belief requirement:

(C)<sub>12</sub> A ought to [not believe that A has sufficient evidence that p *or* believe that p].

Suppose that you ought to believe that you have sufficient evidence that p. From (C)<sub>12</sub> and *normative detachment*, it follows that you ought to believe that p. Hence, the wide-scope account entails that whenever you ought to believe that you have sufficient evidence for a proposition, you also ought to believe this proposition itself. On a roughly evidentialist picture of what we ought to believe, this seems plausible. If our evidence requires us to believe that we have sufficient evidence for p, it plausibly also requires us to believe p. Evidence for evidence is evidence – at least if the evidence in question is possessed by one and the same person.<sup>22</sup> Hence, (C)<sub>12</sub> does not seem to allow for detachment of implausible normative conclusions.

As Stephen Finlay has argued, things might look more problematic with respect to the self-governance requirement:<sup>23</sup>

<sup>22</sup> What I say here is thus compatible with Fitelson's (2011) arguments to the effect that evidence that *another* person has evidence for p is not necessarily evidence that p. For further discussion of the assumption that sufficient evidence for sufficient evidence for p is sufficient evidence for p, see Chapter 9.5.

<sup>23</sup> See Finlay (2010, 71).

(A)<sub>12</sub> A ought to [not believe that she ought to  $\phi$  or intend to  $\phi$ ].

For (A)<sub>12</sub>, together with *normative detachment*, entails the contentious claim that whenever you ought to believe that you ought to  $\phi$ , you ought to intend to  $\phi$ . On an evidentialist picture of what we ought to believe, this suggests that whenever our evidence requires us to believe that we ought to do something, this belief (or at least the belief about the corresponding intention) must be true. In contrast to Finlay, however, I do not find this untenable; indeed, I think we are forced to draw this conclusion for independent reasons. It follows from the evidence-relative view of reasons and ‘oughts’ that I defend in Chapter 8.<sup>24</sup> So while the implication of (A)<sub>12</sub> seems contentious, I think it can be justified on independent grounds.<sup>25</sup> It is nevertheless worth mentioning that wide scope advocates seem committed to it.

Similarly, I think that if you ought to intend to  $\phi$ , and you ought to believe that  $\psi$ -ing is a necessary means for  $\phi$ -ing, it is plausible that you ought to intend to  $\psi$ . Of course, our evidence might require us to believe that  $\psi$ -ing is a necessary means for  $\phi$ -ing without this actually being the case. But given the evidence-relative account of ‘oughts’ that I shall argue for, this does not threaten the inference suggested by the instrumental requirement (B)<sub>12</sub> and *normative detachment*.<sup>26</sup>

Finally, let us consider how the normative wide-scope *modus ponens* requirement (D)<sub>12</sub> harmonises with *normative detachment*. Is it true that we ought to believe q

<sup>24</sup> More exactly, it follows from the *evidence principle* defended in Chapter 7.8 together with the evidence-relative view defended in Chapter 8. Interestingly, Gibbons (2009, 171–73) in effect argues for the claim that we ought to  $\phi$  if we ought to believe that we ought  $\phi$  by assuming (A)<sub>12</sub> and *normative detachment*. Finlay’s *modus tollens* is Gibbons’s *modus ponens* here. I accept neither of these arguments. Finlay’s argument presupposes that Gibbons’s conclusion is false, but I argue in Chapters 7 and 8 that it is (roughly) true for reasons independent of (A)<sub>12</sub> and *normative detachment*. Gibbons’s argument, in contrast, presupposes (A)<sub>12</sub>, which is rejected in Chapters 4.6–6.6.

<sup>25</sup> As the discussion in Chapter 8 makes clear, details about time matter. What is true is that if you ought (at t) to believe at t that you ought (at t) to  $\phi$  at t (and certain other conditions about the relation between ‘ought to act’ and ‘ought to intend’ discussed in Chapter 7.9 are satisfied), then you ought (at t) to intend at t to  $\phi$  at t. This connection can break down if different points of time are involved. This is why Finlay is wrong to claim that (A)<sub>12</sub> (given *normative detachment*) “implausibly entails the impossibility of having false but justified normative beliefs” (2010, 71). This claim is an overgeneralisation not only because there are normative beliefs other than beliefs about what one ought to do, but also because it remains possible to have false justified beliefs about what one ought (now) to do at a later time, or about what one ought (at a later time) to do (at that time or any time after that). See Chapter 8 for details.

<sup>26</sup> I defend this conclusion for reasons independent of (B)<sub>12</sub> and *normative detachment* in Chapter 10.1. See the discussion on *evidence-based transmission*.

whenever we ought to believe  $p$  and we ought to believe  $p \rightarrow q$ ? According to an evidentialist account of 'ought to believe', this principle can be true only if whenever the evidence requires us to believe  $p$  and to believe  $p \rightarrow q$ , it also requires us to believe  $q$ . This might seem plausible at first sight: in normal cases, sufficient evidence for  $p$  and sufficient evidence for  $p \rightarrow q$  constitute sufficient evidence for  $q$  as well. However, there seem to be exceptions to this rule, at least if we assume (as seems plausible) that we can have sufficient evidence for a proposition that has an evidential probability below 1.

The problem is that probability does not necessarily transmit to logical implications. The probabilities of  $p$  and  $p \rightarrow q$  might each be slightly higher than the probability of  $q$ . Consequently, it seems possible that the evidential probabilities are such that we ought to believe each  $p$  and  $p \rightarrow q$ , but not  $q$ . This is also the source of the much-discussed lottery and preface paradoxes. I discuss these issues at greater length in Chapter 9.6, and can here do no more than to summarise the results of that discussion. As I shall argue, *if* probability considerations really do undermine the claim that we ought to believe  $q$  whenever we ought to believe  $p$  and believe  $p \rightarrow q$ , then the very same considerations likewise undermine the claim that it is necessarily irrational to believe  $p$ , believe  $p \rightarrow q$ , but not believe  $q$ , i.e. claim (D). I shall also defend a more fine-grained account of *modus ponens* irrationality, which takes into account credences and evidential probabilities, and which is immune to the objection based on the probability considerations.

It follows that *normative detachment* poses a problem for the normative wide-scope requirement (D)<sub>12</sub> only if this requirement is independently questionable because it does not correspond to a tenable structural irrationality claim in the first place. We can leave open at this point whether this is in fact so, as I am inclined to think. The point is that *if* there is a problem, then the wide-scope account is not committed to (D)<sub>12</sub>, and cannot be rejected on the ground that (D)<sub>12</sub> allows for detachment of unacceptable conclusions.

#### 4.6 Wide-scope and *liberal transmission*

We have seen no sufficient reason to think that the wide-scope account allows for detachment of unacceptable normative conclusions via *normative detachment*. It has been argued that a further principle also raises problems for the wide-scope strategy, namely:

*Liberal transmission*: If A has an intrinsic reason to  $\phi$ , and  $\psi$ -ing is a means for A to  $\phi$ , then A has a reason to  $\psi$ .<sup>27</sup>

I shall start with some clarificatory remarks about this principle, before discussing its implications for the wide-scope account. First, I employ a broad conception of ‘means’, according to which everything that helps to bring about a response can count as a means. I follow Niko Kolodny here, who also emphasises that helping to bring about need not be *causing*, but may also be constituting, or preventing something that would prevent a response.<sup>28</sup>

Second, I will call a reason that is explained by the fact that the favoured action is a means to something else we have reason to do an *instrumental reason*; a reason that is not explained in this way an *intrinsic reason*; and a reason that is referred to in the explanation of an instrumental reason a *source reason*. Given this definition, a source reason (but not an intrinsic reason) might itself be an instrumental reason. However, according to the present formulation of *liberal transmission*, only intrinsic reasons can be source reasons for instrumental reasons generated by this principle. This restriction is

---

<sup>27</sup> Compare: “One has reason to take the means to what one has ultimate reason to do” (Bedke 2009, 678). “People have reason to do what will bring them into conformity with reasons which apply to them” (Raz 2005b, 3). “If you have a reason to do A and doing B is a sufficient means to doing A, you have a reason to do B” (Way 2010a, 224). I borrow the name “liberal transmission” from Rippon (2011, 6–7), who uses it for a related principle that is entailed by the one that I here refer to by that name. Raz seems to presuppose *liberal transmission*, or some similar principle, in his argument against the wide-scope account (cf. Raz 2005a, 11–14), but *liberal transmission* should be distinguished from his “facilitative principle” (see Raz 2005a, 5–6; and, for a more accurate statement, Raz 2011, 148), which is much more restrictive. Raz (2005a, 13, n. 18) notes himself that the inference on which his argument against the wide-scope account relies is not entailed by the facilitative principle.

<sup>28</sup> See Kolodny (forthcoming, §1).

necessary in order to avoid what Matthew S. Bedke calls “the problem of subversion”<sup>29</sup>, and which can be illustrated by the following example. Suppose that I own a lottery ticket and have reason to further raise my chances to win the lottery.<sup>30</sup> Assume, for the sake of the argument, that this reason is intrinsic. A means to raising my chances is to buy a further lottery ticket, so *liberal transmission* entails that I have reason to do this. But now a means to buying the second lottery ticket is to sell my first one and then use the money to buy the other ticket. Certainly, however, I have no reason to sell my first lottery ticket in order to buy another one – doing this would be utterly pointless. It is only because the source reasons are restricted to intrinsic reasons that *liberal transmission* does not have this implication. The reason to buy a further lottery ticket is instrumental; it derives from the reason to raise one’s chances to win. And since selling the first ticket is no means to what I have intrinsic reason to do (namely raising my chances to win), *liberal transmission* does not entail that I have reason to sell the ticket.

Third, some authors suggest further restrictions. In particular, sometimes the source reasons of *liberal transmission* are restricted to sufficient or undefeated reasons, and sometimes the means mentioned are restricted to sufficient means.<sup>31</sup> Both these restrictions strike me as unwarranted, given the general rationale for *liberal transmission*. As far as I can see, the best arguments for a liberal principle of instrumental transmission favour a principle that is not restricted in this way, and this is why I do not include these restrictions in my formula. Consider, for example, Raz’s point that intentionally taking a means to an action one has intrinsic reason to perform amounts to acting for a reason.<sup>32</sup> This consideration applies independently of whether the means in question is sufficient, and even independently of whether the source reason is defeated. So if this argument succeeds, then it establishes a principle that is not restricted in the

<sup>29</sup> See Bedke (2009, 679, n. 12), who credits the point to Dreier.

<sup>30</sup> Rippon (2011, 17) puts forward a similar example to argue against *liberal transmission*; he does not consider the possibility of restricting the antecedent reasons to intrinsic ones.

<sup>31</sup> Raz’s *facilitative principle* (Raz 2005a, 5–6; restated in Raz 2011, 148) restricts source reasons to undefeated reasons, and Way’s *means-end transmission principle* (Way 2010a, 224) applies only to sufficient means (contrast, however, the transmission principle in Way 2012, 494, which also applies to insufficient means). Raz’s general position on this question is not entirely clear. On the one hand, he states: “there is no reason to facilitate conformity with a defeated reason” (2011, 145). On the other hand, the inference that he uses to argue against the wide-scope account of instrumental rationality (1B)-(3B) in Raz (2005a, 12; and 2011, 152) relies on a version of *liberal transmission* that does not restrict source reasons to undefeated reasons.

<sup>32</sup> Compare Raz (2005a, 8–9).

ways suggested. Having said this, I believe that the argument of this section is independent of this question, and would apply to more restricted versions of *liberal transmission* as well.

Finally, I will understand *liberal transmission* such that the instrumental reasons that it generates are cancelled once A conforms to the relevant intrinsic reason. I believe this is already reflected in the above formulation, for once A conforms to the intrinsic reason to  $\phi$ , what have been means for A to  $\phi$  before are no longer means for A to  $\phi$ . It follows that *liberal transmission* does not license the generation of instrumental reasons from reasons one already conforms to.<sup>33</sup>

So understood, *liberal transmission* is at least *prima facie* plausible: it seems that taking means – and not just necessary means – to what we have reason to do, is something we have reason to do. Yet *liberal transmission* is not uncontroversial. Opponents point out that it has counterintuitive implications in cases of very ineffective, immoral or imprudent means. If I have a reason to avoid feeling hungry, does it really follow that I have a reason to kill myself, even if I could achieve the same result by eating a sandwich instead?<sup>34</sup> Proponents of *liberal transmission* reply that instrumental reasons to take inefficient, immoral, or imprudent means will be massively outweighed, and they provide pragmatic explanations of why it could seem to us that they do not exist.<sup>35</sup>

An assessment of these and other arguments for and against *liberal transmission* would go beyond the scope of this thesis; I have to leave this for another occasion. What seems clear is that there needs to be some kind of transmission principle that goes over and above means that are necessary. Clearly, we sometimes have reasons to take means to something we have reasons to do even if the means is not necessary. Opponents of *liberal transmission* therefore face the challenge to come up with an alter-

---

<sup>33</sup> I take it that this is also what Raz has in mind when he states that “we have reason to perform any one (*but only one*)” of the actions that facilitate conformity with the source reason (Raz 2005a, 5, my emphasis). As Raz’s response to Broome (2005b, 6–8) makes clear, he really intends this statement to mean that we have a reason to perform *each* of the facilitative actions, though we do not have a reason to perform more than one that is sufficient for conforming to the source reason (Raz 2005b, 3, n. 8; see also Rippon 2011, 5, n. 13). This denies agglomeration for reasons (‘If A has a reason to  $\phi$ , and A has a reason to  $\psi$ , then A has a reason to [ $\phi$  and  $\psi$ ]). However, reasons agglomeration is implausible anyway. For example, if you have a reason to  $\phi$  and a reason not to  $\phi$ , it does not seem to follow that you also have a reason to [ $\phi$  and not  $\phi$ ].

<sup>34</sup> This is Broome’s (2005b, 7) counterexample to *liberal transmission*. For further examples of this sort, see also Rippon (2011, 17).

<sup>35</sup> See e.g. Bedke (2009, 684–685); Raz (2005b, 3–4); Schroeder (2005b, 7–9; 2007a, 92–97).

native to it. In any case, I shall here not further discuss whether *liberal transmission* is true. Rather, supposing that there is at least some reason to believe that it is true, I shall directly discuss its implications for the wide-scope account.

*Implications for the wide-scope account*

Whenever you have a disjunctive, or wide-scope, reason to  $[\phi \text{ or } \psi]$ , both  $\phi$ -ing and  $\psi$ -ing are sufficient means to conform to this reason. Apparently, then, according to *liberal transmission*, every reason to give a disjunctive response entails reasons to give *each* of the responses that are part of the disjunction. It has been argued that this creates a problem for the normative wide-scope account, which is even worse than the bootstrapping problem. It will suffice to examine this problem by means of the self-governance requirement. Recall:

(A)<sub>12</sub> A ought to [not believe that she ought to  $\phi$  or intend to  $\phi$ ].

Since ‘oughts’ imply reasons, (A)<sub>12</sub> entails that A has a reason to [not believe that she ought to  $\phi$  or intend to  $\phi$ ]. Since intending to  $\phi$  is a means to conforming with that reason, (A)<sub>12</sub> seems to entail that A has a reason to intend to  $\phi$ , for *any* action-type  $\phi$ -ing. Apparently, we do not even need to bootstrap such reasons into existence by adopting some attitude. If *liberal transmission* is correct, then (A)<sub>12</sub> entails that we already have, independently of any of our attitudes, reason to intend *every* possible action. This is what Simon Rippon calls the “All Possible Ends Problem”.<sup>36</sup> This conclusion must be false. No doubt, it is at least *sometimes* correct to say “you have no reason to intend this”.

This argument is overhasty, however. Recall the point that *liberal transmission* does not license the generation of instrumental reasons from reasons one already conforms to. If you do not believe that you ought to  $\phi$ , then intending to  $\phi$  is not a means for you to satisfy (A)<sub>12</sub>, for you already satisfy it. Since for the vast majority of possible

---

<sup>36</sup> Rippon (2011, 8). Broome (2005b, 5–6) first notes this implication in his reply to Raz (2005a), and Raz, ascribing it to Rippon, embraces it in the revised version of his original article (see Raz 2011, 152, n. 17).

actions, it is true that we do not believe that we ought to perform them,  $(A)_{12}$  and *liberal transmission* do not actually entail that for each possible action, we have reason to intend it.

They do entail, however, that we have reason to intend to  $\phi$  whenever we *do* believe that we ought to  $\phi$  and do not intend to  $\phi$  – i.e., whenever we violate  $(A)_{12}$  – for in this case intending to  $\phi$  is a means to satisfying  $(A)_{12}$ . This implication is not worse, as some have claimed, but rather less bad than the original bootstrapping problem. It does not follow that we can create reasons to intend to  $\phi$  just by believing that we ought to  $\phi$ , but only that we can create such reasons by believing that we ought to  $\phi$  while not intending to  $\phi$ . However, this implication is still bad enough. Recall one of the cases that motivated the original bootstrapping objection. You believe, for no reason, that you ought to put all your efforts into building a time-machine that will enable you to meet Elvis Presley. Let us suppose further that you have no intention to do that. If *liberal transmission* is true, we can now conclude from  $(A)_{12}$  that you have reason to intend to put all your efforts into building the time-machine. In fact, however, you might not have any such reason. It is implausible that just by adopting a crazy belief about what you ought to do, and then akratically not forming the corresponding intention, you can make it the case that you have reason to have that intention. So if *liberal transmission* is true, then the wide-scope strategy fails to block the detachment of unacceptable normative conclusions.

A possible reply on behalf of the wide-scope account is to say that the reasons implied by  $(A)_{12}$ - $(D)_{12}$  are instrumental rather than intrinsic reasons, which would mean that *liberal transmission* – applying to intrinsic reasons only – does not apply to them. I have not so far discussed the question of what the reason to satisfy structural requirements of rationality could be and will not do so until the next chapter. There is good reason to believe, however, that even if we grant that the relevant reason is instrumental, this does not help the wide-scope account. Recall that the restriction to *intrinsic* reasons in the antecedent of *liberal transmission* was motivated by the fact that some means that serve the satisfaction of an instrumental reason do not serve the satisfaction of the intrinsic reason it derives from. Recall the lottery case: selling lottery

ticket 1 serves the satisfaction of your instrumental reason to buy lottery ticket 2, but it does not serve the satisfaction of your reason to raise your chances in the lottery. The problem is that it is difficult to see how any intrinsic reason could give rise to an instrumental reason for a disjunctive response, while at the same time responding in one of the two ways that would constitute such a response failed to be a means to satisfying that intrinsic reason. In such a case, we seem always warranted to conclude that the instrumental reason was not disjunctive after all. Hence, I believe that an instrumental account of the normativity of structural requirements of rationality does not help advocates of the wide-scope account out of the detachment problems that *liberal transmission* poses for their account. However, there is a more promising reply available, which is worth considering in more detail.

*Wide-scope, instrumental transmission, and the object-given/state-given distinction for reasons for attitudes*

In defence of the wide-scope account of the instrumental requirement, Jonathan Way argues that principles of instrumental transmission, while applying to reasons for action and state-given reasons for attitudes, do not apply to object-given reasons for attitudes. Moreover, he argues that the reason for satisfying the instrumental requirement is object-given.<sup>37</sup> If both of these assumptions are true, and if Way's account can be extended to other structural requirements of rationality, this would provide a rationale for rejecting the application of *liberal transmission* to  $(A)_{12}$ - $(D)_{12}$ .

I shall grant Way's claim that transmission principles do not apply to object-given reasons. This claim does indeed seem plausible. Recall that a reason for a mental state is object-given if it is provided by facts that bear on the state's *object*, while it is state-given if it is provided by properties of the state itself (typically benefits of having it).<sup>38</sup> First, it seems questionable that object-given reasons give rise to instrumental

<sup>37</sup> See Way (2010a, esp. §4).

<sup>38</sup> Way holds that the same is true for reasons to have combinations of attitude-states; these reasons, too, are object-given if they are provided by facts that bear on the objects of the relevant attitude-states, and state-given if they are provided by properties of having the combination of attitude-states (see Way 2010a, 218). I find it less clear that a reason for a combination of attitudes could be object-given at all (see Chapter 10.4 on this point).

reasons at all. Suppose your evidence gives you reason to believe that it will rain tomorrow, and suppose that your putting on the radio would help to bring it about that you have this belief. Does it follow that you have a reason to put on the radio? Plausibly not. Whether you have a reason to put on the radio depends solely on practical considerations. You may have an epistemic reason to believe that it will rain, and no reason to put on the radio at all. Second, it seems clear that object-given reasons can never be instrumental reasons themselves, for an instrumental reason for an attitude would depend on the instrumental relation in which this attitude stands to some other response, and it would thus depend on properties of the attitude rather than its object. So it seems that object-given reasons are never outputs, and plausibly also not inputs of instrumental transmission principles. Contrast state-given reasons: if the fact that you have been promised a reward to believe that it will rain really is a reason to believe that it will rain, and turning on the radio will causally effect your having this belief, then it is plausible to think that this gives you a reason to turn on the radio. So state-given reasons (if there are any) can be inputs of transmission principles, and as a matter of course, they can also figure as outputs of such principles.

The question then turns on whether the reason that is supposed to back up claims (A)<sub>12</sub>-(D)<sub>12</sub> is object-given. I shall now argue that for general considerations concerning the form of wide-scope requirements, a reason to satisfy such a requirement cannot be object-given. A brief discussion of Way's account of the instrumental requirement will help to illustrate this point. To begin with, recall the standard wide-scope account of the instrumental requirement of rationality:

(B)<sub>5</sub> Rationality requires A to [not intend to  $\phi$  *or* not believe that  $\psi$ -ing is a necessary means for  $\phi$ -ing *or* intend to  $\psi$ ].

Way rejects this account in favour of what might be called a "medium scope" approach:

(B)<sub>13</sub> *Instrumental requirement, medium-scope*: If A believes that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then rationality requires A to [not intend to  $\phi$  *or* intend to  $\psi$ ].

So according to Way, the means/end-belief is part of the *condition* of the requirement rather than part of its content. It follows that the reason to conform to this requirement must be a reason to [not intend to  $\phi$  or intend to  $\psi$ ]. Way claims that the reason in question is the fact that  $\psi$ -ing is a necessary means to  $\phi$ -ing itself. Since this fact concerns the objects of the relevant attitude-states (namely the actions of  $\phi$ -ing and  $\psi$ -ing) rather than the states themselves, Way maintains that this reason is object-given.<sup>39</sup>

The first thing to note about this approach is that a similar move is simply not available for other structural requirements of rationality. If, for example, we take the belief-element of the self-governance requirement to be a condition of the requirement rather than a part of its content, we automatically end up with a narrow-scope account.

The second and more important point in our context is that Way's account cannot vindicate the normativity of rationality, not even in the case of the instrumental requirement. One reason for this is that the medium-scope account of the instrumental requirement licenses unacceptable bootstrapping when we assume the normativity of rationality. Consider:

(B)<sub>14</sub> *Normative instrumental requirement, medium-scope*: If A believes that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then A ought to [not intend to  $\phi$  or intend to  $\psi$ ].

Suppose, again, that you believe – for no reason – that in order to impress your date, you need to dress as a parrot and tell a lot of sexist jokes. According to (B)<sub>14</sub>, we have to conclude that you ought either not to intend to impress your date or to intend to dress as a parrot. But plainly, this does not follow: you may well have perfect reasons to intend to impress your date and not to intend to dress as a parrot, and you cannot change this simply by adopting a crazy means/end-belief.

Another reason why Way's account is incapable of vindicating the normativity of rationality is that the reason that according to Way speaks in favour of satisfying the instrumental requirement (namely the fact that  $\psi$ -ing is a necessary means to  $\phi$ -ing) is present only if the agent's means/end-belief is true, while the rational requirement

---

<sup>39</sup> See Way (2010a, §§1-2).

holds independently of this assumption. It follows that if an agent's means/end-belief is false, he might well have no reason at all to be instrumentally rational.

Now Way himself does not aim to vindicate the claim that we necessarily have reasons to be rational; he is a proponent of the mixed view rather than the structuralist view.<sup>40</sup> Yet I am here concerned with the question of whether the structuralist view can be defended against the charge that it allows for detachment of implausible normative conclusions. And with respect to *this* purpose, we can draw two lessons from the discussion of Way's account: First, in order to avoid the bootstrapping objection, a normative wide-scope account has to include the belief-element involved in structural irrationality claims as part of the disjunctive response that is required. Second, the reason to satisfy the requirement cannot be given by the content of the belief, because the requirement holds independently of whether the belief is true or false.

Indeed, we can go further. The reason to satisfy the requirement must also be independent of any *evidence* that counts in favour or against the content of the belief-element, for the requirement also holds independently of the presence of such evidence. Given this, however, it seems clear that reasons to satisfy wide-scope requirements of rationality cannot be object-given. Since all these requirements have a belief-element as part of the required response, and object-given reasons for or against believing something are always provided by evidence for or against the content of the belief, and a reason to satisfy a structural requirement must be independent of evidence for the belief-element involved, any such reason must be state-given. If there is such a reason, it must be provided by properties of the attitude-states in question, for example their coherence or their value for autonomy.<sup>41</sup>

This point will be discussed in more detail in the next two chapters, but I hope that the basic idea is clear. Since the wide-scope account needs to include the belief-element as part of the required response, and since the requirement must hold inde-

---

<sup>40</sup> See Way (2010a, 220). More exactly, Way is a proponent of an *apparent reason account* of the normativity of rationality, which claims that whenever an agent is rationally required to do something, he *would* have reason to do it *if* his beliefs were true. The apparent reasons account has sometimes been presented as a vindicating account of the normativity of rationality; Way himself suggested this interpretation in an earlier article (Way 2009). I argue in Chapter 5.6 that the apparent reasons account is an undermining rather than vindicating account of the normativity of rationality.

<sup>41</sup> See also Kolodny (2005, 551).

pendently of evidence, a reason for satisfying the requirement, which is necessarily in place, cannot be object-given. Now, I have argued in Chapter 1 that there are good reasons to think that all reasons for attitudes must be object-given. So it already follows from what I have said so far that there are no reasons to satisfy wide-scope requirements of rationality. This thesis will be further substantiated in the next chapter. However, the conclusion of *this* section does not rely on the view that there are no state-given reasons. What I have argued here is merely that considerations that call into question that *liberal transmission* can be applied to object-given reasons do not provide any reason to doubt that *liberal transmission* can be applied to wide-scope requirements of rationality. It follows that if *liberal transmission* can be maintained (in a version that does not apply to object-given reasons), then the wide-scope strategy does not succeed in escaping the bootstrapping objection, for the wide-scope account itself licenses the detachment of unacceptable normative conclusions.

#### 4.7 Wide-scope and *necessary means transmission*

In this final section I shall consider how the normative wide-scope account deals with the necessary means transmission principle:

*Necessary means transmission:* If A ought to  $\phi$ , and  $\psi$ -ing is a necessary means for A to  $\phi$ , then A ought to  $\psi$ .<sup>42</sup>

---

<sup>42</sup> Compare: “If you should do *E*, all things considered, and doing *M* is a necessary means to doing *E*, you should do *M*, all things considered, too” (Setiya 2007a, 660). “If *X* objectively ought to do *A*, and to do *A* *X* must do *B*, it follows that *X* ought objectively do *B*” (Schroeder 2009, 234). See also the following similar principles about reasons: “Reasons for me to make something my end are [...] equally reasons for me to take the necessary means to it” (Darwall 1983, 16). “If *X* has an objective reason to do *A* and to do *A* *X* must do *B*, then *X* has an objective reason to do *B* of equal weight to *X*’s objective reason to do *A*” (Schroeder 2009, 245). “If *R* is a practical reason in favor of *X*, *X* is attainable by the agent, and *M* is a necessary means to or necessary constitutive element of *X*, then *R* is a practical reason in favor of *M*” (Bratman 2009a, 424). “If you have a reason to do *A*, and doing *B* is a necessary means to doing *A*, you have a reason to do *B*, which is at least as strong as your reason to do *A*” (Way 2010a, 225). A similar transmission principle seems presupposed by Street’s claim that it is conceptually necessary to believe in accordance with it: “Just as it is constitutive of being a parent that one have a child, so it is constitutive of taking oneself to have conclusive reason to *Y* that one also [...] take oneself to have reason to take what one recognizes to be the necessary means to *Y*” (Street 2008, 228).

Despite its great intuitive appeal, this principle has recently been met with objections.<sup>43</sup> However, I have elsewhere defended *necessary means transmission* against these objections, and provided an independent argument for it.<sup>44</sup> I shall here take it for granted.

Kieran Setiya has argued that there are cases in which this principle allows us to detach unacceptable normative conclusions from normative wide-scope requirements of structural rationality.<sup>45</sup> I shall focus on his argument against the instrumental requirement:

(B)<sub>12</sub> A ought to [not intend to  $\phi$  *or* not believe that  $\psi$ -ing is a necessary means for  $\phi$ -ing *or* intend to  $\psi$ ].

Setiya's example is as follows. Suppose that you intend to smoke and believe that in order to do so, you have to buy a pack of cigarettes. (B)<sub>12</sub> tells you that you ought to either give up one of these attitudes or intend to buy a pack of cigarettes. But now suppose that your intention and means/end-belief are psychologically unalterable; you can do nothing to change these attitudes. It follows that a necessary means to satisfying (B)<sub>12</sub> is to intend to buy a pack of cigarettes. Applying the transmission principle, we reach the conclusion that you ought to intend to buy a pack of cigarettes. But this might not be true. Perhaps you are suffering from asthma, or you have promised someone to stop smoking, so you ought not intend to smoke, nor intend to buy a pack of cigarettes. So as this example seems to show, the normative wide-scope instrumental requirement allows for the detachment of unacceptable normative conclusions. And analogous cases are easily constructed for other structural requirements.<sup>46</sup>

Michael Bratman offers an interesting response to this challenge, which takes off from a different question, namely the question of what, given that we have reason to satisfy structural requirements of rationality, this reason is. According to Bratman, it is

<sup>43</sup> See Broome (2013, 126), Heuer (2010), Kolodny (forthcoming, §§2-3), Raz (2005a, §1).

<sup>44</sup> See my "Instrumental normativity" (unpublished).

<sup>45</sup> See Setiya (2007a, 654–663).

<sup>46</sup> The same conclusion follows from Greenspan's principle that a wide-scope 'ought' to make a conditional true allows detachment of an 'ought' to make the consequent true if the antecedent proposition describes a fact that is unalterable by the agent who is subject to the 'ought' (cf. Greenspan 1975, 265). This principle is entailed by *necessary means transmission*, but it might also seem independently plausible.

a reason of autonomy or “self-governance” as he calls it.<sup>47</sup> The idea is, roughly, that we have reason to govern ourselves, that conformity with structural requirements of rationality is a necessary constitutive element of self-governance, and that we therefore have reason to satisfy such requirements.<sup>48</sup> I shall discuss this position in the next chapter in more detail, but will assume it here for the sake of the argument. Bratman’s reply to Setiya is that the reason of self-governance is itself conditional on the agent’s capacity to govern herself in the domain in which the rational requirement applies, and that in Setiya’s example this condition is not satisfied.<sup>49</sup> When an agent is in a state of psychological compulsion such that she is unable to modify her attitudes in the light of her reflection, she is not capable of governing herself in the relevant domain, and therefore has no reason to satisfy the requirement in question. Hence as Bratman conceives of the normativity of structural wide-scope requirements of rationality, no conclusions to the effect that we have reason to take the means to any unalterable intention we have can be detached.

One might wonder whether Bratman’s account does not actually sacrifice the idea that rationality is necessarily normative by making the reasons to satisfy rational requirements conditional on further factors. And indeed Bratman seems to think that the rational requirement to be means/end-coherent remains in place even if the reason to satisfy it vanishes in case the agent is incapable of self-governance. Arguably, however, Bratman’s argument applies not only to reasons, but to rational requirements as well. The capacity to modify one’s attitudes in the light of reflection seems not only a precondition for an agent’s being subject to a reason to be structurally rational in a given domain, but a precondition for being subject to the corresponding requirement of rationality as well. Admittedly, it is not uncommon to label incoherent attitudes resulting from psychological compulsions as “irrational”, but this shows only that the boundaries between psychological compulsion and irrationality are not sharp in ordinary

---

<sup>47</sup> Note that Bratman’s conception of self-governance is not primarily concerned with what I have called the *self-governance requirement*, or is at least much broader than that.

<sup>48</sup> See Bratman (2009a, esp. 430–434). Bratman focuses solely on *practical* requirements of rationality, and it is far from obvious that his strategy can be employed for similar problems that we encounter in the case of theoretical requirements. Since I am going to argue that his response ultimately fails to rescue the wide-scope account anyway, I ignore this point here.

<sup>49</sup> See Bratman (2009a, 434–435).

language. The notion of irrationality we are interested in when asking for the normativity of rationality – the one that is associated with legitimate criticism – does, I think, require the capacity to modify one’s attitudes in the light of reflection, and thus the absence of compulsion.<sup>50</sup> So understood, Bratman’s response poses no threat for the claim that rational requirements are necessarily normative.

It does, however, fail to rescue the normative wide-scope account from detachment problems for a different reason. The problem is that the fact that it is impossible for an agent to modify one of his attitudes may not be due to psychological compulsion or lack of capacity for self-governance, as shown by a thought experiment put forward by John Brunero:<sup>51</sup> Suppose that Smith intends to smoke, believes that buying a pack of cigarettes is necessary in order to do so, but does not intend to buy a pack of cigarettes. Smith is capable of revising his intention to smoke as well as forming the intention to buy cigarettes; he has the relevant capacity of self-governance necessary in order to count as violating a reason to be instrumentally rational according to Bratman. Next, suppose that Jones is exactly like Smith in every psychological respect. Jones differs from Smith only in that the following counterfactual claim is true about him: were Jones to start revising his intention to smoke, someone (perhaps a hypnotist or a neuroscientist) would prevent this from happening. Since Smith and Jones share the same psychology, and are equally guided by their deliberation, it seems that since Smith is self-governed so is Jones. Hence they are both subject to the instrumental requirement. However, Jones can satisfy the wide-scope requirement in only one way: by intending to buy a pack of cigarettes. It follows from *necessary means transmission* that we can detach an ‘ought’ (or reason)<sup>52</sup> to buy cigarettes. But this is an unacceptable result, for we would have to draw similar conclusions in cases in which agents intend to

---

<sup>50</sup> It is interesting, especially in this context, that Setiya seems to share this standpoint: “Irrationality in the narrow sense stands to defects of reason as moral culpability stands to moral wrongdoing; it is circumscribed by our capacities. To say that someone is irrational, in this sense, is to ascribe to them a failure of reason that they could legitimately have been expected to avoid” (Setiya 2007a, 660, n. 25).

<sup>51</sup> See Brunero (2010b, 584–587).

<sup>52</sup> Bratman purports to defend only the claim that we necessarily have a *pro tanto* reason to be rational; his discussion therefore focuses on reasons rather than ‘oughts’. However, it seems clear that if *necessary means transmission* is true for ‘oughts’, then a similar principle is true for reasons as well; see my “Instrumental normativity” (unpublished, §4) for an argument to this effect. Bratman (2009a, 424) explicitly embraces a necessary means transmission principle for reasons.

commit mass murder or other cruel, insane or stupid actions. Bratman's suggestion that the reason to be means/end-coherent is conditional on the agent's capacity to govern herself is of no help here.

In defence of the normative wide-scope account, one might opt for the following principle:

*Disjunctive 'ought' implies conjunction of 'cans'*: If A ought to [ $\phi$  or  $\psi$ ], then it is possible for A to  $\phi$ , and it is possible for A to  $\psi$ .

This would mean that (B)<sub>12</sub> applies only to Smith but not to Jones, and would thus block the conclusion that Jones ought to buy a pack of cigarettes. I find it a plausible principle, but I think it is of no help for the normative wide-scope account. The problem is that it undermines the idea that rationality is necessarily normative rather than supporting it. The move we considered before is not available here: that agents both can revise an intention for the end and can form an intention to take the means may be a precondition of it being the case that they *ought* to do either. But it is no precondition of agents' being irrational in not intending the means to an end they intend, *not even in the relevant sense of 'irrational' which implies criticisability*. This is shown by the case of Jones. The fact that another person *would* have intervened had he decided otherwise makes no difference for the fact that, as things happened, he is just as responsible for his condition as Smith, and legitimately to be criticised as irrational. If the wide-scope 'ought' applies to him, we can detach unacceptable normative conclusions. But if the wide-scope 'ought' does not apply to him, then we either need a different account of the normativity of rationality or must conclude that rationality is not, after all, necessarily normative.<sup>53</sup>

---

<sup>53</sup> Given that it is true that *disjunctive 'ought' implies conjunction of 'cans'*, it is, of course, open to wide-scope advocates to defend a restricted account of the normativity of rationality along the lines of the following claim: If rationality requires A to  $\phi$ , and if – in case  $\phi$ -ing is a disjunctive response to [ $\phi_1$  or ... or  $\phi_n$ ] – it is possible for A to  $\phi_1$ , ..., and possible for A to  $\phi_n$ , then A ought to  $\phi$ . This would certainly be an interesting result, but it would fail to meet several desiderata that motivate belief in the normativity of rationality in the first place. Such a view could not generally account for the criticism involved in irrationality ascriptions and would still be left with the obscure notion of a non-normative requirement of rationality.

#### 4.8 Summary

In this chapter I have been concerned with a series of difficulties for the structuralist view, which claims that there are structural requirements of rationality and that these requirements are normative. The general problem is that such requirements seem to allow for the detachment of unacceptable normative conclusions. I have argued that if structural requirements of rationality take narrow scope, they cannot be normative because this would allow for unacceptable forms of bootstrapping. So the only way to defend the structuralist view is by taking the requirements at issue to have wide scope. But I have argued that wide scope requirements, too, allow for problematic instances of detachment via independently plausible bridge principles. While conclusions to be drawn from *normative detachment* seem defensible, the implications that follow from *liberal transmission* amount to implausible bootstrapping. Last but not least, I have argued that the wide-scope account yields unacceptable conclusions because of *necessary means transmission*. These are significant problems for the structuralist view, for it seems that on both possible interpretations of structural requirements of rationality, these requirements cannot be understood as normative. In the next chapter, I will substantiate this conclusion by discussing a second kind of problem that applies to the structuralist view independently of questions of detachment.



## The why-be-rational challenge

---

In the last chapter, I have argued that the assumption that structural requirements of rationality are normative leads to various detachment problems. This chapter focuses on a further difficulty for the structuralist view: the challenge to identify a reason that counts in favour of conforming to rational requirements. As far as I can see, there are three possible ways to account for this challenge. The first is to present instrumental reasons to conform to rational requirements (5.1). The second is to argue that rational requirements are themselves reasons (5.2). The third is to give some kind of buck-passing account of rational requirements, according to which such requirements are verdictive statements about reasons that exist independently of them (5.3-5.4). I argue that none of these strategies succeed. Finally, I consider two accounts that have been claimed to explain the normativity of structural rationality without assuming that there are reasons to be rational: the transparency account (5.5), and the apparent reasons account (5.6). I argue that these accounts may be capable of explaining why rationality *appears* to be normative from the first-personal perspective, but that they can neither vindicate this appearance, nor give a plausible error theory for the fact that rationality seems normative from the standpoints of advice and criticism as well.

### 5.1 Are there instrumental reasons to be rational?

Suppose that rationality is normative in either the strict or the contributory sense. Suppose, in other words, that we have decisive or at least *pro tanto* reason to conform to the rational requirements that apply to us. This poses the question of what the reason is – a question that turns out to be quite difficult to answer. I call this the *why-be-rational challenge*, in honour of Niko Kolodny, who first raised this question in his paper “Why be rational?” (2005).

When we seek reasons to conform to the requirements of rationality, two general approaches suggest themselves. One might provide an *instrumental* explanation of the normativity of rationality in terms of reasons that descend from other sources than rationality itself, such as self-interest or morality. Alternatively, one might argue that rational requirements are *in themselves* reasons to conform to them. This section considers the first of these approaches.

According to the instrumental approach, we have reason to satisfy the requirements of rationality because doing so is a means to doing other things we have reason to do. This is a natural idea, but John Broome and Niko Kolodny have both pointed out a general problem for it.<sup>1</sup> Since nothing guarantees that conforming to a rational requirement is an effective means to doing what one otherwise has reason to do *in each particular case*, it also does not follow that we have reason to conform to rational requirements in each particular case. Yet we are looking for a necessary connection between normativity and rationality, not only a regularity.

Suppose, for example, that I have the false but very robust belief that I ought to have another drink. The likeliest way for me to satisfy the self-governance requirement and become rational, then, is by intending to have another drink, and this will in turn very likely cause me to have another drink. But since my belief was false, we can suppose that it is in fact very imprudent of me to have another drink. Hence, in a particular case like this, my reason to be rational cannot be that it is prudent or helps me to do what I ought to do.

It might be argued that although we do not have instrumental reasons to be rational in each particular case, we do have a general reason to acquire and promote the *disposition* to be rational, for being disposed to be rational is in the long run the most effective means to conform to what we have reason to do independently of rationality. But this is not the question at issue. We criticise people for being irrational, not – at least not only – for not being disposed (or not sufficiently promoting our disposition) to be rational. So even if we have general reasons to dispose ourselves to be rational, we

---

<sup>1</sup> See Broome (2007a, 171–5) and Kolodny (2005, 543).

still need reasons to intend and believe rationally in order to capture the point of such criticism.<sup>2</sup>

A more promising line of argument takes it that conforming to rational requirements is a *constitutive* element of something else we have reason to do, such as functioning as an agent or believer, or exercising our autonomy.<sup>3</sup> On the basis of this assumption, it can be argued that the reason to be a functioning or autonomous agent transmits to a reason in favour of conforming to rational requirements, because the latter is a necessary means to the former.

The problem with an argument along these lines is that its premises seem too strong, and its conclusion too weak. Let us begin with the case of agency. First, we have to assume that we necessarily have a reason to function as an agent. Second, we have to assume that conforming to each and every rational requirement is constitutive of functioning as an agent. It seems, however, that the second of these assumptions is true only if the first is understood in a way that begs the question. Let us grant that we necessarily have a reason to function as an agent. This assumption may seem plausible, at least unless we take the notion of ‘functioning as an agent’ to be very demanding. But then the second assumption will ring false. One can certainly function as an agent in a natural sense of ‘agent’ without conforming to each and every rational requirement – we all do! On the other hand, if we take the notion of agency to be so demanding that one ceases to function as an agent whenever one violates a rational requirement, the assumption that we necessarily have reason to be an agent in *this* sense just begs the question. The question “Why be rational?” can then simply be reiterated as the question “Why be an agent in this very ambitious sense?”<sup>4</sup>

---

<sup>2</sup> See also Kolodny’s (2008b) critique of the idea that we have reasons to dispose ourselves to be rational.

<sup>3</sup> Both Davidson (1985, 196–197) and Korsgaard claim that conformity with requirements of rationality is constitutive of acting and believing. However, they do not suppose that there is a primary reason to function as an agent or believer, but seem to think that the normativity of rationality can be grounded in the bare fact that we cannot but act and believe. Cf. Korsgaard (2009b, 32): “the laws of practical reason govern our actions because if we don’t follow them we just aren’t acting, and acting is something that we must do”. For the claim that conformity with rational requirements is constitutive of autonomy, see esp. Bratman (2009a, 430–434).

<sup>4</sup> Enoch (2006) makes a similar point against theories according to which (practical) normativity in general is grounded in constitutive facts about agency.

The prospects of giving such a constitutivist argument look better in the case of autonomy. Drawing on Harry Frankfurt's concern with the question of which of an agent's attitudes have "agential authority"<sup>5</sup> in the sense that they 'speak' or 'stand' for this agent, Michael Bratman develops the following argument:<sup>6</sup> An agent with incoherent attitudes about a certain subject matter has not yet determined where she stands with respect to that subject matter. Unless there is such a place where the agent stands, this agent cannot govern herself with respect to that domain, for self-governance (at least partly) consists in the agent's stance's guiding her attitudes. Therefore, conforming to structural requirements of rationality with respect to a given domain is a constitutive element of exercising self-governance in this domain.

This argument provides a good case for the second assumption of the constitutivist argument. Moreover, it seems to me that it does so without running into the danger of begging the question with respect to the first assumption, namely that we have reason to govern ourselves in this sense. Bratman's notion of self-governance is not simply defined in a way such that conforming to rational requirements is a constitutive element of it (as it appeared to be the case for the demanding sense of agency); it is a natural conception, and one that does not seem to presuppose that the normative question about rationality is already settled, because it is an open, substantive question whether we have reason to govern ourselves in this sense.

The question, then, is whether it is really true that we necessarily have reason to govern ourselves in each and every domain in which it is possible for us to violate rational requirements. And here I have my doubts. Let us assume that the value of autonomy often provides us with reasons to govern ourselves. It still seems that self-governance in each and every domain is neither possible, nor desirable. Note that Bratman's notion of self-governance does not merely require the absence of weakness of will (as what I have called the "self-governance requirement" does), it positively requires that one's attitudes be guided by certain higher-order attitudes, plans and policies, which speak for the agent. And while I agree that we often have reasons to guide

---

<sup>5</sup> This is Bratman's term, see especially Bratman (2001). For Frankfurt's discussion of this question, see especially Frankfurt (1987).

<sup>6</sup> See Bratman (2009a, 430–432).

ourselves in this way, I just find it difficult to see why this should always be the case, and with respect to each and every attitude that is subject to rational requirements. Sometimes we just have to let go.

A further point to note is that agents often lack a determinate stance without committing any irrationality, and that Bratman's argument would yield the conclusion that there is a reason for determining one's stance even in those cases. Consider the case of withholding belief with respect to a certain issue. Withholding belief with respect to  $p$  means that one has not determined where one stands with respect to a certain domain in which rational requirements apply. But this is often the most rational response we can give, and it seems just false to say that we always have a reason to settle the matter whether to believe  $p$  or  $\neg p$  for any proposition  $p$  (consider "the number of stars is equal"). A similar point applies to intentions. It seems clear that we do not have a reason to determine our stance with respect to intending an action *for each possible action*. But if having determined where one stands in a given domain counts as a prerequisite for governing oneself in that domain, it would follow that a reason to govern oneself in each domain does indeed give rise to a reason to determine for each possible action whether to intend it or not. In sum, it seems unclear how Bratman's justification of a reason against incoherence can avoid condemning cases in which it seems intuitively unproblematic, rationally advisable, or even psychologically necessary that agents do not determine where they stand.

Finally, the conclusion that Bratman's argument yields seems just too weak to sufficiently substantiate the normativity of rationality. If we really have a reason to govern ourselves in each particular domain where this is possible, it seems clear that this reason would be just one among others, easily outweighed by other considerations. The simple fact that our psychological resources are limited shows that we will often have strong reasons *not* to govern ourselves in a particular domain. Consequently, the derivative reason to conform to rational requirements would also quite regularly be defeated or outweighed. As a result, Bratman's approach fails to account for the criticism that is associated with irrationality ascriptions, for defeated reasons cannot ground criticism. In order to account for a necessary, or at least very regular, connection

between irrationality and criticisability, we need a reason that is (at least very regularly) decisive.

In sum, the instrumental explanation of the normativity of rationality seems to fail: while it can be agreed that conforming to structural requirements of rationality often helps us to conform to other kinds of reasons, this is not true in each particular case. And while it might be argued that conforming to rational requirements is a constitutive element of functioning as an agent in a very demanding sense, or of exercising autonomy, the assumption that we always have reason to function as an agent in this sense just begs the question, and the assumption that we always have reason to exercise autonomy strikes me as false.

## 5.2 Are rational requirements themselves reasons?

It thus seems that in order to maintain a necessary connection between rationality and reasons, we need to claim that the fact that rationality requires something is *in itself* a reason for doing it.<sup>7</sup> However, as especially Niko Kolodny has emphasised, this view has some serious defects.<sup>8</sup> To begin with, the idea that rational requirements are themselves reasons seems to turn rationality into some kind of fetish. Rational agents do not typically (if ever) act, intend or believe *for the reason that their actions or beliefs are rationally required*. We can underpin Kolodny's point here by considering the case of an agent who satisfies the self-governance requirement by forming an intention to do what she

---

<sup>7</sup> Or, as Broome phrases it, that rationality is “non-derivatively normative” (2013, 204). This is Broome's current position on the matter, although he concedes that he is without an argument for it. It is also a natural interpretation of the views of Southwood (2008, esp. 17–18) and Hussain (unpublished, esp. §5), who both charge Broome and Kolodny with committing the same kind of mistake with respect to rationality that Prichard (1912) thought was involved in asking for an external justification of morality. I believe this charge to be unjustified. Kolodny explicitly addresses the question whether rational requirements are themselves reasons, and he gives good arguments against this view, which Southwood in turn does not seem to address at all. Hussain in fact oscillates between the view that rational requirements are themselves reasons and the view that rational requirements exhibit a unique kind of normativity that is independent of reasons. On the first view, his position is vulnerable to the objection discussed in this section. On the second view, it cannot account for the criticisability claim defended in Chapter 2. For further objections to Hussain's positive proposal that rational requirements are norms that apply only to processes of reasoning, see Southwood (2008, 19–22). For objections to Southwood's positive proposal that rational requirements are genuinely normative “standpoint-relative demands” with “first-personal authority”, see Broome (2008, 96–100) and Coons and Faraci (2010). I find these objections convincing and could do no more than to repeat them here.

<sup>8</sup> Kolodny (2005, 547–51). See also Scanlon (2007, 85–86).

believes she ought to do. It is very natural to think that this agent, at least if she is rational, will form this intention for the same reason for which she formed her belief. Suppose, for example, she believes that she ought to donate money to a charity organisation because this will help people in need. When she now, in accordance with the self-governance requirement, forms the intention to donate money, she will typically do this for the same reason, namely that donating money will help people in need.<sup>9</sup> It would be *very* awkward of her to form this intention *for the reason that rationality requires her to have this intention* or to have coherent attitudes in general. This would amount to treating rationality as a fetish.

This argument does not by itself show that rational requirements cannot be reasons, but it puts a lot of pressure on this idea. If it is very untypical, and indeed inappropriate, to act for certain putative reasons, this calls into question that such reasons exist. But more can be said. As Kolodny convincingly argues, it is not only untypical but often *impossible* for us to give responses we are rationally required to give *for the reason that they are rationally required*, namely in case the response that is called for is a belief. Beliefs, insofar as they are guided by reasons, are guided by evidential considerations – considerations that have to do with the truth of the belief's content.<sup>10</sup> The fact that a belief, or a disjunction of beliefs, is required by a structural requirement of rationality, however, is not an evidential consideration. Consider an agent who believes that  $p$  and that  $p \rightarrow q$ . On the narrow-scope view, rationality requires him to believe  $q$ . But this fact does not itself provide any evidence for  $q$ . What supports the truth of  $q$  may be the *truth* of  $p$  and  $p \rightarrow q$ , or the *evidence* in favour of  $p$  and  $p \rightarrow q$ , or independent evidence in favour of  $q$ , but the structural requirement holds independently of any assumption that we make about these facts. The same point applies if the agent is required to [not-believe  $p$  *or* not-believe  $p \rightarrow q$  *or* believe  $q$ ], as the wide-scope view claims. This fact also does not support the truth or falsity of the

---

<sup>9</sup> Coates (2013), too, stresses that 'enkratic' agents base their intentions and normative beliefs on the same reasons. On his account, the self-government requirement requires one to intend to  $\phi$  for the same reasons for which one believes one ought to  $\phi$ . But note that this does not account for the why-be-rational challenge, as long as we do not know what reasons we have to satisfy *this* requirement.

<sup>10</sup> See e.g. Moran (1988, 143).

content of any of the attitude-states mentioned, and an agent cannot *for this reason* adopt either of them.

So if rational requirements were reasons, they would in many cases constitute reasons to which we could not respond. However, as mentioned in Chapter 1, I take it to be a conceptual truth that if R is a reason for A to  $\phi$ , then it is possible for A to  $\phi$  for that reason. If this is true, then rational requirements cannot generally be reasons in themselves.

Moreover, this point applies to attempts to justify the normativity of rationality by *instrumental reasons* as well. The general idea of such a justification is that we have reason to satisfy requirements of structural rationality because doing so helps to bring about, or is a constitutive element of, our doing something else we have reason to do. Now, all theoretical requirements of rationality require us to adopt certain belief-states, and at least the wide-scope versions of practical requirements of rationality allow us to conform with such requirements by adopting a belief-state.<sup>11</sup> However, the fact that adopting a required state, or collection of states, is a means to doing something else we have reason to do, such as functioning as an agent, acting autonomously, or being prudent, is not an evidential consideration – a reason for which we could adopt a belief-state. So if these instrumental considerations were reasons to satisfy rational requirements, they would also in many cases constitute reasons we could not respond to by giving the responses *for* these reasons. I conclude that, in addition to the worries I have discussed in the foregoing section, both approaches to the why-be-rational challenge fail for the general reason that the alleged reasons they present in favour of satisfying rational requirements are reasons for which we typically do not, and often cannot, give the responses that are required by rationality.

### 5.3 Buck-passing and the normative question

Both approaches to account for the why-be-rational challenge seem to fail: the normativity of rationality can be vindicated neither by instrumental reasons to be rational, nor

---

<sup>11</sup> As before, I use ‘belief-state’ as a general term for both the state in which one believes something and the state in which one does not believe something.

by assuming that rational requirements are reasons in themselves. I believe it is helpful at this point to consider the parallel situation in the case of morality. Here, too, moral discourse presses us to think of moral requirements as normative requirements; and given the reasons claim, we face the challenge to cite reasons for satisfying the requirements of morality, i.e. to account for the question “Why be moral?”. Our options to answer this question seem to be just the same: on the one hand, we can try to come up with an instrumental explanation of the normativity of morality in terms of non-moral reasons; on the other hand, we may claim that moral obligations are reasons in themselves.<sup>12</sup>

As in the case of rationality, both of these options have considerable disadvantages. Again, it seems hopeless to believe that prudential or other non-moral reasons could exhaust the demands that morality makes on us, even aside from the fact that they would provide us with the *wrong kind* of reasons to account for moral motivation as well as moral criticism. However, the alternative answer comes with problems similar to the ones we encountered in the case of rationality. To consider moral requirements as reasons in themselves also seems to involve painting a misguided picture of the moral agent as someone who is concerned with a moral fetish. A moral agent does not perform a moral action, such as helping people in need, for the reason that the action is morally required – she does it because these people need her help.<sup>13</sup>

I shall now turn to a theoretical option to account for the normative question about moral, rational and other possible requirements that I have so far neglected. This option is to regard requirements not as independent truths that may either be reason-providing in themselves or accompanied by instrumental reasons to satisfy them, but as *verdictive* statements, i.e. verdicts to the effect that the balance of reasons of a particular kind, different from the requirement itself, is in favour of the required response, or to the effect that the reasons of a particular kind decisively count in favour of the required

---

<sup>12</sup> A famous example of the first of these options is the theory defended by Gauthier (1986). A natural way to interpret Prichard’s (1912) rejection of an external justification of morality is to understand him as endorsing the second of these options. Broome (2007b, 3–6) explicitly advocates the view that moral requirements provide reasons by themselves. Similarly, both Parfit (2011a, 201) and Scanlon (1998, 148) hold that the fact that an action would be morally wrong provides a reason against this action.

<sup>13</sup> See e.g. Smith (1994, 75–76).

response.<sup>14</sup> This gives us a kind of *buck-passing account* of requirements. Take moral requirements as an example. According to the buck-passing account of moral requirements, stating that morality requires you to help someone in need in a particular situation is stating that the moral reasons that are relevant in the situation count decisively in favour of helping.

How does this answer the question of why we should be moral? It does not really answer it. Rather, it shows that the question is incoherent. To call an action morally required, according to this account, is to say that there are reasons of a particular kind that count decisively in favour of it. Asking the question “Why should I be moral?” therefore amounts to asking “Why should I do what I have decisive reasons to do?” And as I noted in Chapter 1, this latter question is not a coherent one to ask.

This buck-passing account of normative requirements is attractive for at least two reasons. First, it avoids both of the answers to the normative question that we considered above and which seemed unsatisfactory. Second, it provides us with an analysis of what it means to say, in the case of normative requirements, that something is required by something else. In contrast, an account that considers normative requirements as themselves *providing* reasons, as John Broome suggests it, seems committed to the existence of independent truths about requirements that can neither be understood in merely descriptive terms (as non-normative requirements, such as the rules of etiquette, plausibly can) nor in terms of normative reasons.

The buck-passing account is thus a promising model for understanding moral, prudential, evidential and other normative requirements, and it provides an attractive response to the normative question about them. However, I do not think that the structuralist view can make use of the buck-passing strategy in order to account for the why-be-rational challenge. In Chapter 7, I shall propose something like a buck-passing view about rational requirements myself, but the requirements that I shall be concerned with are *non-structural* requirements of rationality to respond to available reasons, not structural requirements of rationality. While it is relatively obvious how such non-

---

<sup>14</sup> See e.g. Dancy (2000b, 165–167 and 172); Smith (1994, 182–184). I follow Dancy’s use of the term ‘verdictive’. It is also used by Foot (1978, 182) and Stratton-Lake (2000, 88), though not necessarily in exactly the sense at issue.

structural response requirements can be understood as verdictive statements about existing reasons, it is far from clear how structural requirements could be understood in this way.

First, structural requirements simply do not appear to be verdictive: we do not reach conclusions about such requirements by weighing considerations of a certain sort and then come to an overall (or domain-specific) verdict on the matter, as it seems natural in the case of moral, prudential, and other normative requirements.<sup>15</sup> Second, since structural irrationality occurs independently of the antecedent reasons that agents have for and against particular attitudes, structural requirements cannot be understood as verdictive claims about such reasons. They could still be verdictive claims about a particular group of reasons (call them “rational reasons”) that are specifically concerned with mental coherence.<sup>16</sup> But this just poses the question again what these reasons are supposed to be. We have a grip on the idea of what moral, prudential, or evidential reasons are: the fact that clouds are coming is an evidential reason for believing that it will rain; the fact that filing a return would save me a lot of money is a prudential reason to do it; and the fact that making fun of your mother would hurt your feelings is a moral reason to refrain from doing it. However, there does not seem to be a special group of reasons like that, which constitutes the domain of ‘rational reasons’ in contrast to moral or prudential ones.

In any case, I do not know of a viable interpretation of structural requirements as verdictive statements about reasons that exist independently of them. I shall consider a possible interpretation along these lines for *theoretical* requirements in the next section, but I shall reject it. So it seems that the buck-passing account is not an available strategy in meeting the why-be-rational challenge for structural requirements of rationality.

---

<sup>15</sup> This is not to say that we *always* reach conclusions of moral, prudential or other normative requirements in this way. The point is that this is a common and typical way to reach conclusions about such requirements, while it is not in the case of structural requirements of rationality.

<sup>16</sup> On an alternative reading than the one I suggested above, this might be what Southwood has in mind with “reasons *of* rationality” (2008, 18). Southwood never says, however, what these reasons are supposed to *be*, which is just the very question at issue.

#### 5.4 Are there epistemic reasons to be theoretically rational?

Andrew Reisner has recently put forward an argument designed to show that we necessarily have *epistemic* reasons to satisfy wide-scope structural requirements of theoretical rationality.<sup>17</sup> If this argument succeeds, we may be able to understand such requirements as verdictive statements about a certain kind of epistemic reasons, and thereby avoid some of the most severe problems discussed so far. Epistemic reasons are exactly the kind of reasons for which we can, and regularly do, hold beliefs. If, as Reisner claims, “the reasons to conform to a requirement of rationality are ordinary, truth-directed, epistemic reasons”<sup>18</sup>, we would neither have to appeal to instrumental reasons to be rational, nor would we have to claim that rational requirements are themselves reasons in order to vindicate the normativity of rationality.

Let us have a look at his argument. Reisner discusses a requirement not to have beliefs that are mutually inconsistent, and a version of the *modus ponens* requirement, according to which rationality requires us not to [believe  $p$  and believe  $p \rightarrow q$  and believe  $\neg q$ ].<sup>19</sup> I shall focus on the latter. The argument Reisner puts forward can be summarised as follows:

- (1) Evidence requires us not to believe [ $p$  and  $p \rightarrow q$  and  $\neg q$ ].
- (2) If evidence requires us not to believe [ $r$  and  $s$ ], then evidence requires us not to [believe  $r$  and believe  $s$ ].
- (3) Therefore, evidence requires us not to [believe  $p$  and believe  $p \rightarrow q$  and believe  $\neg q$ ].

On the assumption that evidential requirements are verdictive statements about evidential reasons, or at least correspond to such reasons, this argument would show that there necessarily are evidential reasons to conform to the structural requirement of

---

<sup>17</sup> See Reisner (2011, 46–49).

<sup>18</sup> Reisner (2011, 47).

<sup>19</sup> This requirement is equivalent to the disjunctive requirement to [not believe  $p$  or not believe  $p \rightarrow q$  or not believe  $\neg q$ ]. It is thus less demanding than the requirement I focused on, which requires us to [not believe  $p$  or not believe  $p \rightarrow q$  or believe  $q$ ].

rationality not to [believe  $p$  and believe  $p \rightarrow q$  and believe  $\neg q$ ]. We might understand this requirement as a verdictive statement about the existence of such reasons, thereby vindicating its normativity.<sup>20</sup>

There is a question of what exactly this argument amounts to. As Reisner argues himself, the prospects of providing a similar justification for structural requirements of *practical* rationality look dim.<sup>21</sup> His account therefore does not provide a general answer to the why-be-rational challenge. It would nevertheless be interesting if this challenge could partially be answered for the theoretical realm.

However, there is a second worry that Reisner's argument cannot even vindicate the normativity of all structural requirements of *theoretical* rationality, as he claims it does.<sup>22</sup> Recall the *modus ponens* requirement that we have so far focused on, the wide-scope version of which requires us not to [believe  $p$  and believe  $p \rightarrow q$  and not believe  $q$ ]. This requirement is stronger than Reisner's, for when we believe  $p$  and believe  $p \rightarrow q$ , we already violate this requirement if we *lack* a belief in  $q$ , while we violate Reisner's requirement only if we *have* a belief in  $\neg q$ . Whether this particular requirement is preferable to Reisner's may be open to discussion, but it seems that lack of belief can at least sometimes be structurally irrational. It is unclear, however, how Reisner's argument could even in principle justify such requirements. The argument starts off with the assumption that evidence requires us not to believe in a certain conjunction of propositions (1), and then proceeds to a requirement not to have the corresponding collection of beliefs in the conjuncts (3). Yet the attitude of lacking a belief has no content that could figure as part of the conjunction figuring in premise (1). It is thus doubtful that any requirement that prohibits lack of belief (in combination with other attitude-states) could be justified in the way that Reisner suggests.

---

<sup>20</sup> Note that this verdictive interpretation is not part of Reisner's view. However, without this assumption it seems that his argument, if successful, would not amount to a partial vindication of the normativity of rationality (as he claims); it would merely show that some rational requirements coincide with evidence requirements. See also Broome (2013, 198–99) for this problem. Broome does not consider the possibility that rational requirements are normative in the sense that they are verdicts about reasons.

<sup>21</sup> See Reisner (2011, 51–53).

<sup>22</sup> "This line of argument can be extended quite straightforwardly to any requirement of theoretical rationality, as long as it is a consistency requirement" (Reisner 2011, 47).

Leaving this issue aside, I shall now argue that Reisner's argument fails even for the restricted cases to which it can be applied. This is not because the argument is invalid: iterated applications of (2) validate the inference from (1) to (3). Moreover, I shall grant that premise (1) is true. Since the conjunction of  $p$ ,  $p \rightarrow q$ , and  $\neg q$  is necessarily and obviously false, it does indeed seem inevitable to say that evidence requires us to abstain from a belief in this conjunction. The problem is with premise (2):

(2) Necessarily, if evidence requires us not to believe [r and s], then evidence requires us not to [believe r and believe s].

On a natural view, evidence requires us to have a belief in  $p$  only if the evidential probability of  $p$  reaches a certain threshold (which might be vague or context-sensitive), and evidence requires us to abstain from a belief in  $p$  only if the evidential probability of  $p$  does not reach this threshold. Now, suppose that the threshold is reached for each of two propositions  $r$  and  $s$ , and evidence therefore requires us to believe  $r$  and it requires us to believe  $s$ .<sup>23</sup> On the assumption that evidence does not issue incompatible requirements, which seems certainly plausible, this means that the consequent of (2) is falsified: evidence cannot require us not to [believe  $r$  and believe  $s$ ] if it requires us to believe  $r$ , and it requires us to believe  $s$ . Does it follow that the threshold is reached also for the conjunctive proposition [r and s], such that the antecedent of (2) is likewise falsified? No, it simply does not follow. The probability of a conjunction of two propositions is usually lower than the probability of each of the propositions. As a result, evidence might require us to abstain from a belief in a conjunction [r and a], without requiring us to abstain from believing each of the conjuncts.

To illustrate this point, suppose that it is sufficiently likely that it will rain tomorrow, and sufficiently likely that you will receive a letter tomorrow, such that evidence requires you to have both of these beliefs. It still might not be sufficiently likely that it will rain and you will receive a letter tomorrow. That both of these things

---

<sup>23</sup> Plausibly, for it to be the case that we stand under an evidential requirement to believe  $p$ , further conditions beyond the threshold being reached have to be satisfied: for example, it might need to be the case that the propositions in case are relevant for us, that we attend to them, that we are interested in whether they are true, etc. We may here assume that this is the case. See Chapter 7.7 for a discussion of this question.

happen is less likely than that either happens. In such a case, it seems that evidence can require you to abstain from a belief in the conjunction, even though it does not require you to abstain from a belief in each of the conjuncts.

One might object that in the case that Reisner focuses on, the belief in the conjunction is not only insufficiently likely, but guaranteed to be false. So perhaps some variant of premise (2) can be justified, according to which evidence requires us not to have a collection of beliefs when the conjunction of their contents is guaranteed to be false. I have my doubts. Taken at face value, the moral of the lottery and the preface paradox is that we might well have sufficient evidence for each of a number of beliefs, even if we know that the conjunction of their contents is guaranteed to be false and we are thus required by evidence not to believe in this conjunction.<sup>24</sup> In any case, there are good reasons to doubt even this more restricted version of (2), and Reisner does not say anything in its support.<sup>25</sup>

Moreover, if we accepted this argument and concluded that there really is an evidential requirement not to [believe  $p$  and believe  $p \rightarrow q$  and believe  $\neg q$ ], we would still have to identify the normative reason that counts in favour of the required response. Reisner claims that there is such a reason, and that the reason is epistemic, but he does not say what the reason is supposed to be. Now, given that this reason must be necessarily in place, it cannot be a reason that counts particularly against believing  $p$ , believing  $p \rightarrow q$ , or believing  $\neg q$ ; it must be a reason that counts only against the combinations of these beliefs. Given that this reason is supposed to follow from the reason that is presupposed in premise (1), namely a reason not to believe [ $p$  and  $p \rightarrow q$  and  $\neg q$ ], it is natural to think that this reason consists in the fact that the conjunction of  $p$ ,  $p \rightarrow q$  and  $\neg q$  is necessarily false. But while this fact is a reason for which one can rationally give up a belief in this conjunction, it is not a reason for which

---

<sup>24</sup> This point is discussed in more detail in Chapter 9.6.

<sup>25</sup> Actually, he says that this premise “relies for its plausibility on the view that we can agglomerate under evidential requirement” (Reisner 2011, 49). It is not exactly clear what this is supposed to mean, however. By ‘agglomeration’, we usually mean that two modal claims which employ the same modal operator entail one modal claim about the conjunction of the contents of the former two claims. In our context, this would amount to the following claim: If evidence requires us to be in belief-state (a), and evidence requires us to be in belief-state (b), then evidence requires us to [be in belief-state (a) and be in belief-state (b)]. This claim seems to me both correct and uncontroversial, but it does not support Reisner’s premise (2).

one can rationally give up a belief in one of the conjuncts. So if a reason is a consideration *for which* one can give the response favoured by it, the idea that there is an epistemic reason against a combination of beliefs – as opposed to a reason against a belief in a conjunction, or a combination of reasons against beliefs in the conjuncts – is very hard to make sense of.

This gives us further and independent reason to doubt that the inference from (1) to (3) is valid. In any case, as long as Reisner does not say what the epistemic reason is supposed to *be*, he has not answered the why-be-rational challenge, even not for the particular subset of theoretical requirements of rationality on which he focuses.

## 5.5 The transparency account

Let me recapitulate. As I have argued in the last chapter, the assumption that we have decisive reason to conform to structural requirements of rationality leads to various detachment problems, which have no satisfying solution. Moreover, as I have argued in the present chapter, even if these detachment problems could be solved, there would still be no satisfying answer to the question of what these reasons *are*. We seem forced to conclude that there is not necessarily any reason to satisfy structural requirements of rationality. This raises a couple of questions: If rational requirements are not requirements of reasons, what are they then? Is there another way to account for the normativity of rationality? And if not: How can we account for the fact that rationality still *appears* to be normative? In this and the next section, I will discuss two accounts that aim to answer these questions.

I shall start with Niko Kolodny's *transparency account*, which has much in common with T.M. Scanlon's approach to structural rationality.<sup>26</sup> This account comes in two steps. The first of these steps is the claim that instances of structural irrationality always involve a conflict between an agent's attitude-state and the way an agent sees her reasons. The idea is that one of the conflicting attitudes present in all cases of structural irrationality is an attitude with *normative content*, and the relevant conflict is one between this content and the presence or absence of another attitude. On the

---

<sup>26</sup> See Kolodny (2005, 557–560) and Scanlon (1998, 18–30; 2007).

simplest and clearest construal of this view, structural irrationality always consists in the conflict of an attitude-state and a judgement to the effect that one has decisive reason against this attitude-state.<sup>27</sup> Let us call this the *normative attitude conception* of structural rationality.

The second step is the claim that statements about (structural) requirements of rationality are to be analysed in terms of the reasons that the agent believes himself to have. As Kolodny puts it:

When we tell someone, in the register of advice, [...] that he ought rationally to have attitude A, or that it would be irrational of him not to have it, we are [...] making the descriptive, psychological claim that he believes that he has conclusive reason for the attitude.<sup>28</sup>

The transparency account can thus be summarised as follows:

*The transparency account:* For it to be the case that A is rationally required to  $\phi$ , is for it to be the case that A believes that she herself has decisive reason to  $\phi$ .<sup>29</sup>

Various difficulties already arise with the *normative attitude conception* of structural irrationality. The first is that there seem to be instances of structural irrationality that do not involve attitudes with normative content at all, namely the first-order instances which I distinguished earlier from the second-order instances.<sup>30</sup> A second worry is that

---

<sup>27</sup> This is implicit in Kolodny's assumption that "all rational requirements are ultimately derived from" what he calls the "core requirements", namely the requirement to have an attitude one believes one has conclusive reason to have, and the requirement not to have an attitude one believes to lack sufficient reason to have (see Kolodny 2005, 557). See also Scanlon (1998, 20 and 25) for similar claims.

<sup>28</sup> Kolodny (2005, 557).

<sup>29</sup> Cf. Way (2009, 4). As the discussion below will show, it is not clear whether Kolodny really holds this view or whether his analysis of rational requirements in terms of psychological claims is restricted to the role such requirements play in advice.

<sup>30</sup> See Chapter 1.3 for the distinction, and Broome (2007d, 364–365 and 368) for this criticism of the transparency account. Kolodny and Scanlon offer different replies to this worry. While Kolodny (2007a; 2008a; 2011a) aims to explain first-order instances of structural irrationality directly in terms of (disjunctions of) reason requirements (see Chapters 9.1 and 10.2 for discussion), Scanlon (2007) seems to suggest that the first-order instances can be explained in terms of second-order requirements of rationality, by providing a constitutive account of the attitudes mentioned in first-order requirements, according to which

the view requires an artificial focus on what Scanlon calls “attitude-directed” (rather than “content-directed”) judgements.<sup>31</sup> If, as is commonly supposed, rationality can only require mental states, then the transparency account requires that the normative attitude involved in structural irrationality be a belief about reasons for *attitudes*. Yet common normative judgements are usually concerned with the contents of beliefs or intentions – i.e. with propositions and actions – rather than with the attitudes of belief and intention themselves. Finally, there is the problem that agents might judge themselves to have state-given rather than object-given reasons for attitudes, and it is doubtful that such judgements render them irrational if they do not have the corresponding attitudes.<sup>32</sup> An agent who has been offered a large reward for intending to drink a glass of toxin might believe that he has decisive reason to have this intention and yet fail to form this intention without irrationality if he does not believe that he has any reason to actually *drink* the toxin – or so it seems plausible to say.<sup>33</sup>

These problems might partly be answered by giving an alternative account of first-order structural requirements of rationality, by broadening the scope of the relevant normative attitudes, and excluding beliefs about state-given reasons. I shall not discuss these possibilities here in detail. Instead, I shall focus on the question whether – given that there is a version of the normative attitude conception of structural rationality which is both tenable and suitable for the purposes of the transparency account – the latter provides a satisfying explanation of the apparent normativity of rationality.

There is some confusion, I think, about what the explanation that both Kolodny and Scanlon aim to give is supposed to accomplish. Scanlon starts out by announcing that he wants to “examine [...] the kind of normativity involved in requirements of structural rationality”; he explicitly maintains that “claims about what it would be irrational to do are normative claims”.<sup>34</sup> Yet he goes on to claim that “normativity enters only from the point of view of the person” who has the normative

---

they have some kind of normative content (see Chapter 10.3). Broome (2007d) does not address either of these replies.

<sup>31</sup> See Scanlon (2007, 90–91), who notices this problem and sees the need to account for it.

<sup>32</sup> See Broome (2007d, 365–366) and Scanlon (2007, 90) for this problem.

<sup>33</sup> See Kavka (1983) for this example.

<sup>34</sup> Scanlon (2007, 84–85)

attitude in question.<sup>35</sup> He seems to take himself to account for the normativity of rationality by showing that “the relevant normativity is provided by what the agent sees as reasons”.<sup>36</sup> This strikes me as an instance of wishful thinking. Claims about what agents *see* as reasons are psychological claims, not normative claims. A radical error theorist about the normative will happily accept the truth of such claims. Scanlon just helps himself to the notion of normativity that he wants to account for.

Kolodny, too, claims that the transparency account “explains the normativity of rationality”, but he is quick to add that this normativity is only apparent.<sup>37</sup> As it stands, the transparency account is a straightforward psychologistic reduction of an apparently normative phenomenon; it does not purport to explain the normativity of rationality, but rather the *erroneous* impression that rationality is normative. It is an error theory.

So how does the transparency account explain the normative *appearance* of rationality? According to it, we are rationally required to have an attitude-state if, and only if, we believe there is decisive reason for us to have that state. Consequently, whenever we are rationally required to give some response, it must seem to us as if we had decisive reason for this response. From the first-personal perspective, then, rationality and reasons cannot come apart. This explains why we feel normative pressure to revise our attitudes in accordance with rational requirements; the pressure in question is simply the pressure of the reasons we believe ourselves to have.

So far, so good. The problem is that rational requirements appear to have normative force not only from the first-personal perspective of a person who is presently required to give a response. As Jason Bridges points out, rationality seems to have normative force also from a retrospective standpoint of an agent who has changed her views about reasons, or from the standpoint of a different person who does not share the agent’s assessments of reasons.<sup>38</sup> In these cases, the apparent normative force of rationality cannot be explained in the way that the transparency account suggests.

---

<sup>35</sup> Scanlon (2007, 87).

<sup>36</sup> Scanlon (2007, 88).

<sup>37</sup> Compare the headline of §5 of Kolodny (2005) on p. 557 with his remarks on p. 559.

<sup>38</sup> Cf. Bridges (2009, 357–362).

Kolodny himself stresses that “claims about rationality can function as advice”.<sup>39</sup> His analysis of the kind of advice that we give agents when we recommend a response as rational is as follows:

We are telling him, as we might put it, that *from his point of view, or as it seems to him*, he has conclusive reason to have the attitude. Thus, when we tell him that he ‘ought rationally’ to have attitude *A*, we are not ourselves offering him a reason to have *A*. How, then, are we advising him to have *A*? By drawing his attention to a reason for *A* that *he believes* he has.<sup>40</sup>

This explanation is puzzling for several reasons. Why would we draw a person’s attention to the reasons that he believes he has? At least apart from exceptional cases, the person will already know what he believes. And what would be the point of this venture given that we believe neither that the person’s beliefs are true, nor that he has independent reasons to adjust his attitudes to his beliefs about reasons? Drawing someone’s attention to the reasons that he believes to have cannot plausibly by itself count as advice if one does not share this person’s assessment of reasons. On any plausible conception, advice must involve at least some kind of normative endorsement on the part of the adviser.<sup>41</sup>

Moreover, if we allow for a notion of advice that amounts to no more than stating what the advisee believes about his reasons, it is unclear how this notion could capture what is ultimately at issue, namely the apparent normative force of rationality from the second-personal standpoint. If the notion of ‘advice’ in question does not involve any kind of normative endorsement, then it cannot help to explain why it will

---

<sup>39</sup> Kolodny (2005, 555).

<sup>40</sup> Kolodny (2005, 557)

<sup>41</sup> See Bridges (2009, 362) for a similar point. In his response to Bridges, Kolodny seems to withdraw his earlier position that “claims about rationality can function as advice” (2005, 555) in favour of the weaker thesis that “some charges of irrationality look more like advice than a mere appraisal could” (2009, 373). On the face of it, however, the stronger thesis seems plausible. Moreover, the weaker thesis faces the objection that I go on to discuss.

seem to the adviser that there is anything wrong with forming irrational attitudes.<sup>42</sup> But clearly, it seems to us that there is something wrong with it.

The transparency account is also incapable of accounting for the criticism involved in irrationality ascriptions. This is because criticism must involve at least some kind of normative judgement, while the transparency account analyses rational requirement claims as purely descriptive claims about what an agent sees as reasons. To illustrate this point, recall Nadeem Hussain's case of his well-educated Muslim friends who believe in God while freely admitting that they do not have sufficient reason to do so. As he describes the situation, his friends feel criticised when he calls them irrational, and they perceive this charge as a challenge. Hussain puts the point against Kolodny as follows:

Now imagine that in the face of the initial heated response, I say, "Calm down, I'm just making a descriptive psychological claim and one that you already agree with. Look, after all, you granted that you don't have sufficient reason to have *A*, and that you have *A*. Don't worry, my claim that you're irrational, doesn't [...] mean that *I* think there is anything wrong with the attitudes you have". I am sure we can all imagine the look of utter perplexity that would cross their faces. If the conversation were to continue in some intelligible fashion, *per impossible* no doubt, we can imagine them responding as follows: "Yes, we know you know that we think we lack sufficient reason for believing in God, but calling us irrational is saying that there is something wrong with us and that's what we deny. Don't pretend that you don't think there is something wrong with us when you call us irrational".<sup>43</sup>

---

<sup>42</sup> See Bridges (2009, 360–361) again. As Kolodny (2009, 369–371) points out, Bridges' criticism is partly based on the misunderstanding that the transparency account aims for an explanation of the appearance that rational requirements are *themselves* reasons or that we have certain reasons *because* we are rationally required. Kolodny denies, convincingly to my mind, that this assumption is part of ordinary discourse. The argument presented here, however, relies only on the assumption that in ordinary discourse, rational responses are normatively endorsed (not that they are normatively endorsed *because* they are rationally required), which seems hard to deny.

<sup>43</sup> Hussain (unpublished, 29).

This passage emphasises two aspects of rational requirements that the transparency account cannot capture. First, claims about rational requirements are subject to disagreement even if the participants of the dispute agree about all descriptive claims. The transparency account here shares with all other theories which aim to reduce apparently normative claims to merely descriptive claims the great burden of explaining how this can be so. Second, irrationality ascriptions are naturally understood as criticism, and it is obscure why this is so, and how such criticism could ever be legitimate, if rational requirements are understood in the way the transparency account suggests.

Kolodny might respond that both of these aspects of rational requirements are not supposed to be explained by the transparency account, but by an independent assumption, which, in a later article, he expresses as follows:

When *S* judges that *H* is being irrational in refusing to have *A*, *S* negatively appraises *H*, as functioning poorly, or manifesting some vice.<sup>44</sup>

One might wonder how this assumption is compatible with the reduction that Kolodny proposes in the context of his argument for the transparency account. Perhaps he thinks that statements about rational requirements have merely descriptive content when they are future-directed (as in advice), and evaluative content when they are directed at the past. Or perhaps he thinks that such statements actually have both evaluative and descriptive content. Be that as it may, complementing the transparency account with this claim might help to explain why disagreements about rational requirements are not exhausted by disagreements about descriptive psychological claims. It cannot, however, explain why irrationality is, or seems, criticisable. As I have argued in Chapter 2, functioning poorly or manifesting some vice is criticisable only if we assume that in doing so, we violate reasons that apply to us. But Kolodny aims to explain the phenomena in a way that does not presuppose that rational requirements correspond to reasons; quite the contrary, he maintains that this is not the case. So if the violation of such requirements nevertheless amounts to a vice or some kind of

---

<sup>44</sup> Kolodny (2009, 372).

malfunctioning, we are forced to conclude that we do not generally have a reason to refrain from manifesting such a vice or to function properly. We might then still disapprove of such a behaviour, but we cannot legitimately criticise it.

Let me sum up the discussion of this section. Far from providing an explanation of the normativity of rationality (as at least Scanlon seems to think), the transparency account actually entails that statements about rational requirements are not normative. It amounts to a radical error theory about the normativity of rationality. While it can explain why it must *seem*, from the first-personal standpoint of an agent who is presently rationally required to have some attitude, that he has decisive reason for this attitude, the transparency account fails to explain why rationality appears to be normative from other standpoints as well, such as the second-personal standpoint of advice. For this reason, it also cannot account for the fact that the attribution of irrationality is commonly understood as a form of criticism. If the arguments from Chapter 2 are correct, the transparency account is committed to the claim that ordinary discourse about rationality is deeply misguided, because irrationality is not criticisable after all. This is a conclusion that we might well seek to avoid.

## 5.6 The apparent reasons account

Let me turn to a second proposal about the relation between reasons and rationality. As we have seen, the transparency account explains rationality in terms of the reasons that an agent *believes* himself to have. According to what I shall call the *apparent reasons account*, rationality is to be explained instead in terms of the reasons that an agent *would* have *if* his beliefs were true. For example, Derek Parfit claims that “when we have beliefs whose truth would give us a reason to act in some way, we have [...] an *apparent* reason to act in this way”, and “what it would be rational for people to do depends on their apparent reasons”.<sup>45</sup> Mark Schroeder makes similar claims, and Jonathan Way, drawing on both Parfit and Schroeder, suggests the following generalisation:

---

<sup>45</sup> Parfit (2011a, 35). See also Parfit (2001, 25): “We are rational insofar as we respond to [...] apparent reasons. We have some *apparent* reason when we have some belief whose truth would give us that reason.”

*The apparent reasons account:* For it to be the case that A is rationally required to  $\phi$ , is for it to be the case that A would have decisive reason to  $\phi$ , if A's beliefs were true.<sup>46</sup>

Whenever the transparency account entails that a response is rationally required, the apparent reasons account also has this implication. This is because whenever an agent believes himself to have decisive reason to  $\phi$ , it is trivially true that if his beliefs were true, he would have decisive reason to  $\phi$ . However, the converse is not true. To illustrate this point, consider Parfit's famous example of a man who is indifferent to pains on future Tuesdays.<sup>47</sup> Suppose that this man believes that by changing his appointment for an operation from Tuesday to Wednesday, he would avoid a great amount of pain. It seems that if this belief were true, he would have decisive reason to change the appointment. However, suppose that this man does not believe that he has decisive reason to change his appointment, because he believes in a bizarre normative theory according to which pains on future Tuesdays are normatively irrelevant. In this case, the apparent reasons account, but not the transparency account, implies that he is rationally required to change his appointment. More generally speaking, whenever A believes that p, and p's truth would give A decisive reason to  $\phi$ , but A does not believe that he has decisive reason to  $\phi$ , the apparent reasons account, but not the transparency account, implies that A is rationally required to  $\phi$ .

The transparency and the apparent reasons account thus differ in extension: the latter deems many more responses irrational than the former. This gives rise to the worry that the apparent reasons account is too demanding. A person might believe something which would constitute a decisive reason just because of other facts that are not available to this person, in which case it does not seem irrational for the person not to respond to her apparent reasons. John Broome presents the following example:

---

<sup>46</sup> Cf. Way (2009, 4). Following Schroeder (2009), Way calls this view the "subjective reasons account". Schroeder does not himself speak of "rational requirements" but only of "subjective 'oughts'", but as he points out, his distinction between objective and subjective 'oughts' is equivalent to Parfit's distinction between reasons and rationality (see Schroeder 2009, 230, n. 15).

<sup>47</sup> See Parfit (1984, 124).

Suppose you believe this liquid is orange juice. Suppose that, if it is indeed orange juice, that is a perfect reason for you not to drink it, because you have overnight become so allergic to orange juice that drinking it will kill you. But suppose that nothing has given you any inkling on this allergy's onset. Then you might drink the liquid and not be irrational in doing so.<sup>48</sup>

This objection presupposes that what we have decisive reasons to do can be determined by facts to which we have no epistemic access. While Parfit, Schroeder, and Way seem to share this assumption, I will reject it in Chapter 8. If, however, reasons depend on available evidence in the way that I shall argue for, it is unclear whether an objection of this sort can be made to work. So I shall disregard it in what follows.

There are other problems with the apparent reasons account. First, as in the case of the transparency account, it is unclear how the analysis can account for first-order instances of structural irrationality. It is at least not obvious that an instrumentally irrational person necessarily has beliefs the truth of which would give her decisive reason to respond in a different way, especially since we are here assuming that there is no general reason to avoid instrumental irrationality as such.<sup>49</sup> Second, the apparent reasons account shares with the transparency account the problem that being unresponsive to judgements about state-given reasons does not seem to instantiate irrationality. Third, there is the problem that the beliefs of an agent might be inconsistent, and that since contradictions entail everything, the apparent reasons account seems to entail that agents with inconsistent beliefs stand under all possible rational requirements.

---

<sup>48</sup> Broome (2007d, 373). Broome credits the general line of this objection to Reisner. Ross (2006, ch. 7.4.4) makes a similar point.

<sup>49</sup> Schroeder (2009) and Way (2010a; 2012b) both have defended sophisticated accounts of instrumental rationality in terms of apparent reasons. Both rely on controversial assumptions, however. According to Schroeder, an intention to  $\phi$  necessarily involves the belief that one ought to  $\phi$ , which entails that agents always have decisive apparent reason to take the means believed necessary to their ends. According to Way, the fact that  $\psi$ -ing is a necessary means to  $\phi$ -ing is necessarily a (wide-scope) reason to [intend to  $\psi$  or not-intend to  $\phi$ ]. This entails that agents who believe that an action is a necessary means to one of their ends always have apparent reason to [intend the means or give up the end]. See Chapters 10.3–10.4 for a discussion of these views.

These problems might at least partly be solved by introducing further restrictions into the analysis.<sup>50</sup> As in the case of the transparency account, I shall not discuss these possibilities here in detail, but focus on the question of whether, assuming that an extensionally adequate version of the apparent reasons account is available, such an account could provide a satisfying explanation of the normativity, or apparent normativity, of rationality.

Again, I believe that there is some confusion about what this explanation amounts to. Proponents of the apparent reasons account do not hesitate much to call apparent reasons “subjective reasons”<sup>51</sup>, and to jump from facts about what we would have decisive reason to do if our beliefs were true to conclusions about what we “ought subjectively”<sup>52</sup>, “ought rationally”<sup>53</sup> or “ought in the belief-relative sense”<sup>54</sup> to do. This suggests that they take themselves to provide an “account of the normativity of rationality”, as Way puts it – an account of rationality according to which rationality is indeed normative.<sup>55</sup> They seem to hold that claims about apparent reasons are genuinely normative claims. Parfit states that “normativity involves reasons or apparent reasons”<sup>56</sup>, and Way claims that the apparent reasons account “implies that rationality is normative”, because it “explains rational requirements in terms of facts about reasons”.<sup>57</sup>

Now, there is indeed an interesting difference between the transparency account and the apparent reasons account regarding the question of normativity. In contrast to the transparency account, the apparent reasons account does not reduce claims about rational requirements to psychological claims, but to *hypothetical claims about reasons* of the form “If A’s beliefs were true, A would have reason to  $\phi$ ”. Way

---

<sup>50</sup> Way (2010a; 2012b), for example, restricts the relevant apparent reasons to object-given apparent reasons. For discussion of these and further problems of the apparent reasons account, see Ross (2006, ch. 7.3-7.4), Schroeder (2009, §5), and Vogelstein (2012).

<sup>51</sup> Schroeder (2009) and Way (2009, 3; 2012b, 499).

<sup>52</sup> Schroeder (2009, 228–230).

<sup>53</sup> Parfit (2011a, 34).

<sup>54</sup> Parfit (2011a, 163).

<sup>55</sup> See Way (2009).

<sup>56</sup> Parfit (2011a, 144; 2011b, 268).

<sup>57</sup> Way (2009, 4).

argues that such claims are “genuinely normative”, because an error theorist about normativity is committed to denying that such claims can be true.<sup>58</sup>

As I have noted elsewhere, I agree with Way that there is a technical sense in which hypothetical claims about reasons are normative, but I do not think that they are normative in the relevant sense.<sup>59</sup> Way’s point can be put as follows. Given that there are no deductive inferences from non-normative premises to normative conclusions, every claim that together with a non-normative claim entails a normative claim, must itself be a normative claim. This is true for all hypothetical claims of the form “If some non-normative claim were true, then A would have reason to  $\phi$ ”. Apparent reason claims such as “If A’s beliefs were true, A would have reason to  $\phi$ ” are a subclass of such claims, provided that the beliefs in question do not have normative content.<sup>60</sup> An apparent reason claim together with the non-normative claim that A’s non-normative beliefs are true entails a normative claim, and must therefore itself be normative.

We need, however, to distinguish two ways in which we refer to a claim as normative. In the narrow sense, a claim is normative when it is about what matters, what is important, what is right or wrong, advisable, criticisable, what we ought or have reason to do, etc. In a wider and rather technical sense, a claim is already normative whenever together with a non-normative claim it entails a normative claim in the narrow sense. Trivially, if reason claims are normative in the narrow sense, then all kinds of hypothetical claims about reasons, including apparent reason claims, are normative in the technical sense. But this technical sense is simply not at issue here. The kind of normative judgements we are interested in – the judgements that figure in deliberation, advice, or criticism – must involve some kind of *normative endorsement*: a person who judges such claims to be true will thereby (at least normally) express some kind of approval or disapproval. Only normative judgements in the narrow sense involve such endorsement, however. To think that someone *would* have reason for a

---

<sup>58</sup> Way (2009, 4).

<sup>59</sup> See Kiesewetter (2012, 470–471). The next four paragraphs constitute an adjusted and slightly revised version of this passage.

<sup>60</sup> This is an important restriction: A normative error theorist *can* accept claims of the form “If A’s *normative* beliefs were true, A would have reason to  $\phi$ ” – those claims are thus not even normative in the technical sense I am discussing here.

response *if* some condition obtained is not to endorse this response. In other words, hypothetical truths about reasons do not, as such, matter. They are truths merely about what *would* matter under certain conditions. But when we were asking the normative question about rationality, we wanted to know whether rationality *actually* matters.

Note that in the technical sense, the fact that you would have reason to  $\phi$  if *your* beliefs were true is just as normative as the fact that you would have reason to  $\phi$  if *anyone else's* beliefs were true. It should be clear, however, that for the purposes of deliberation, advice or criticism, we can perfectly well ignore truths about what you would have reason to do if some random person's beliefs were true. Such truths are in a technical sense normative and would have to be denied by a normative error theorist, but they may have absolutely no significance for anyone. So if rational requirements are normative in any sense that matters in deliberation, advice, or criticism, this normativity cannot be traced back to the alleged fact that rational requirements are hypothetical statements about reasons. Explaining rational requirements in terms of apparent reasons therefore cannot help us with meeting what Way himself concedes as the challenge for any theory about rationality, namely to explain that "the charge of irrationality is, in and of itself, a serious criticism".<sup>61</sup>

In order to meet this challenge, proponents of the apparent reasons account need to explain why some hypothetical truths about reasons matter and some do not. They have to tell us why it matters that a person would have reasons to  $\phi$  if her beliefs were true, but does not matter that a person would have reasons to  $\phi$  if some arbitrary other condition were fulfilled. It is difficult to see, however, how this difference could be spelled out other than in terms of *actual* reasons, by maintaining, in effect, that we have reasons to avoid being in the state in which we fail to have an attitude that would be required by reasons if our beliefs were true. Yet this is just the answer that has been challenged throughout the last two chapters; it raises the same problems of detachment and of identifying the reason in question that I have discussed above. So we seem to have come no further.

---

<sup>61</sup> Way (2009, 1).

There is an alternative interpretation of the apparent reasons account, which dispenses with the idea that apparent reason claims are in a relevant sense normative because they are hypothetical claims about reasons. The view I have in mind is that normativity is divided between (at least) two fundamental notions of ‘ought’, one of which corresponds to decisive reasons and the other of which corresponds to decisive apparent reasons.<sup>62</sup> Accordingly, the idea that claims about the second kind of ‘ought’ (the *subjective ‘ought’*, as it is often called) correspond to claims about the decisive reasons an agent would have if her beliefs were true is not supposed to explain or ground the normativity of such claims, but simply to state an interesting relation between these two independently normative kinds of ‘oughts’. Call this *the division view*.

The division view faces severe objections. First, as many others have pointed out, the idea that agents are subject to different ‘oughts’ that can issue incompatible directives is hard to make sense of.<sup>63</sup> Second, as I have argued in the last chapter, the division view is vulnerable to the bootstrapping objection. According to the division view, whenever we have beliefs whose truth would give us decisive reason to respond in a certain way, there is a genuinely normative ‘ought’ that requires us to give this response. The view thus allows us to generate genuinely normative demands simply by believing that we have reasons to act in these ways. This seems implausible. Third, the division view simply begs the question. We were looking for a justification of the intuition that we ought to avoid irrationality. Introducing a primitive normative notion which guarantees this claim’s truth does not seem to be a convincing response to this challenge.

One way of substantiating the view would be to show that the division is necessary not only in order to account for the normativity of rationality, but to account for normativity in general. And indeed, as we have seen earlier, both Parfit and Schroeder argue that in criticism, we generally need to appeal to a subjective sense of ‘ought’,

---

<sup>62</sup> On another construal, the two fundamental notions of ‘ought’ correspond to two fundamental notions of ‘reason’, where claims about reasons of the one kind (“subjective reasons”) correspond to claims about what would be reasons of the other kind if the agent’s beliefs were true. This latter construal is more in the vein of Schroeder (2009), while the one mentioned above is more in the vein of Parfit (2011a, 162–163).

<sup>63</sup> Compare Dancy (2000a, 61); Kolodny (2005, 512); Zimmerman (2008, 6–8).

which corresponds to our apparent rather than actual reasons, and they justify the division view by pointing to the different purposes that different practices such as criticism, decision-making, and advice have.<sup>64</sup> While I agree that deliberation, advice, and criticism do in part pursue different purposes, I think this is no reason to call into question that the concept of ‘ought’ central to these purposes is just one and the same. Intuitively, this seems to be the case: if it is correct to criticise someone for a response, it also seems correct to decide against it or advise others to refrain from it. I have already argued in Chapter 2 that truths about apparent reasons are irrelevant for criticisability; hence the assumption of a subjective ‘ought’ cannot be founded on its putative indispensability in criticism. The division view therefore strikes me as unmotivated.<sup>65</sup>

It thus seems that apparent reason claims are not relevantly normative, neither because they are hypothetical claims about reasons, nor because they correspond to an independently normative notion of ‘ought’. Like the transparency account, the view turns out then to be an error theory about the normativity of rationality, which is supposed to explain the illusory impression that rationality is normative. This view faces the same problems as the transparency account, however. The apparent reasons account might be seen as providing an explanation of why agents feel a normative pressure to revise their attitudes when they discover that they are being irrational, because it will then *seem to them* that they have decisive reasons against these attitudes. But it cannot explain why rationality seems normative from a second- or third-personal standpoint, because from these standpoints, the normative situation might seem different. It cannot explain why irrationality in others presses us to *advise* a revision of attitudes, and why it seems a legitimate ground for *criticism*. It must see these aspects of our common understanding of rationality and irrationality as misguided.

---

<sup>64</sup> See Parfit (2011a, 35–37); Schroeder (2009, 229).

<sup>65</sup> The division view is partly motivated by a conception of actual reasons, according to which reasons are provided by the facts alone, independently of an agent’s epistemic access. This is because it seems clear that the ‘ought’ of criticism cannot be determined by facts the agent has no access to. As I shall argue in Chapter 8, however, the notion of ‘ought’ or ‘decisive reason’ in decision-making and advice must also be understood in an evidence-relative way; there is thus no reason to think it is a different sense than the one used in criticism.

## 5.7 Summary

I have argued that the prospects of finding an answer to the question of why we should be structurally rational look dim. None of the possible strategies considered here appear capable of meeting this challenge. It does not seem that there are necessarily instrumental reasons to conform to structural requirements of rationality, nor that these requirements are reasons in themselves, nor that they are verdicts about independently existing reasons. The transparency and the apparent reasons accounts might provide an explanation of why rationality *seems* normative to us from the first-personal perspective. But they undermine rather than vindicate this appearance, and they are incapable of explaining the apparent normativity of rationality as it presents itself from the second- or third-personal standpoint.

This adds to the detachment problems for the structuralist view discussed in Chapter 4. Even if the proponent of this view could somehow circumvent these problems, he would still be incapable of meeting the why-be-rational challenge. I conclude that the structuralist view has proven to be indefensible.



## The myth of structural rationality

---

The last two chapters prompt the conclusion that structural requirements of rationality cannot be normative. This chapter goes one step further: I shall call into question that there are any structural requirements of rationality to begin with. The argument is that both narrow- and wide-scope interpretations of structural requirements face insurmountable problems quite independently of whether these requirements are understood as normative. Starting with the narrow-scope interpretation, I discuss the problem that narrow-scope requirements lead to bootstrapping (6.1), that they can be inconsistent with one another (6.2), and that they undermine each other in a counter-intuitive way (6.3). Turning to the wide-scope interpretation, I discuss the charge that wide-scope requirements cannot capture an important asymmetry involved in structural irrationality (6.4-6.5), and that they are incapable of guiding our responses (6.6). I argue that all of these objections pose serious problems for the respective accounts. This lends considerable support to the conclusion that there are no structural requirements of rationality (6.7).

### 6.1 Narrow-scope problem (i): bootstrapping

Let me start with the objection that narrow-scope requirements license implausible bootstrapping. In the traditional form – which was discussed at greater length in Chapter 4 – this objection presupposes that rational requirements constitute or provide decisive or at least *pro tanto* reasons.<sup>1</sup> The argument then proceeds as follows: Suppose you have no reason to intend some particular action, say spreading your lunch on the desks of your colleagues. But now suppose that you come to have the silly idea that you ought to spread your lunch on the desks of your colleagues. According to the narrow-scope version of the self-governance requirement (read as a requirement of reasons),

---

<sup>1</sup> For the traditional argument, see e.g. Bratman (1987, 24–27) and Broome (2001a, 98).

you now have a (possibly decisive) reason to intend to spread your lunch on the desks of your colleagues. But this is absurd. We started from the assumption that you had no such reason. You cannot make it the case that you now have a reason to do such a silly thing simply by believing that you ought to do it. You cannot, that is, “bootstrap” a reason into existence that easily. Therefore, the narrow-scope interpretation of the self-governance requirement must be false. And similar considerations yield the analogous result with respect to the other narrow-scope versions of structural requirements of rationality.

Since the argument in this traditional form presupposes the normativity of rationality, it merely shows that a *normative* reading of narrow-scope requirements leads to implausible bootstrapping. In other words, the only conclusion we are entitled to draw is that the narrow-scope account is false *or* rationality is not normative. It remains possible to defend the narrow-scope account and conclude that rationality is not normative.<sup>2</sup> Some authors have suggested that the bootstrapping objection is effective against the narrow-scope account independently of the assumption that we have reasons to be rational.<sup>3</sup> As far as I can see, however, this version of the objection has never been explicitly spelled out, which is what I aspire to do in this section.

To begin with, let me note a methodological problem that we encounter when running the bootstrapping objection against rational requirements independently of their normativity. The traditional bootstrapping objection appeals to our *normative* intuitions, both about the *genesis* and about the *content* of reasons. The central claim is that one cannot bring a reason or an ‘ought’ into existence simply by adopting a belief or an intention, independently of the reasons one has for this belief or intention in the first place. But this genealogical point is usually reinforced by examples of reasons with implausible content, which we would have to embrace if we accepted such bootstrapping, such as a reason to spread one’s lunch on the desks of colleagues. The general form of the traditional objection may be summarised as follows: “The narrow-scope account implies that people have all kinds of normative reasons to do the most stupid

---

<sup>2</sup> This is, in a nutshell, Kolodny’s central argument against the normativity of rationality (2005, 514–542).

<sup>3</sup> See Hussain (unpublished, §4.1), Parfit (2011a, 119), Reisner (2009, 248) and Way (2011, 228–229, esp. n. 4).

things, and they have these reasons through a simple decisionistic act or through believing something for no good reason. Since this is implausible, we need to reject the narrow-scope account.” Now, since a reason-independent bootstrapping objection obviously cannot appeal to intuitions about reasons, it seems that it must appeal to intuitions about rational requirements alone. However, it is unclear what intuitions about rational requirements can accomplish, for ‘rational requirement’ is a philosophical term of art to which we have no intuitive pretheoretical access. So which intuitions can we use to decide whether the genesis and the content of a particular rational requirement are acceptable?

The answer can be found in a source that is particularly unsuspecting in the context of an argument for a reason-independent bootstrapping objection to narrow-scope rational requirements, namely Niko Kolodny. This is because Kolodny subscribes to the traditional bootstrapping objection against the normativity of narrow-scope rational requirements, while at the same time he maintains a narrow-scope account of rational requirements and thus denies that this account is vulnerable to a reason-independent bootstrapping objection. When Kolodny discusses the technical term ‘rational requirement’, he notes that it must be anchored in ordinary judgements about rationality and irrationality:

Roughly, ‘S is rationally required to X,’ ought to be equivalent to more familiar judgments such as, ‘S is, or would be, being irrational in not X-ing,’ or ‘It is, or would be, rational for S to X.’<sup>4</sup>

Following Kolodny, we can thus say that the intuitions that we use to decide whether it is plausible to assume that A is rationally required to  $\phi$  are intuitions about whether it would be rational for A to  $\phi$  or irrational for A not to  $\phi$ .

It might be objected that statements of the form “It would be rational for A to  $\phi$ ” are simply statements about the property of rationality and thus cannot be equivalent to rational requirements, since – as I have argued in Chapter 1 – rational require-

---

<sup>4</sup> Kolodny (2005, 515).

ments cannot be reduced to statements about the property of rationality. However, it is not true that statements of the form “It would be rational for A to  $\phi$ ” are simply statements about the property of rationality. Rather, they express some kind of recommendation or advice from the standpoint of rationality (while we can leave open in this context whether this standpoint is a normatively binding one or not). Indeed, they behave exactly like rational requirement statements in terms of the logical relations to statements about the property of rationality. Just as “A is rationally required to  $\phi$ ”, “It would be rational for A to  $\phi$ ” does *not* entail the property statement that A would be rational if A  $\phi$ -ed. It makes perfect sense, e.g., to say that it would be rational for a person to intend a means to a valuable end that she deeply cares about, even though one knows that this person would not be rational if she did so, for example because she has some inconsistent beliefs that may not even stand in any interesting relation to the means/end-concern. Property statements to the effect that A is or would be rational are global, while statements of the form “It would be rational for A to  $\phi$ ” can be meaningful in more local contexts that are independent of whether A has (or would have) the property of rationality. And while “It would be rational for A to  $\phi$ ” (again just as “A is rationally required to  $\phi$ ”) *does* entail the property statement “A would be irrational if A did not  $\phi$ ”, it is *not* equivalent to it. For it makes perfect sense to say that A would be irrational if A did not  $\phi$  *because* it would be rational for A to  $\phi$ , while this could not make sense if the statements were equivalent. I mention another reason in a footnote below.<sup>5</sup> In sum, there seems to be no reason to doubt Kolodny’s suggestion that talk of rational requirements is ultimately anchored in more ordinary judgements of the form “It would be rational/irrational for A to...”. Consequently, we can use our intuitions about these more familiar judgements to test the plausibility of rational requirements.

---

<sup>5</sup> If the statements were equivalent, then (1) “It would be rational for A to omit  $\phi$ -ing” would also be equivalent to (2) “A would be irrational if A  $\phi$ -ed”. But (2) entails (2)’ “A would be irrational if A [ $\phi$ -ed and  $\psi$ -ed]”, where  $\psi$ -ing can be any arbitrary response type. By the equivalence assumption, (2)’ in turn entails (1)’ “It would be rational for A to omit [ $\phi$ -ing and  $\psi$ -ing]”. So the equivalence assumption implies that whenever (1) is true and it would thus be rational for A to omit  $\phi$ -ing, then if A omits any arbitrary response *other* than  $\phi$ -ing, A does something that would be rational for her to do, for according to the assumption, (1)’ follows from (1), and by omitting any arbitrary response  $\psi$ -ing, A omits the conjunctive response [ $\phi$ -ing and  $\psi$ -ing]. This, however, is highly counterintuitive.

So let us come back to the bootstrapping objection. Consider the following example of a young student called Jane. Jane has always been a true believer in the word of the bible, but she is also an ambitious biology student. Because of that, Jane is tempted to believe the conjunctive proposition that mankind was created within a day and mankind developed over million of years. We can suppose that it would be irrational of her to believe this proposition. We need not appeal to intuitions about reasons to do that, intuitions about structural rationality alone suggest that it is irrational to have a belief with contradictory content. Next, let us suppose that Jane comes to believe that she has sufficient evidence for this conjunctive proposition. Again, we can suppose that Jane acquires this evidence-belief on no rational basis. In fact, she comes to have it out of wishful thinking; she thinks that she has evidence for the contradiction only because she wants to believe in it and wants to think of herself as being rational in having this belief. Recall the narrow-scope version of the belief-guidance requirement (C)<sub>2</sub>:

(C)<sub>2</sub> If A believes that she has sufficient evidence that p, then rationality requires A to believe that p.

According to (C)<sub>2</sub>, Jane is now *in fact* rationally required to believe the contradiction that mankind was created within a day and mankind evolved over a million of years. That is, simply through the adoption of an irrational belief, Jane made it the case that a belief that seems clearly irrational is now required by rationality. This, I submit, is unacceptable bootstrapping.

Let us test this by Kolodny's own standards concerning which ordinary judgments go along with statements about rational requirements. Do we think that it would be rational for Jane to believe that mankind was created within a day and mankind evolved over million of years? Most certainly not.<sup>6</sup> Do we think that it would be

---

<sup>6</sup> I am not denying here that the narrow-scope advocate could *at the same time* maintain that it would be rational for Jane to drop this belief. At this point, the narrow-scope account faces another problem, namely that it allows for inconsistent requirements of rationality (see the next section). But this is irrelevant here: All that is needed for the argument presented above is that the narrow-scope account entails that it would be rational for Jane to believe the contradiction, which it clearly does according to Kolodny's own standards.

irrational for Jane not to adopt this belief? Again, the answer seems “no”, for it seems clear that it would be rational of her to drop her belief in the evidence and thereby avoid the irrationality in question. So it is implausible to assume that simply by adopting an attitude, independently of the rationality of this attitude, we can make it the case that it would be rational for us to have this attitude. The narrow-scope account thus entails implausible bootstrapping of highly counterintuitive requirements of rationality. In contrast, the wide-scope account avoids this implication, for according to this account, Jane is only rationally required to believe the contradiction *or* give up the belief that she has evidence for it. So we cannot conclude from the wide-scope account that it would be rational for Jane to believe the contradiction.

In his defence of a narrow-scope account of both the self-governance and the belief-guidance requirement, Kolodny simply assumes that the bootstrapping objection has force only if we grant that there are decisive reasons to do what rationality requires.<sup>7</sup> But as the example has shown, there is little reason to accept this. Insofar as we can abstract from the normativity of rational requirements at all, it seems that intuitions about rationality and irrationality alone suggest that there is something wrong with bootstrapping. This is a good reason to deny that there are any rational requirements that allow for such bootstrapping.

## 6.2 Narrow-scope problem (ii): inconsistency

The second objection against the narrow-scope account is that it leads to inconsistent requirements of rationality. Here is an example inspired by John Broome.<sup>8</sup> Peter is a catholic scientist who believes that he is facing a deontic dilemma: He believes he ought to investigate embryonic stem cells, but he also believes he ought not to investigate embryonic stem cells. According to a narrow-scope account of the self-governance requirement, Peter is rationally required to intend to investigate embryonic stem cells and Peter is rationally required to intend *not* to investigate embryonic stem cells. Plausibly, however, rationality requires us not to have logically inconsistent intentions,

---

<sup>7</sup> Kolodny (2005, 518–519, n. 10). See Hussain (unpublished, §4.1) for a helpful critical discussion of this passage.

<sup>8</sup> Cf. Broome (2007c, 364–365; 2007b, 28–33).

i.e. rationality requires Peter not to both intend to investigate *and* to intend *not* to investigate embryonic stem cells. Taken together, these requirements are inconsistent – Peter cannot satisfy all of them. According to the wide-scope account, in contrast, Peter is required only to have the intentions at issue *or* else give up his beliefs. So the wide-scope account does not give rise to inconsistent requirements on intentions.

It is implausible to assume that rationality could put us under inconsistent requirements in the way that the narrow-scope account implies. This is not because requirements in general must be consistent. Rationality is particularly associated with consistency, and consistency is generally seen as some kind of uncontroversial minimal condition of rationality. This makes it particularly implausible to assume that rationality can issue inconsistent requirements.

This is a widely shared view in philosophy, and it also gains support from pretheoretical intuitions about rationality. Recall Kolodny's claim that the technical phrase "A is rationally required to  $\phi$ " is roughly equivalent to the more familiar statement "It would be rational for A to  $\phi$ ". Intuitively, "It would be rational for A to  $\phi$ " seems to exclude "It would be rational for A to not- $\phi$ ". Suppose Peter asks his friend, "What do you think is the rational thing for me to do?" His friend would probably not answer, "Well, it would be rational for you to intend to investigate embryonic stem cells, *and* it would be rational for you to intend *not* to investigate embryonic stem cells; *moreover*, it would also be rational for you to *not* both intend to investigate *and* intend not to investigate embryonic stem cells". Such an answer would certainly seem paradoxical, but the narrow-scope account entails that it is the correct answer to give. This is a strong reason to deny that there are narrow-scope requirements of structural rationality.

In response, it might be argued that since it is assumed that Peter believes in a deontic dilemma, the example rests on the presupposition that Peter is irrational in the first place. And given Peter's irrationality, it might not seem so implausible that rationality imposes conflicting requirements on him. As Kolodny would put it, Peter has "painted [him]self into a corner".<sup>9</sup>

---

<sup>9</sup> Kolodny (2007b, 383).

For the sake of argument, let us grant that a belief in a deontic dilemma is necessarily irrational. This seems a tenable position, because there is good reason to believe that a belief in a deontic dilemma is not only substantially false but also contradictory: “A ought to  $\phi$ ” plausibly implies “A is permitted to  $\phi$ ” which is in turn equivalent to “It is not the case that A ought not to  $\phi$ ”. Consequently, Peter’s belief that he ought to investigate and ought not to investigate embryonic stem cells is contradictory.<sup>10</sup> However, it does not seem to me correct that Peter’s irrationality is a good reason to believe that rationality can require Peter to give inconsistent responses. As John Broome remarks, “we should expect rationality to require you to get out of your irrational state, not to get in deeper, into the further irrationality of having contradictory intentions”.<sup>11</sup> When we find ourselves in an irrational position, we can still sensibly ask which response would be rational for us to give, and it still seems paradoxical in such a situation to say that it would be rational for us to do one thing and it would be rational for us to do the opposite.

John Brunero presents a different response to the inconsistency challenge. Drawing on a terminology introduced by Kolodny, he distinguishes “local” from “global” requirements of rationality:

Local requirements of rationality specify what rationality requires of us in the light of some *specific* conflict of attitude-states, while global requirements of rationality specify what rationality requires of us in the light of our entire set of attitude-states. Global requirements, one might say, tell us what rationality requires of us, all attitude-states considered.<sup>12</sup>

---

<sup>10</sup> Note that I am assuming this here without much argument only to make my opponent as strong as possible. My opponent’s starting point is that Peter is irrational because he believes in a deontic dilemma. This premise seems without support if deontic dilemmas exist, or the fact that they do not exist is a substantive rather than conceptual truth.

<sup>11</sup> Broome (2007c, 365).

<sup>12</sup> Brunero (2010a, 33). I take it that both local and global requirements are supposed to be *structural* requirements of rationality in the sense that they are concerned with an irrational combination of attitudes independently of whether any of the particular attitudes involved is rational or supported by reasons. I also assume that the “specific conflict” need not be actual but can be a potential conflict (thus allowing the requirements in question to apply to rational as well as irrational persons).

The (A)-(D)-corresponding requirements of structural rationality at issue here are concerned with a potential conflict among *specific* attitude-states; they are thus intended by Brunero to fall under the category of *local* requirements. Brunero now claims that the distinction between local and global requirements of rationality “parallels a distinction often made by philosophers, following David Ross, between *prima facie* obligations and all-things-considered obligations”.<sup>13</sup> On this assumption, Brunero argues that since there is no problem with conflicting *prima facie* obligations, there is no problem with conflicting local requirements of rationality: “If this analogy is apt, then a conflict of local rational requirements need be no more mysterious than a conflict of *prima facie* obligations.”<sup>14</sup>

The problem with this argument is that the analogy is not apt. As Ross himself notes when he introduces this notion, a *prima facie* duty is not really a duty, but “something related in a special way to duty”<sup>15</sup>. The concept that Ross has in mind makes a *contribution* to duty (or obligation) by counting morally in favour or against an action, but does not itself constitute a duty, because it can be outweighed by other considerations. I agree with those who think that a *prima facie* duty is best understood as a *pro tanto* moral reason, but what is important here is only that *prima facie* duties are contributory. As a consequence, conforming to a *prima facie* duty is not a necessary condition for being moral. This is the very point of introducing this concept: Ross wants to be able to say that one may be morally required to do something (e.g. lying), even though there is a *prima facie* duty not to do it. Someone who lies to a murderer in order to save the lives of innocents does not conform to all his *prima facie* duties, but he might nevertheless be entirely moral.

Contrast structural requirements of rationality, such as those corresponding to (A)-(D). The very idea that such requirements exist is based on the assumption that some combinations of attitudes are necessarily irrational. Conforming to a rational requirement that applies to one therefore *is* a necessary condition for being rational. And the idea of contributing to an overall-judgement or being outweighed by other

---

<sup>13</sup> Brunero (2010a, 33).

<sup>14</sup> Brunero (2010a, 34).

<sup>15</sup> Ross (1930, 20).

considerations does not seem to play a role. In short, Brunero's analogy fails because what he calls local requirements are supposed to be real requirements, while *prima facie* duties are not supposed to be real duties. Conflicts between *prima facie* duties are unproblematic, because it is the very point of *prima facie* duties to make a contribution that can be outweighed. The requirements that are the topic of our discussion, however, make claims on agents that cannot be outweighed in the sense that it is possible for an agent to violate such a requirement and still be entirely rational. In contrast, a putative rational requirement is falsified if it is possible to violate it without irrationality. Since such requirements are really supposed to be requirements, a conflict between them is problematic in a way that conflicts between *prima facie* duties are not.<sup>16</sup>

In sum, it seems to me that the objection against the narrow-scope account is sound. Narrow-scope versions of structural requirements of rationality allow that the requirements of rationality are inconsistent. Even if such inconsistency occurs only in cases in which the agent is irrational in the first place, this is a very implausible result. This is partly because, contrary to what Brunero maintains, the requirements in question do not parallel *prima facie* duties, but are supposed to be strict requirements, the violation of which is necessarily irrational.

### 6.3 Narrow-scope problem (iii): undermining

The third objection to the narrow-scope account is that narrow-scope requirements undermine each other in a counterintuitive way. This can be demonstrated by means of an example that John Brunero presents as an objection to the narrow-scope account – an objection that is independent of the assumption that rationality cannot issue incon-

---

<sup>16</sup> Bratman (2009a, 413) suggests that structural requirements of rationality are 'pro tanto'. For my part, I find it not so easy to make sense of the idea of a contributory coherence requirement, and Bratman does not elaborate on this idea very much. One obvious worry is that such *pro tanto* requirements cannot explain the stringency associated with structural irrationality, cf. Broome's "strictness test" (1999, 405) and Schroeder (2009, esp. 232–233) on stringency. Consequently, Bratman's *pro tanto* requirements cannot explain the structural irrationality claims (A)-(D), as long as the latter are not supplemented with a *ceteris paribus*-clause. In any case, Bratman has to deny that the violation of a rational requirement necessarily constitutes irrationality, and this is a claim that Brunero presupposes himself in his argument (cf. Brunero 2010a, 36, and the discussion in the next section), and which – as far as I can see – is shared by all other participants in the debate. Given that this claim is accepted, the analogy with *prima facie* duties must fail.

sistent requirements. The example is based on the claim that it is irrational to believe something while believing that one lacks sufficient evidence to believe it – a variant of *weakness of belief*(C):

(C-) If A believes that she herself lacks sufficient evidence that p, and A believes that p, then A is irrational.<sup>17</sup>

We are currently investigating the possibility of accounting for such claims in terms of structural requirements of rationality with narrow scope. The claim of the narrow-scope account is that (C-) corresponds to the following requirement:

(C-)₂ If A believes that she lacks sufficient evidence that p, then rationality requires A to not believe that p.

Recall as well the narrow-scope version of the self-governance requirement:

(A)₂ If A believes that she ought to  $\phi$ , then rationality requires A to intend to  $\phi$ .

Here is Brunero's example: Annie believes that she ought to go to the lecture, but she also believes that she lacks sufficient evidence that she ought to go to the lecture. According to (A)₂, rationality requires Annie to intend to go to the lecture. According to (C-)₂, rationality requires Annie to not believe that she ought to go to the lecture. In contrast to Peter's case, these requirements are consistent: Annie can satisfy both requirements, namely by giving up her belief that she ought to go to the lecture and adopting the intention to go to the lecture. However, it is implausible to say that Annie is required to do both of these things. By requiring the absence of an attitude-state that is presupposed by another requirement, one of these requirements undermines the other in a counterintuitive way. To put it somewhat differently, there is nothing to be

---

<sup>17</sup> "Rationality requires one not to believe that p, if one believes that there is not sufficient evidence that p" (Kolodny 2005, 521). See also Brunero (2010a, 29) and Lord (2011, 394). I follow these authors in marking this claim with a minus behind the "C", because the corresponding narrow-scope requirement (see below) requires absence of belief, while the narrow-scope version of the evidence-belief requirement requires belief.

said from the standpoint of rationality in favour of *both* not believing that one ought to go to the lecture *and* intending to go to the lecture. We can thus claim:

- (I) The narrow-scope account implausibly entails that Annie is *both* rationally required to not believe that she ought to go to the lecture *and* rationally required to intend to go to the lecture.

Brunero goes further than this:

Suppose [Annie] drops her belief that she [ought] to attend and also does not form the intention to attend – a perfectly sensible thing for her to do. According to [the narrow-scope account], she has failed to do something she was rationally required to do.<sup>18</sup>

So according to Brunero, the narrow-scope account not only entails that Annie is required to *both* not believe that she ought to go *and* intend to go; even worse, it entails that if Annie, in reaction to her situation, drops her belief without forming the intention, she is irrational:<sup>19</sup>

- (II) The narrow-scope account implausibly entails that unless Annie *both* drops her belief that she ought to go to the lecture *and* forms the intention to go there, Annie is irrational.

In defence of the narrow-scope account, Errol Lord presents an objection to (II). According to Lord, when Annie drops her belief that she ought to go to the lecture, then she not only satisfies  $(C-)_2$ , but also *exits* the requirement  $(A)_2$ , for  $(A)_2$  is conditional on her having the belief. Since exiting a requirement does not entail violating it,

<sup>18</sup> Brunero (2010a, 35).

<sup>19</sup> I am assuming that “failing to do something one is rationally required to do” implies irrationality. As becomes clear in later passages, Brunero shares this assumption (see e.g. the third paragraph on p. 36).

the apparent conclusion cannot be drawn.  $(A)_2$  and  $(C-)_2$  do not together entail that Annie is irrational if she only drops her belief, but does not adopt the intention.<sup>20</sup>

According to Lord's diagnosis of the disagreement, Brunero falsely assumes that exiting a requirement entails violating it.<sup>21</sup> And if Brunero's objection (II) really rested on this assumption, we would have to reject it, because exiting a requirement clearly does not entail violating it. However, Brunero does not need to make this assumption, and is better interpreted as not making it. The real issue between Brunero and Lord is whether their interpretation of the relevant narrow-scope requirements allows for exiting requirements in the first place. And this depends, in part, on whether the requirements are supposed to be synchronic or diachronic.<sup>22</sup>

Consider the option where  $(C-)_2$  and  $(A)_2$  are synchronic requirements:

- (1) For all times  $t$ , if  $A$  believes at  $t$  that she lacks sufficient evidence that  $p$ , then rationality requires  $A$  to not believe at  $t$  that  $p$ .
- (2) For all times  $t$ , if  $A$  believes at  $t$  that she ought to  $\phi$ , then rationality requires  $A$  to intend at  $t$  to  $\phi$ .

<sup>20</sup> For this argument, see Lord (2011, 395–396).

<sup>21</sup> Cf. Lord (2011, 393): "Brunero is right that narrow-scoping is doomed *if* we make a particular assumption about the logic of requirements, *viz* that all non-compliance states are violation states. His argument depends on making this assumption." Cf. also Lord (2011, 395): "I have isolated a crucial assumption of Brunero's argument, namely, that by exiting from a narrow-scope conditional requirement one is violating that requirement."

<sup>22</sup> A note on the notion of "exiting a requirement": It seems to me that, strictly speaking, one cannot exit a rational requirement in the sense that one makes it the case that *the very same* conditional requirement that applied to one before no longer applies to one. This is because the conditions of structural requirements of rationality make reference to occurrences of mental states and thus implicitly to time. That is, we have to understand statements such as "If  $A$   $\phi$ -s, then rationality requires  $A$  to  $\psi$ " as short for "For all times  $t$ , if  $A$   $\phi$ -s at  $t$ , then rationality requires  $A$  to  $\psi$  at  $t$ " (for reasons of simplicity, I focus on the synchronic case here, but the same goes for diachronic requirements as well). But note that this latter statement is not literally a requirement, but a general form for an infinite set of requirements. So what we have is actually an infinite set of requirements such as (1) "If  $A$   $\phi$ -s at  $t_1$ , then rationality requires  $A$  to  $\psi$  at  $t_1$ ", (2) "If  $A$   $\phi$ -s at  $t_2$ , then rationality requires  $A$  to  $\psi$  at  $t_2$ ", etc. Now suppose that  $A$  satisfies the antecedent of (1), but not of (2). As a consequence, while  $A$  is required to  $\psi$  at  $t_1$ , she is no longer required to  $\psi$  at  $t_2$ . So did  $A$  exit a requirement? But which requirement would that be? She did not exit (1), for the antecedent of (1) remains true. But she also did not exit (2), for the antecedent of (2) has never been true. Hence there is no particular requirement that applies to  $A$  at one time, but not at another time, which is what the very idea of exiting seems to presuppose. What actually happens is that one avoids a requirement that has the same general form as *another* requirement that applied to one earlier. We may, of course, use the expression "exiting a requirement" in this loose way, and this is what I do when I use it here.

We are supposing that:

- (3) Annie believes at  $t$  that she lacks sufficient evidence that she ought to go to the lecture.
- (4) Annie believes at  $t$  that she ought to go to the lecture.

We can conclude that:

- (5) Rationality requires Annie to not believe at  $t$  that she ought to go to the lecture (*from 1 and 3*).
- (6) Rationality requires Annie to intend at  $t$  to go to the lecture (*from 2 and 4*).

Since

- (7) If rationality requires  $A$  to  $\phi$  at  $t$ , and  $A$  does not  $\phi$  at  $t$ , then  $A$  is irrational at  $t$ ,

we can further conclude that:

- (8) If Annie does not *both* not believe at  $t$  that she ought to go to the lecture *and* intend at  $t$  to go to the lecture, then Annie is irrational at  $t$ . (*from 5, 6, and 7*)

But this is not a very interesting conclusion. For (8) already follows from the fact that Annie is irrational at  $t$ . And that Annie is irrational at  $t$  is uncontroversial if we assume a synchronic account of rationality. It follows from (3), (4), and (C-) alone so long as we read (C-) as a synchronic claim, independently of whether one accounts for (C-) by structural requirements of rationality and whether these requirements have wide or narrow scope.

An interesting conclusion would be that Annie is irrational in not intending to go to the lecture *at some later time at which she no longer believes that she ought to go to the lecture*. But nothing like this follows from a *synchronic* version of (A)<sub>2</sub>. Quite the contrary, (A)<sub>2</sub> does not even apply at that later time, because it is conditional on Annie's

having the belief. All of this is a consequence of the fact that was demonstrated in Chapter 3: synchronic versions of wide- and narrow-scope requirements do not differ with respect to their implications for whether an agent is rational or irrational at a particular time.

As regards the disagreement between Brunero and Lord, we can conclude that on the assumption of synchronicity, Brunero's objection (II) is indeed false. Since synchronic requirements do not require anything beyond the time at which the antecedent attitude holds, they allow – to use Lord's term – “exiting”. As a consequence, they do not entail that Annie is irrational when she drops her belief without forming the intention.

However, we should keep two things in mind. The first is that nothing we have said casts any doubt on objection (I), an objection that Lord does not address at all. The second thing to note is that Brunero directs his argument against Kolodny, and proceeds on Kolodny's assumption that requirements of structural rationality are *not* synchronic.<sup>23</sup> Brunero does not mention how exactly he conceives of the requirements in question, but for the reasons given in Chapter 3, I think it is best to construe them as diachronic state-requirements rather than as process-requirements.<sup>24</sup> So let us consider the diachronic versions of the relevant narrow-scope requirements:

- (A)<sub>7</sub> If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , then rationality requires A to intend at  $t'$  to  $\phi$ .
- (C-)<sub>7</sub> If A believes at  $t$  that she lacks sufficient evidence that  $p$ , and if there is a time  $t'$  (possibly later but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , then rationality requires A to not believe at  $t'$  that  $p$ .

<sup>23</sup> Cf. Brunero (2010a, 35): “Kolodny and I [...] are both concerned to specify what Kolodny [...] calls ‘process requirements’ – requirements which say ‘how, going forward, one is to form, retain, or revise one’s attitudes’”.

<sup>24</sup> The difference does not matter for my argument to the effect that diachronic requirements do not allow for exiting. It is clear that the same holds for narrow-scope *process*-requirements, since such requirements prescribe a unique way of how to proceed in time.

Assuming that there is a time  $t'$  that is later than and appropriately related to  $t$ , an argument analogous to the one given above would yield the following conclusion:

(8)\* If Annie does not *both* not believe at  $t'$  that she ought to go to the lecture *and* intend at  $t'$  to go to the lecture, then Annie is irrational at  $t'$ .

And (8)\* is indeed an interesting conclusion about whether Annie is rational or irrational. In contrast to (8), (8)\* is not implied by (3) and (4), together with (C-).<sup>25</sup> Hence, (8)\* is a particular implication of the narrow-scope account, and it is a particularly implausible one. Indeed, it is exactly the one that Lord regards as false. For he agrees with Brunero that if Annie reacts to her situation “by dropping her belief and *not* forming the intention to go”, this “does not seem irrational for Annie”.<sup>26</sup>

So it seems that on the assumption of diachronicity, Brunero’s objection (II) can be maintained. And this is not because exiting a requirement generally entails violating it, it is because the diachronic requirements (A)<sub>7</sub> and (C-)<sub>7</sub> do not allow exiting. If you have held the antecedent attitude at  $t$ , you cannot escape the requirement by giving up the antecedent attitude at  $t'$ . The requirements in question apply independently of what you do at  $t'$ .

In order to avoid Brunero’s objection (II), then, narrow-scope advocates either have to maintain that the requirements in question are synchronic, or they have to suggest an account of diachronic narrow-scope requirements which in contrast to (A)<sub>7</sub>–(D)<sub>7</sub> are escapable. We have already seen, however, that both of these options have unacceptable results. As regards the first option, recall the case of Michael, a brilliant thinker who receives new information and adjusts his attitude-states in the fastest rational way. A synchronic account of structural requirements of rationality implausibly condemns Michael as irrational. As I have argued in Chapter 3, this is not acceptable; any plausible account of rational requirements must be diachronic. As regards the

<sup>25</sup> Nor is it implied by (3) and (4), together with the following *time-sensitive* version of (C-), (C-)\*: If A believes at  $t$  that she lacks sufficient evidence that  $p$ , and if there is a time  $t'$  (possibly later but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , and if A believes at  $t'$  that she lacks sufficient evidence that  $p$ , and A believes at  $t'$  that  $p$ , then A is irrational at  $t'$ .

<sup>26</sup> Lord (2011, 395).

second option, the only way to make a diachronic narrow-scope requirement escapable is to include, as a further condition of the requirement, that the agent holds the antecedent attitude not only at  $t$ , but also at  $t'$ . The self-governance requirement, for instance, would read:

(A)<sub>15</sub> *Diachronic self-governance requirement, narrow scope + escapable*: If A believes at  $t$  that she ought to  $\phi$ , and if there is a time  $t'$  (possibly later, but not earlier than  $t$ ) such that  $t'$  is appropriately related to  $t$ , and if A believes at  $t'$  that she ought to  $\phi$ , then rationality requires A to intend at  $t'$  to  $\phi$ .

The requirement (A)<sub>15</sub> is escapable in a way that (A)<sub>7</sub> is not. According to the former, Annie may not be required to intend at  $t'$  to  $\phi$  if she drops her belief immediately after  $t$ . As a consequence, this requirement avoids Brunero's objection (II). According to (A)<sub>15</sub>, it is possible that Annie is entirely rational in dropping her belief that she ought to  $\phi$  without forming the intention to  $\phi$ . But as I also have argued in Chapter 3, including such a condition comes with the uncomfortable consequence that the requirement in question is principally incapable of guiding the very processes of rational adjustment that it was supposed to make room for in the first place. For the relevant condition requires the presence of an attitude state at  $t'$ ; as a result, the requirement does not apply before  $t'$ . But in order for the requirement to be capable of guiding the relevant adjustments, it must apply at the time the process of adjustment is supposed to *begin*, not just at the time at which it is supposed to be finished. The narrow-scope account thus faces a dilemma: Either it is vulnerable to Brunero's objection (II), or it avoids this objection by including a condition for requirements which make them incapable of guiding the very rational processes that diachronic requirements are supposed to make room for.

In any case, it is worth noting that just like the synchronic narrow-scope account, both diachronic narrow-scope accounts are vulnerable to objection (I). To see this, suppose that Annie continually believes between  $t$  and  $t'$  that she ought to go to the lecture and that she lacks sufficient evidence that she ought to go to the lecture.

On any account that aims to account for something like (C-), it follows that Annie is irrational at  $t'$ . However, it is a particular implication of the narrow-scope accounts – whether synchronic as in (A)<sub>2</sub>-(D)<sub>2</sub>, diachronic as in (A)<sub>7</sub>-(D)<sub>7</sub>, or diachronic and escapable as in (A)<sub>15</sub>-(D)<sub>15</sub> – that Annie is *both* rationally required to not believe at  $t'$  that she ought to go to the lecture *and* rationally required to intend at  $t'$  to go to the lecture. And as I have argued above, this is an implausible result, because these requirements undermine each other. There is no point in doing both of these things, and hence it is not plausible to say that they are both required by rationality.

To sum up this discussion, I have argued, drawing on the case of Annie, that both (I) and (II) present forceful objections to the narrow-scope account. Whether (II) can be maintained depends on whether my arguments against synchronic accounts of structural rationality and against escapable diachronic narrow-scope requirements are sound. Objection (I), however, applies independently of these arguments.

#### 6.4 Wide-scope problem (i): asymmetry

We have seen that there are weighty objections against narrow-scope structural requirements of rationality. Wide-scope requirements avoid these objections. Since wide-scope requirements do not require the agent to be in the consequent state, but allow instead the revision of the antecedent state, they do not generally license the kind of bootstrapping – neither of reasons, nor of rational requirements – entailed by narrow-scope requirements. For the same reason, they do not issue inconsistent or undermining requirements.

The wide-scope account is confronted with other problems, however. Yet some of them have already been solved along the way. Recall the original version of the wide-scope self-governance requirement:

(A)<sub>3</sub> Rationality requires of A that [if A believes that she ought to  $\phi$ , then A intends to  $\phi$ ].

It has been argued that requirements of this form rely on the awkward idea that a conditional (or its truth) – rather than an action or some other agential response – is required, and that due to their unconditional form, they allow for weird instances and posit inescapable and unexplainable requirements.<sup>27</sup> But as we have seen, the main point of the wide-scope account, namely that structural requirements of rationality do not allow for the detachment of a requirement to be in the consequent state because they can also be satisfied by dropping the antecedent state, can be spelled out in ways in which these requirements are understood as relating to agential responses rather than propositions, and as being conditional rather than unconditional. In the remainder of this chapter, I shall discuss two further objections to the wide-scope account, which I think are not so easily avoided.<sup>28</sup>

In this section, I shall start with the asymmetry challenge. Wide-scope requirements pose a symmetry between different ways of satisfying the requirement. This is just another way of saying that they are disjunctive: one can satisfy such requirements by exhibiting either of the responses that are part of the disjunctive response that is required. This feature is part of the attraction of the wide-scope account; it is, for example, the reason why wide-scope requirements do not license detachment via *modus ponens*. However, the symmetry of wide-scope requirements also poses a problem for the wide-scope account, for it seems that there is an *asymmetry* between different ways of satisfying structural requirements of rationality that wide-scope requirements cannot capture. Niko Kolodny and Mark Schroeder have independently raised a version of this objection to the wide-scope account.<sup>29</sup>

I shall briefly illustrate this point on the basis of the four requirements I am focusing on. First, consider the self-governance requirement. Suppose you believe that you ought to do the dishes now, but you do not intend to do so. One way to satisfy the wide-scope requirement is to form the intention to do the dishes now. This seems plausible. Another way to satisfy the wide-scope requirement is to cease to believe that you ought to do the dishes now. Indeed, the wide-scope account postulates a strong

---

<sup>27</sup> See Schroeder (2004) and the discussion in Chapters 3.2 and 3.3.

<sup>28</sup> There are other objections which I do not address here, see e.g. Evers (2011) and Dreier (2009).

<sup>29</sup> Kolodny (2005); Schroeder (2004). Bedke (2009, 687–689) and Finlay (2010, 70–71) subscribe to this objection as well.

symmetry between intending to do the dishes and ceasing to believe that you ought to do them as two equally good ways of satisfying the requirement. Intuitively, however, these reactions are not rationally on a par. *Prima facie*, it does seem to be a rational reaction to your situation to form the intention, but it seems less of a rational reaction to drop the belief.

Next, consider the instrumental requirement. Suppose that you intend to take the bus in five minutes, you believe that in order to do so, you have to leave now, but you do not intend to leave now. You are irrational. According to the wide-scope account, you can satisfy the instrumental requirement by ceasing to believe that in order to catch the bus in five minutes, you have to leave now. Intuitively, however, this is not a rational reaction to your situation. There seems to be a relevant asymmetry concerning the satisfiability of the instrumental requirement: you can become instrumentally rational by intending the means, and perhaps by abandoning your end, but not by simply dropping your means/end-belief.<sup>30</sup>

Similar considerations hold for the theoretical realm, as the case of the evidence-belief requirement shows. If you believe that you have sufficient evidence for  $p$ , but you do not believe that  $p$ , it seems there is an asymmetry between believing  $p$  and dropping your belief in the evidence as ways of escaping your irrational state. But the wide-scope account suggests a strong symmetry between these two reactions.

The *modus ponens* requirement might not be seen as asymmetric in this way. It seems that if you find yourself believing that  $p$  and that  $p \rightarrow q$ , but not  $q$ , starting to believe  $q$  and dropping one of your antecedent beliefs really are on a par as rational reactions to your situation. But remember Gilbert Harman's point that the requirement, if it is not supposed to require us to clutter our minds with irrelevant information, needs to be supplemented with a condition according to which you take it to be relevant whether  $q$  (or something similar). If we take this into account, the wide-scope

---

<sup>30</sup> The instrumental requirement might be modified in the following way, in order to escape this objection: "If you intend to  $\phi$ , and if you believe that  $\psi$ -ing is a necessary means for  $\phi$ -ing, then rationality requires you to [not intend to  $\phi$  or intend to  $\psi$ ]." This requirement is not satisfiable by dropping the means/end-belief. Way (2010a, 223) suggests a "medium-scope" account of the instrumental requirement along these lines, although for different reasons. Note, however, that this solution is not available for other requirements. Moreover, as I argue below, the medium scope account entails implausible bootstrapping (see note 60 below).

view entails that you can satisfy the *modus ponens* requirement by stopping to care about whether  $q$ , or stopping to attend to  $q$ , and this again seems like a reaction that is not rationally on a par with the other options.<sup>31</sup>

Intuitions about the asymmetry between different ways of satisfying rational requirements are widely acknowledged, by both advocates of the narrow-scope and advocates of the wide-scope account.<sup>32</sup> Wide-scope advocates hold, however, that they pose no threat to the wide-scope account of structural requirements of rationality. One response on behalf of the wide-scope account is based on the distinction between local and global requirements that was briefly discussed above. Recall that the (A)-(D)-corresponding requirements are local rather than global; they are concerned with a potential conflict between particular attitude-states rather than with the question of what is the most rational thing to do given the whole set of one's attitudes. If a person violates such a local requirement, we can conclude that she is irrational, but we cannot conversely conclude that a person is fully rational when she satisfies such a requirement. As John Brunero argues in defence of the wide-scope account, the following two claims are therefore compatible: (1) the local requirement concerning attitude-states  $A$  and  $B$  can be satisfied by dropping  $A$  or by dropping  $B$ , and (2) in most cases of conflicts between  $A$  and  $B$ , it is more rational on the whole to drop  $B$  rather than to drop  $A$ . While (1) is the commitment of the wide-scope account, (2) is supposed to capture the asymmetry intuition.<sup>33</sup>

Asymmetry is thus explained by the contingent fact that rational requirements combine in certain ways more often than in others. For example, in most cases of instrumental irrationality, the requirements of theoretical and practical rationality

---

<sup>31</sup> See Schroeder (2004, 355) for this point. Note that Schroeder argues for the asymmetry between stopping to care whether  $q$ , on the one hand, and believing  $q$ , or ceasing to believe  $p$  or  $p \rightarrow q$ , on the other hand, as – in the case of the instrumental requirement – he argues for the asymmetry between dropping your means/end-belief, on the one hand, and intending the means or dropping the end, on the other hand. This is surprising, because his favoured narrow-scope account poses a much stronger asymmetry, namely between all the antecedent attitudes, on the one hand, and the sole consequent attitude, on the other.

<sup>32</sup> For narrow-scope advocates, see Kolodny (2005, 527–539) and Schroeder (2004, esp. 339–340 and 346–347). For wide-scope advocates, see Broome (2013, 139), Brunero (2012, 126), Fink (2010, 113–114) and Way (2011, 231–232). For advocates of neither account, see Bedke (2009, 687–689). For advocates of both accounts, see Shpall (2013, 722).

<sup>33</sup> See Brunero (2012, 130). Fink (2010, 128) similarly claims that asymmetry intuitions can be explained by contingent facts about how different rational requirements combine.

combine in such a way that giving up one's means/end-belief is not the most rational thing to do. According to Brunero, this explains our asymmetry intuition in this case, but it does not follow that the local instrumental requirement cannot be satisfied by giving up one's means/end-belief.

There are two problems with this response. The first is that the notion of a global structural requirement of rationality – a requirement to respond in a particular way that is most rational given all one's attitudes – cannot be taken for granted, especially not by wide-scope advocates. Since wide-scope requirements can always be satisfied in at least two different ways, it is hard to see how they could combine in such a way that a particular response turns out to be the most rational one.<sup>34</sup> So in order to give the envisaged response, wide-scope advocates first have to tell us how their account can be augmented by some non-disjunctive rational requirements, without falling into the traps of the narrow-scope account. I am not saying that this is impossible; I only maintain that it cannot be taken for granted, especially since wide-scope advocates themselves have argued that rational requirements are generally disjunctive.<sup>35</sup>

The second problem of the response is the assumption that asymmetry is a merely contingent matter, namely a matter of how different local requirements usually combine to form global requirements. It seems that there is a deeper explanation for the asymmetry intuitions, one that has to do with the nature of the attitudes involved, and the possibility of *basing* attitudes on one another in a rational way.<sup>36</sup> Take any of the two attitude-states (or pairs of attitude-states) mentioned in (A)-(D). What really stands out is that it is always possible to rationally base one of the involved attitude-states on the other, but not the other way around. It is possible, for example, to rationally base an intention to do something on a belief that you ought to do it.<sup>37</sup> But if you cease to believe that you ought to do something on the basis of your lack of intention to do it, this is not a rational process; it is wishful thinking. Similarly, it seems plausible that you can rationally base an intention to  $\psi$  on an intention to  $\phi$  and a belief that  $\psi$ -ing is a necessary means to  $\phi$ -ing. This is the paradigmatic case of instrumental reason-

---

<sup>34</sup> See Section 6.6 for further discussion of this point.

<sup>35</sup> See e.g. Broome (2005a, 323): "it is very plausible that rational requirements generally have a wide scope."

<sup>36</sup> See Broome (2013, 186–91) for a helpful discussion of rational basing.

<sup>37</sup> Broome (2001b) calls this "normative practical reasoning".

ing.<sup>38</sup> But it does not seem that you can rationally drop your belief that  $\psi$ -ing is a necessary means to  $\phi$ -ing *on the basis of* your intention to  $\phi$  and your lack of intention to  $\psi$ . Quite the contrary, if you drop your belief in a means/end-relation on the basis of your intentions, what has happened was most certainly an irrational process of wishful thinking. This is a significant and non-contingent asymmetry. Let us call it a *basing asymmetry*.<sup>39</sup>

Narrow-scope advocates believe that rational requirements should reflect such basing asymmetries. They maintain, in Matthew S. Bedke's words, that "rationality has a built in directionality that wide scope requirements do not capture".<sup>40</sup> As Schroeder puts it:

The first problem for Wide-Scoping is that it is symmetric. It doesn't distinguish between acting in accordance with your moral beliefs and adopting moral beliefs in accordance with your actions, and as a result it fails to distinguish between following your conscience and the distinctive vice of rationalization. Rationalization is the vice of changing your beliefs about what you ought to do, because you are not going to do it, anyway. According to the Wide-Scope view, this is precisely as good a way of satisfying this requirement as is actually paying attention to what you believe and acting accordingly.<sup>41</sup>

The difference between following one's conscience and rationalisation is a basing asymmetry. Schroeder is right that wide-scope requirements cannot account for this asymmetry. The question is, however, whether that is their job. Note that basing an attitude-state on another is a process. The requirements in dispute are state-requirements; so far, neither the narrow-scope nor the wide-scope requirements tell us anything about whether we are required or permitted to *base* any attitude on another. Wide-scope advocates can thus argue that basing asymmetries have to be accounted for

---

<sup>38</sup> See Broome (2002) for an account of such reasoning.

<sup>39</sup> Way (2011, 232) uses this expression in the same context.

<sup>40</sup> Bedke (2009, 687).

<sup>41</sup> Schroeder (2009, 227).

by independent basing principles. Broome has defended several such principles. A plausible example is:

Rationality requires A not to drop her belief that she ought to  $\phi$  on the basis of her not intending to  $\phi$ .<sup>42</sup>

Given that it is irrational to drop one's 'ought'-belief on the basis of a lack of an intention (or on the basis of a belief in a lack of intention), such independent basing principles are needed anyway: neither the symmetric wide-scope requirements nor the asymmetric narrow-scope requirements entail by themselves that such basing is prohibited. Wide-scope advocates can thus argue that the state-requirements that correspond to structural irrationalities such as (A)-(D) need not account for basing asymmetries, because this job is done by other principles.<sup>43</sup>

Some philosophers have articulated the worry that as a result of this suggestion, wide-scope requirements ultimately collapse into narrow-scope requirements. For instance, when Schroeder discusses the idea that the wide-scope account could be complemented by independent principles that account for basing asymmetries, he complains that "if there really is a principle according to which [an agent] ought not to change his belief, then in any circumstances in which that principle applies, the only way for him to fulfil his Wide-Scope requirement to either-do-A-or-not-believe-that-he-ought is for him to do A".<sup>44</sup> The worry seems to be that a wide-scope requirement and an additional requirement accounting for asymmetry could together be satisfied only in the very way that the respective narrow-scope requirement prescribes. But Schroeder proceeds on the assumption that the additional principle must be a prohibition to drop the belief, and he overlooks the possibility that it is a prohibition only to drop the belief *on the basis of a lack of intention*. One can abide by this latter prohibition and the relevant wide-scope requirement in ways that are not allowed by the respective narrow-scope requirement, for example by dropping the 'ought'-belief on the basis of

---

<sup>42</sup> Cf. Broome (2013, 141). My formulation deviates slightly from Broome's.

<sup>43</sup> For this response to the asymmetry challenge, see Broome (2013, 138–41) and Way (2011).

<sup>44</sup> Schroeder (2009, 227).

available evidence. There is thus no reason to suppose that the response considered here threatens to collapse the wide-scope account into a narrow-scope account.<sup>45</sup>

To sum up, wide-scope advocates can rebut the objection that the wide-scope view cannot account for basing asymmetries by introducing independent basing principles. Insofar as basing asymmetries pose a problem for the wide-scope view at all, it must be because of a principled reason that structural requirements of rationality that correspond to claims such as (A)-(D) have to reflect such basing asymmetries. In order for asymmetry to pose a problem, a principle along the following line must be true:

*Asymmetry reflection:* If a rational requirement governs attitude-states A and B, between which there is a basing asymmetry such that it is possible to rationally base B on A, but not possible to rationally base A on B, then one can satisfy the requirement only by being in state B, and not by being in state A.

Niko Kolodny employs a similar principle in his argument against the wide-scope account:

*Reasoning Test.* The [...] requirement governing the conflict between A and B is wide scope – that is, one is rationally required (either not to have A or not to have B), only if, from a state in which one has conflicting attitudes A and B, (i)

---

<sup>45</sup> See also Way (2011, 234). There might be another worry with the response. According to Shpall (2013, 722), the asymmetry to be captured by a good account of rationality includes the thought that “there is always something rational about forming the intention to x, assuming that you do so in response to [...] the fact that you believe that you ought to x”. Insofar as this is correct, it might be argued that it cannot be accounted for by basing *prohibitions* alone, but needs to be accounted for by basing *permissions* (see Broome 2013, 189–91, on suggestions for such permissions). Shpall argues, however, that the notion that “rationality permits A to base her intention to x on her belief that she ought to x” (which Broome seems to embrace) involves implausible bootstrapping in cases of irrational (or otherwise unacceptable) ‘ought’-beliefs (cf. Shpall 2013, 725–726). He concludes that the wide-scope view cannot account for asymmetry. I am not convinced by this argument. If “rationality permits A to base her intention to x on her belief that she ought to x” is understood as entailing that A is permitted to intend to x, or to believe she ought to x, then bootstrapping lingers. But Broome is cautious enough to reject this entailment, and Shpall does not object. However, if a rational permission to base an intention to x on a belief that one ought to x is compatible with a rational prohibition to intend to x as well as a rational prohibition to believe that one ought to x, it is difficult to see why there should be a bootstrapping problem.

one can reason from the content of A to dropping B and (ii) one can reason from the content of B to dropping A.<sup>46</sup>

As should be clear by now, the (A)-(D)-corresponding requirements I have been discussing all along do not pass this test, because there is a basing asymmetry between the attitudes involved in them. Kolodny exemplifies this by reference to the self-governance requirement. You can reason from your belief that you ought to  $\phi$  to the intention to  $\phi$ . But you cannot reason from your not having an intention to  $\phi$  to giving up the belief that you ought to  $\phi$ . First of all, and quite obviously, reasoning has to start with some content, but an absent mental state such as your non-intention does not have content. Secondly, and perhaps more significantly, even if you could come to not believing that you ought to  $\phi$  in light of your non-intention to  $\phi$  (say, by starting from the content of your existing belief that you do not have this intention), this process would not be reasoning, but rather self-deception.<sup>47</sup>

Hence, if the *reasoning test*, or some other version of *asymmetry reflection*, is correct, then the wide-scope account has to go. But what is the argument for *asymmetry reflection*, given that there are other principles that can account for basing asymmetries? Kolodny's own argument for the *reasoning test* is based on the assumption that structural irrationalities have to be accounted for by *process-requirements* – an assumption which, as I have argued in Chapter 3, is neither supported by Kolodny's arguments, nor plausible on further reflection.<sup>48</sup> I have also argued, however, that there is a plausible

---

<sup>46</sup> Kolodny (2005, 520–521). Kolodny restricts this test to process-requirements, but he also maintains that all rational requirements are process requirements. As I have argued in Chapter 3, however, there are no good reasons to believe that structural irrationalities such as (A)-(D) could be explained by process-requirements. In my view, the most charitable reading of Kolodny's argument therefore starts from the assumption that *state-requirements* must be satisfiable by rational processes. I develop an argument along these lines in the next section.

<sup>47</sup> Cf. Kolodny (2005, 527–530).

<sup>48</sup> Way (2011, 237–238) argues that Kolodny's argument for the reasoning test fails even if we grant that rational requirements are process-requirements, which seems to me correct. Brunero (2010a, 44–47) argues – or at least appears to argue – that Kolodny's argument for the narrow-scope account fails even if we grant the reasoning test, which seems to me false. Brunero holds that “even if we were to accept the reasoning test as a proper way of restating the rational response test [...], there is a more serious problem for Kolodny's argument [...] concerning his employment of both tests”. The problem is that “his argument for the narrow-scope interpretation [...] relies upon a hidden premise, which I shall term ‘atomism about narrow-scope rational requirements’, according to which an attitude-state which puts in place a narrow-scope rational

idea behind Kolodny's concern for processes, and I shall now consider whether further spelling out this idea can amount to a rationale for *asymmetry reflection*, and thus to an argument against the wide-scope account.

### 6.5 Asymmetry and the rational process constraint

The idea that I think plausibly connects the requirements that are violated in the states described in (A)-(D) to rational processes is what I have called the *rational process constraint*:

*Rational process constraint*: For every rational requirement R, R requires A to be in state S, only if [either A satisfies R or there is a rational process available to A which leads to S].

Indeed, it seems very plausible that one cannot be under a rational requirement unless one can come to satisfy this requirement in a rational way. I have given an argument for this claim and explained in more detail the content of the rational process constraint in Chapter 3. Here, I shall proceed on the assumption that it is true, and ask what conclusions are to be drawn from it concerning the asymmetry challenge.

The rational process constraint is obviously satisfied by all the narrow-scope requirements we have considered so far. Quite generally, these requirements are conditional on an attitude that is a possible rational basis for the required attitude. But the wide-scope requirements meet the rational process constraint as well. As we have seen above, such requirements can either be understood unconditionally or conditionally. If

---

requirement in one context *must* put in place a narrow-scope rational requirement in another context" (Brunero 2010a, 46). Now, it is clear that the claim that Brunero calls "atomism" is implied by the narrow-scope account, but it is difficult to see how Kolodny's argument needs it as a premise in addition to the reasoning test. Let 'a' and 'c' be the antecedent and consequent states involved in a structural irrationality claim. Let 'NS' be the proposition 'If A is in a, then rationality requires A to be in c' and 'WS' be the proposition 'Rationality requires A to [be in c or not be in a]'. Kolodny's argument can then be represented as follows: (1) Either NS or WS (uncontroversial between narrow- and wide-scope advocates). (2) If WS, then one can reason from a to c as well as from c to a (reasoning test). (3) One cannot reason from c to a (uncontroversial). (4) Therefore, not-WS (from 2 and 3). (5) Therefore, NS (from 1 and 4). That is, in addition to the reasoning test no further controversial premise, such as atomism, is needed for Kolodny's argument to go through.

they are understood in the latter way, they are conditional on the same attitude(s) as are the narrow-scope requirements and can thus be satisfied by the very same rational process as narrow-scope ones. If they are understood in the former way (i.e. unconditionally), then we have to consider two cases: the one in which the agent is in the state that provides the rational basis and the one in which he is not. If he is, the requirement can again be satisfied in the same way as the narrow-scope requirement. If he is not in that state, then the requirement is already satisfied, in which case the rational process constraint is met automatically. So the rational process constraint does not provide an argument against neither narrow- nor wide-scope requirements.

Note, however, that according to this argument, wide-scope requirements meet the rational process constraint only because they are either already satisfied, or because they are satisfiable in the same way as their narrow-scope counterpart. This is suspicious because it is the distinctive claim of the wide-scope account that the requirements in question can be satisfied in ways in which the narrow-scope requirements cannot be satisfied. Let us say that *S* is a *satisfaction state* of a rational requirement *R* if, and only if, *S* is an atomic mental state such that necessarily, if *R* applies to *A*, and *A* is in *S*, then *A* satisfies *R*. (A mental state is *atomic* just when it is not composed of a conjunction or disjunction of other mental states.)<sup>49</sup> It is the distinctive claim of the wide-scope account that structural requirements of rationality correspond to at least two satisfaction states. For instance, both the state of not-believing-that-one-ought-to- $\phi$  and the state of intending-to- $\phi$  are satisfaction states of the wide-scope self-governance requirement, whereas only the latter state is a satisfaction state of the corresponding narrow-scope requirement.

Given the possibility that different states count as satisfaction states of a rational requirement, it seems appropriate to extend the rational process requirement in the following way:

---

<sup>49</sup> The restriction to *atomic* states in the definition of a satisfaction state is necessary in order to avoid an infinite proliferation of trivial satisfaction states. For example, suppose that necessarily, if *R* applies to *A*, and *A* believes *p*, then *A* satisfies *R*. A trivial consequence is that necessarily, if *R* applies to *A*, and *A* believes-*p*-and-believes-*q*, then *A* satisfies *R*, and so forth. In such a case, the restriction to atomic states prevents 'believing-*p*-and-believing-*q*' from counting as a satisfaction state.

*Advanced rational process constraint:* For every rational requirement R, R requires A to be in state S, only if [either A satisfies R or for every satisfaction state of R, S\*, there is a rational process available to A which leads to S\*].

The advanced rational process constraint seems justified if we take into account the full commitments of claims about rational state-requirements. Such requirements not only involve the claim that an agent is (under certain circumstances) required to be in a certain state (the “requirement claim”), they also involve the claim that if the requirement applies to an agent, and the agent is in the required state, then he satisfies the requirement (the “satisfaction claim”). In this latter respect, wide-scope requirements involve logically stronger claims than narrow-requirements. Both agree that a particular state counts as a satisfaction state, but the wide-scope requirement claims that there is a further state that counts as a satisfaction state as well. The wide-scope account offers, as it were, an additional way to satisfy a given requirement – a way that is, according to the wide-scope advocate, overlooked in the narrow-scope account. The asymmetry intuitions which I emphasised in the last section (and about which there is wide agreement among both narrow- and wide-scope advocates), show that this additional claim cannot be regarded as independently plausible. It is not independently plausible, for example, that someone who escapes weakness of will by dropping his belief that he ought to do something thereby comes to satisfy a rational requirement. So the claim that one can come to satisfy rational requirements in these ways is in need of justification. It is difficult to see how it could be justified, unless there were a rational process available to the agent by which he could reach the state that according to the wide-scope requirement counts as a satisfaction state. And this just means that the advanced rational process constraint is true.

All non-disjunctive state-requirements that meet the simple rational process constraint also meet the advanced rational process constraint. This is because according to such requirements, there is only one state that constitutes the satisfaction state of the requirement. Not so for disjunctive requirements: since they allow that different states constitute a satisfaction state, they can fail to meet the advanced rational process constraint even if they meet the simple one.

The (A)-(D)-corresponding wide-scope requirements meet the simple, but not the advanced rational process constraint. An agent to whom such a requirement applies either satisfies the requirement, or he can reach its satisfaction state through a rational process. But it is not true that he either satisfies the requirement or can reach *every* state that according to the requirement constitutes a satisfaction state through a rational process. For example, not-believing-that-one-ought-to- $\phi$  is a satisfaction state of the wide-scope self-governance requirement. But it is not true that if the requirement applies to an agent A, then either A satisfies the requirement or there is a rational process available to A which leads to not-believing-that-one-ought-to- $\phi$ . We can stipulate that A does not satisfy the requirement: he believes he ought to  $\phi$ , but does not intend to  $\phi$ . Whether A can give up his belief through a rational process is still an entirely contingent matter, most plausibly it depends on A's evidence or his beliefs in the evidence. But the rational requirement claims we are considering are supposed to hold in all possible worlds; consequently the strong rational process constraint demands that A can give up his belief through a rational process in all worlds in which the wide-scope requirement applies to him and he does not satisfy it.<sup>50</sup> Showing that there is a possible world in which this is not the case suffices to demonstrate that the wide-scope requirement does not meet the advanced rational process constraint.

State-requirements involve certain claims about which states count as satisfaction states. The problem with wide-scope state-requirements is that they allow attitude-states to be satisfaction states, even though there may be no rational processes available that lead to these states. In those possible worlds in which these processes are not available, agents can rationally come to satisfy these requirements only in the same way as the respective narrow-scope requirement. In such cases, it does not seem plausible to say that there really are two ways of satisfying the rational requirement – but this is the distinctive claim of the wide-scope account.

Moreover, note another awkward feature of the wide-scope account. Call a process P a *satisfaction process* of a rational requirement R if, and only if, necessarily, if A passes through P, then A is in the process of satisfying R. For example, necessarily, if

---

<sup>50</sup> Conditional requirements do, of course, only hold in worlds in which the condition is satisfied, but the claims about such requirements (“If p, then rationality requires...”) are supposed to be necessary.

one is forming an intention to  $\phi$  on the basis of a belief that one ought to  $\phi$ , then one is in the process of satisfying the self-governance requirement; hence, this process is a satisfaction process of the requirement. I have mentioned the fact that in many cases, there are no available rational processes that lead agents to satisfy the wide-scope requirement in the way that the corresponding narrow-scope requirement cannot be satisfied. This speaks against wide-scope requirements, which are supposed to be true in all possible worlds. A possible response would be to say that the wide-scope requirements hold only in those possible worlds in which the alternative rational processes *are* available. Note, however, that these processes are never processes that can count as satisfaction processes *of* the respective requirement. This means that whenever one rationally satisfies a wide-scope requirement R in the way that the corresponding narrow-scope requirement cannot be satisfied, the process one is running through is not the satisfaction process of R, it is the satisfaction process of quite another requirement, R\*.

This point can be illustrated by again making reference to the self-governance requirement. Consider the possible worlds in which there is a rational process available to A that would lead A to give up his belief that he ought to  $\phi$ . These would be worlds in which A lacks evidence that he ought to  $\phi$ , or believes that he lacks this evidence, or has other beliefs that, if true, would indicate such a lack of evidence.<sup>51</sup> On grounds of these evidential considerations, A may rationally drop his belief. But the process in which A would be involved is not a satisfaction process of the self-governance requirement, it is the satisfaction process of a rational requirement having to do with belief-changes in the light of evidential considerations. Thus, A can rationally come to satisfy the wide-scope self-governance requirement in the way that he cannot satisfy the correspondent narrow-scope requirement only if he also satisfies a *different* rational requirement and only if he runs through a rational process that is essentially linked to this *other* requirement and has virtually nothing to do with self-governance as such. Is it plausible to say in a case in which an agent just forms his beliefs in accordance with his

---

<sup>51</sup> My assumption here is, of course, that all rational processes of belief-change involve, at least implicitly, reference to evidential considerations, but it seems to me that an argument similar to the one in the main text can be founded on alternative conceptions (though not every alternative conception) of rational belief-change.

evidence, that this agent satisfies a rational requirement of *self-governance*? Of course, by giving up his belief, the agent avoids the irrationality of weakness of will. All accounts of (A) will have this implication. But it is a particular implication of the wide-scope account of (A) that this agent not only avoids weakness of will (and thus avoids the violation of a rational requirement), but positively satisfies a requirement that is particularly concerned with self-governance. And yet the only possible process that he could have run through in a rational way in order to achieve that state is not concerned with self-governance at all, but only with belief-formation in light of evidence.

These considerations seem to me to support a further generalisation of the rational process constraint:

*General rational process constraint:* For every rational requirement R, S is a satisfaction state of R, only if there is a rational process P leading to S such that P is a satisfaction process of R.

In other words, for every state that is, according to a rational requirement, a satisfaction state, there must be a corresponding rational process leading to that state which can be conceived of as a satisfaction process of the requirement in question (rather than as a satisfaction process of some other requirement). The (A)-(D)-corresponding wide-scope requirements do not meet this constraint. There are rational satisfaction processes leading from the antecedent to the consequent attitude-states, but no rational process leading an agent to drop the antecedent state is a satisfaction process of the respective requirement. This is due to the basing asymmetries discussed above. Given the plausibility of the advanced and the general rational process constraint, rational state-requirements have to account for such asymmetries, as asymmetry reflection demands. And this means that the wide-scope requirements are false.

I admit that this argument is contestable. To me, all three versions of the rational process constraint seem plausible, but the considerations I have offered in favour of them are not beyond dispute. What I take to be the upshot of the discussion is that the asymmetry challenge for the wide-scope account is not as easily solved as it may seem. Even if we assume independent basing principles that can explain the basing

asymmetries discussed above, the fact that wide-scope requirements do not reflect such asymmetries remains a problem, for rational state-requirements plausibly *correspond* to rational processes through which we can come to satisfy the requirement. My point is that it is a disconcerting property of wide-scope requirements that they are supposedly satisfiable in two ways, even though one of these ways does not correspond to a rational process specifically connected to this requirement, and often to no rational process at all. This, I think, casts considerable doubt on the wide-scope account.

### 6.6 Wide-scope problem (ii): guidance

Let me now turn to a second challenge for the wide-scope account, namely the challenge to explain how such requirements could be guiding. I shall assume that if there are structural requirements of rationality, it should at least be possible that these requirements guide us in our mental lives, that we can orientate ourselves on them in order to become structurally rational.<sup>52</sup> It might be argued that this assumption presupposes a normative interpretation of such requirements, whereas I have here confined myself to arguments that are independent of this presupposition. But this is not so. To claim that rational requirements are normative is to claim that such requirements *necessarily* involve reasons (possibly decisive ones) to conform to them. Denying this claim does not mean to deny that persons can at least *sometimes* have reasons, or can sometimes *take* themselves to have reasons, to satisfy these requirements. Only this latter claim is presupposed by the assumption that rational requirements must be capable of guiding. It seems to me that this assumption is part of what it means to talk of ‘requirements’ not merely in the sense of a necessary condition for having a certain property, but in the sense of a code or rule that prohibits or demands certain responses. As I shall argue, however, the idea of being guided by a structural wide-scope requirement of rationality is very difficult to make sense of.

---

<sup>52</sup> Cf. Kolodny (2007b, 371–372): “requirements of rationality [...] can function as advice or guide your deliberation”. Bratman (2009a, 417) goes even further when he claims that the capacity to be guided by (wide-scope) requirements of instrumental rationality and intention consistency is essential to our general planning capacity. See also Reisner (2009, 245): “Rationality must have the resources to guide the agent to some particular correct mental state(s), when an agent is in a position such that rationality requires a change in mental states.”

A natural worry is that due to their disjunctive structure, wide-scope requirements never determine a particular option as the rational one, and thus leave us like Buridan's ass, who famously dies of hunger because he is unable to make a rational decision between two equally big and delicious piles of hay. It is unclear, however, whether it really constitutes an objection to the wide-scope account that it does not determine particular responses as rationally required, but leaves it to the discretion of the agent as to which part of the disjunctive response he gives. I think there is a serious problem if the wide-scope account is understood as the view that all requirements of rationality are structural, and all structural requirements of rationality take wide-scope.<sup>53</sup> For it seems that disjunctive structural requirements of rationality never combine in such a way that a particular option turns out to be more rational than another one.<sup>54</sup> So if all rational requirements are structural and disjunctive, agents will always be confronted with at least two equally rational options. And this view, I think, threatens the very idea of a rational decision. It entails that rationality can *never* determine a particular option as required, that there are *always* two ways of avoiding incoherence, and that these two ways are rationally on a par. This is counterintuitive, to say at least. Common sense surely tells us that *some* decisions for particular options are

---

<sup>53</sup> This, I think, is roughly Broome's conception of rationality, cf. Broome (2005a, 323): "it is very plausible that rational requirements generally have a wide scope." The only example of one of the "rare exceptions" that Broome allows for non-disjunctive requirements of rationality is the requirement not to have a contradictory belief, and such requirements not to have particular incoherent attitudes would not solve the problem I mention in the main text. Brunero (2010a, esp. §4) takes a different route, holding that there are "global" non-disjunctive requirements of rationality (while global requirements "specify what rationality requires of us in the light of our entire set of attitude-states", p. 33), though he does not spell out how these requirements fit into his general wide-scope framework. Similarly, Way (2011, 230) defends a wide-scope view while holding that we are "sometimes rationally required to have particular attitudes", but he, too, says nothing about how to understand these occasional non-disjunctive requirements within the framework of a wide-scope account.

<sup>54</sup> This is due to their content, not due solely to the fact that they are disjunctive. As Joseph Barnes pointed out to me, disjunctive requirements *can* in combination determine a particular response. Suppose that A is under the three rational requirements to (1)  $\phi$  or  $\psi$ , (2)  $\psi$  or not- $\pi$ , and (3)  $\pi$  or not- $\phi$ . It follows that rationality determines a particular response, namely  $\psi$ -ing, as rationally necessary. For suppose that A does not  $\psi$ . Then, in order to satisfy (1), A has to  $\phi$ , and in order to satisfy (2), A has to not- $\pi$ . But if A  $\phi$ -s and does not  $\pi$ , then A violates (3). However, while it is conceptually possible that disjunctive requirements exhibit this structure, there are no plausible candidates of wide-scope requirements of rationality that instantiate it. This should not surprise us. In contrast, it would be remarkable if disjunctive requirements that are concerned with the avoidance of incoherence were, when taken together, to entail that particular attitudes (with simple content) are rationally required or prohibited.

more rational than others, and that it is *sometimes* rational to avoid incoherence in one way rather than another.

While I think that this would be a deeply problematic consequence of the wide-scope account, we have already seen that it follows only on the assumption that rational requirements generally have wide-scope. One need not take this radical view in order to maintain that *structural* requirements of rationality have wide-scope, or that the particular requirements which correspond to structural irrationality claims such as (A)-(D) do. So even though I think it poses a challenge for wide-scope advocates, the problem mentioned does not amount to an objection to the existence of particular wide-scope requirements.

It would, I think, constitute an objection to particular wide-scope requirements if their disjunctive nature posed a fundamental obstacle to guidance. Andrew Reisner seems to hold this view when he claims that “rational requirements must be fully determinate to be guiding”, while by ‘determinate’, he means “that the changes required are not disjunctive options amongst multiple mental states”.<sup>55</sup> This claim, however, strikes me as unwarranted. There is no general reason why only non-disjunctive requirements could be guiding. Recall Frank who ought to take the bus or the cab in order to arrive at the station in time. Of course, this requirement cannot guide Frank to take the bus *rather* than the cab, but the fact that there is no reason favouring one over the other should not make it impossible for him to be guided by this requirement. “It would be good to arrive at the station in time, so I shall call for a cab” seems a perfectly appropriate piece of practical reasoning in Frank’s situation. The fact that he could have concluded with equal right with an intention to take the bus does not, I think, pose a problem here.

### *The commonality condition*

We need, however, to take a closer look at what it means to be guided by a disjunctive requirement. Generally speaking, in order for a response to be guided, the agent who gives the response must think of something as counting in favour of the response; he

---

<sup>55</sup> Reisner (2009, 249).

must give the response in the light of a consideration that he takes to support it. One might be inclined to say that, in order for a response to be guided by a *requirement*, the guiding consideration must be that the response is required by it. This, I think, would be too strong. It seems to me that an agent can count as being guided by a moral requirement, for example, even if his guiding consideration is not the fact that he is required to do something, but rather one of the considerations that support the requirement itself, as for example the fact that the required action would help someone in need. So I shall allow for the possibility of an agent being guided by a requirement even if his guiding consideration is not the consideration that the response is required.

So let us consider particularly the idea of being guided by a *disjunctive* requirement. Since one can satisfy such a requirement by giving either of the responses that are part of the disjunctive response, the question is under which conditions we can say that an agent's  $\phi$ -ing is guided by a disjunctive requirement to  $\phi$ -or- $\psi$ . My aim is not to give an account of this, but only to argue for a necessary condition: In order for an agent's  $\phi$ -ing to be guided by a disjunctive requirement to  $\phi$ -or- $\psi$ , it must be true that the consideration that guided him in  $\phi$ -ing could also have guided him in  $\psi$ -ing. In other words, the guiding consideration in virtue of which someone counts as being guided by a disjunctive requirement must be capable of guiding each of the responses that are part of the disjunctive response. Since this condition requires a *common consideration* that can guide each of the responses that are part of the disjunctive response, we can call it the *commonality condition*.

Consider Frank. We are assuming that he is required to take-the-bus-or-take-the-cab. Suppose he takes the bus, and in doing so is guided by this disjunctive requirement. Then the commonality condition tells us that the consideration that guided Frank in taking the bus could also have guided him in taking the cab. And indeed, this seems to be the case. There is a common purpose that supports each of the responses, namely getting to the station in time, and this provides a consideration that could guide both parts of the disjunctive response. This common consideration makes it comprehensible why, in giving only one of the responses that are part of the disjunctive response, Frank can still be regarded as being guided by the disjunctive requirement. For suppose, to the contrary, that the consideration which guides Frank to take

the bus could not also guide him to take the cab. We may suppose that Frank's taking the bus is guided by the consideration that the bus is cheaper, or faster, or that he just likes taking the bus. It is clear that in such cases, Frank would not be guided by a disjunctive requirement to take-the-bus-or-take-the-cab (possibly, though not necessarily, he could count as being guided by a non-disjunctive requirement to take the bus). A consideration that can guide a person in  $\phi$ -ing, but not in  $\psi$ -ing is independent of a disjunctive requirement to  $\phi$ -or- $\psi$ , it thus cannot be the consideration in virtue of which this person can count as being guided by the requirement.

*A general guiding consideration?*

I shall now argue that structural wide-scope requirements of rationality cannot meet the commonality condition, and are thus incapable of guiding agents. Let me start with the possibility that there is a *general* consideration in virtue of which persons could be guided by wide-scope requirements. Since such requirements are supposed to hold in all possible worlds (or, on the conditional account, in all possible worlds in which the agent is in the antecedent state), this is a natural starting point. A possibility mentioned above is that what guides agents when they are guided by a wide-scope requirement of rationality, is just the consideration that the response is required by rationality. A related suggestion is that agents could be guided by wide-scope requirements in light of the ideal of internal coherence. Finally, it has been argued that the value of autonomy provides the guiding consideration we are looking for.<sup>56</sup>

None of these suggestions, I believe, meet the commonality condition, i.e. the condition that the guiding consideration of a disjunctive requirement must be capable of guiding each of the responses that are part of the disjunctive response. Consider the

---

<sup>56</sup> For this latter suggestion, see Bratman (2009a, 433). I use the term "autonomy" instead of his preferred term "self-governance" in order to avoid the risk of confusion with what I have called the "self-governance requirement". Bratman is concerned with the instrumental requirement (and a more general requirement of intention consistency) only, and he focuses on the question of what objective reasons we have to satisfy this requirement, rather than the question of which considerations, provided that agents subjectively *see* them as reasons, could possibly guide them towards satisfying it. A response to the first question, however, can always be taken as a response to the second question as well: If there is a general reason in favour of satisfying structural requirements of rationality, then it should also be possible that this reason can guide agents to satisfy them.

first proposal. Despite being unusual, it seems possible that an agent's intention is guided by the consideration that rationality requires him to not-believe-that-one-ought-to- $\phi$ -or-intend-to- $\phi$ . It is not possible, however, that an agent's disbelief that he ought to  $\phi$  is guided by the consideration that rationality requires him to not-believe-that-one-ought-to- $\phi$ -or-intend-to- $\phi$ . As Richard Moran, among others, has emphasised, the considerations that guide an agent's deliberation about whether to believe  $p$  are "transparent" to considerations about whether  $p$  is the case.<sup>57</sup> Consequently, an agent's belief-state (i.e. belief or disbelief) can be guided only by considerations that have to do with the content of the belief in question, considerations that provide evidence for or against it.<sup>58</sup> The fact that rationality requires one to not-believe-that-one-ought-to- $\phi$ -or-intend-to- $\phi$  does not provide any evidence against its being the case that one ought to  $\phi$ ; it therefore cannot guide one in dropping a belief with this content or continuing to disbelieve that one ought to  $\phi$ . And this point generalises for every wide-scope requirement we have considered so far: each of these requirements can be satisfied by a belief-state, and in no case does it seem that the fact that rationality requires one to be-in-this-belief-state-or-some-other-state provides any kind of evidential support for the content of the belief-state.

The same argument shows that neither the ideal of internal coherence nor the value of autonomy can provide the consideration that is needed in order for a wide-scope requirement to be guiding. For this to be the case, the consideration must be capable of guiding an agent in adopting the belief-state that is part of the disjunctive response of each wide-scope requirement. It can do so only by providing evidential support for this belief-state. But considerations to the effect that certain disjunctive responses are necessary for (or enhance) coherence or autonomy do not provide evidential support for the belief-states in question, for they are entirely independent of their content. Indeed, it is difficult to see how *any* general consideration to the effect that a

---

<sup>57</sup> See e.g. Moran (1988, 143): "within the perspective of the first-person, deciding what to believe is a matter of determining what is true."

<sup>58</sup> Of course, a belief-state could be *caused* by non-evidential considerations, as it is the case in wishful thinking. Also, an agent might be *guided* by a non-evidential consideration to *bring about* a belief-change, e.g. by taking a belief-changing pill in order to receive certain benefits. None of this casts doubt on the claim that a belief-state itself can be guided only by evidential considerations.

disjunctive response would promote a certain aim or value could provide evidential support for the belief-state.

This argument echoes the one given in Chapter 5, which showed that several proposals to substantiate the normativity of rationality by citing values or rationality itself as reasons for satisfying structural requirements of rationality are doomed to fail. I am here concerned only with the subjective side of the coin, however – i.e. with the question of which considerations, provided that agents *see* them as reasons, could possibly guide them towards satisfying a structural requirement of rationality. Even though the failure to find normative reasons for satisfying structural requirements concerns both the wide- and the narrow-scope account of such requirements, the subjective problem of finding a *possible* guiding consideration can arguably be resolved by the narrow-scope account. The general form of a narrow-scope requirement claim is ‘If A is in the antecedent state, then rationality requires A to be in the consequent state’. As we have seen above, the antecedent and the consequent state are such that one can rationally base the consequent state on the antecedent state. It thus seems natural to say that the consideration that guides an agent when he is guided by a narrow-scope requirement is given by the content of the antecedent state. An agent’s intention to  $\phi$  could, for example, count as being guided by the narrow-scope self-governance requirement, if it is guided by the agent’s belief that he ought to  $\phi$ .

I shall not discuss whether this approach is viable here.<sup>59</sup> What is important for the purpose of this section is that it is not open to wide-scope advocates. As argued above, the guiding consideration in virtue of which an agent’s response counts as being guided by a disjunctive requirement must be capable of guiding each of the responses that are part of the disjunctive response. The consideration that one ought to  $\phi$  might provide an agent with guidance for intending to  $\phi$ , but it cannot provide him with

---

<sup>59</sup> Objection 1: An agent is guided by a requirement only if the consideration that guides him supports the requirement itself, but the consideration that one ought to  $\phi$  does not support the claim that if one believes one ought to  $\phi$ , then one is required to intend to  $\phi$ . Possible response: The consideration that one ought to  $\phi$  does not support this conditional claim, but it does indeed support the narrow-scope requirement, which is a requirement to intend to  $\phi$ . Objection 2: This conception of guidance might work in the case of self-governance, but not in other cases like e.g. the instrumental requirement. The content of an intention to  $\phi$  and a belief that  $\psi$ -ing is a necessary means to  $\phi$ -ing does not provide a guiding consideration for intending to  $\psi$ . Possible response: Intending something involves seeing something as a reason for doing it, the content of an intention to  $\phi$  therefore *does* support intending the necessary means to  $\phi$ .

guidance for *not* believing that one ought to  $\phi$ . Generally speaking, the content of an antecedent state might provide a consideration that could guide one in adopting the consequent state, but surely not in dropping the antecedent state.<sup>60</sup>

*A situation-specific guiding consideration?*

So there is neither a general consideration in virtue of which agents could be understood as being guided by structural wide-scope requirements of rationality, nor can the contents of the antecedent states provide such guiding considerations. What remains is the possibility that particular situations provide considerations that can guide agents to give the disjunctive response wide-scope requirements call for. Now, it is difficult to see how such cases could be understood as cases of being guided by requirements of *rationality*. With respect to the suggestions considered so far, it seems possible to come up with an explanation why being guided in this way could count as being guided by rationality. This is obvious in case of the consideration that the response is rationally

---

<sup>60</sup> In this context, it is worth mentioning that Way's "medium scope" account of the instrumental requirement escapes the guidance problem. According to Way (2010a), the instrumental requirement can be satisfied by intending the means or giving up the end, but *not* by giving up the means/end-belief; it has the general form: 'If A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, then rationality requires A to [not intend to  $\phi$  or intend to  $\psi$ ]. In keeping with Way, we may say that the consideration that guides agents when they are guided by the instrumental requirement is the consideration that  $\psi$ -ing is a necessary means to  $\phi$ -ing. And this consideration seems to meet the commonality condition: It could guide an agent in intending to  $\psi$  as well as in discarding an intention to  $\phi$ . This poses the question of whether the wide-scope account could not generally be defended in this fashion. The answer is "no", however. Way's strategy essentially consists in subtracting from the possible satisfaction states of a rational requirement the option to revise the antecedent belief-state. If we do this in the case of the self-governance or the belief-guidance requirement, what we get are simply the narrow-scope versions of these requirements. The medium scope account therefore does not provide a general answer to the guidance problem for wide-scope requirements. Second, the medium scope account is questionable even in the particular case of instrumental rationality. I think it allows for implausible bootstrapping. Suppose, for example, A believes that not- $\phi$ -ing is a necessary means to  $\phi$ -ing. According to Way, A is now rationally required to [not- $\phi$  or not- $\phi$ ], which just means she is required to not- $\phi$ . The medium scope account thus implies that we can very easily generate a rational prohibition on a particular response, simply by adopting an absurd means/end-belief. In other cases, the account allows bootstrapping of implausible disjunctive requirements. Suppose you are rationally intending to impress your date, and you are rationally not-intending to dress as a parrot and tell a lot of sexist jokes when you meet up with one another. Next, you irrationally come to believe that dressing as a parrot and telling a lot of sexist jokes is a necessary means for you to impress your date. According to the medium scope account, it now follows that you are rationally required [to intend to dress as a parrot and tell a lot of sexist jokes *or* not to intend to impress your date]. But this is implausible, it seems you can rationally intend to impress your date and not intend to dress as a parrot, if you just give up your irrational means/end-belief. So not only is the medium scope account incapable of providing a general answer to the guidance challenge for wide-scope requirements, it also fails to give a plausible bootstrapping-avoiding answer in the particular case of the instrumental requirement.

required itself. As regards coherence and autonomy considerations, the idea seems to be that the point of rationality, in the eyes of those who are guided by its requirements, is tied to the ideals of coherence or autonomy. And if the considerations are supposed to be provided by the antecedent states, then it can be pointed out that people are guided by rational requirements simply in all cases of reasoning or rational attitude-change. Neither of these ideas is available if the consideration that guides an agent in giving the response a wide-scope requirement calls for is provided by the particular situation. I will nevertheless consider this possibility here. I shall argue that even if we allow the guiding consideration of a wide-scope requirement of rationality to be provided by the particular situation, the prospects of finding any plausible example of an agent being guided by such a requirement are extremely dim.

The first problem is that the responses that are part of the disjunctive response that structural wide-scope requirements of rationality call for are normally supported by very different, if not contrary, considerations. Consider the requirement to [not believe that one has sufficient evidence that  $p$  *or* believe that  $p$ ]. In order for a consideration to provide an agent with guidance in giving each of these responses, the consideration would have to provide evidential support for  $p$  as well as evidential support against its being the case that there is sufficient evidence that  $p$ . It seems to me that there simply is no consideration that could possibly meet these criteria. One and the same consideration cannot provide sufficient evidence in order to guide an agent in believing  $p$ , while at the same time providing sufficient evidence in order to guide an agent in *not* believing that he has sufficient evidence that  $p$ .

In cases of other requirements it seems possible, but certainly unusual, that one and the same consideration could guide each of the responses. Consider the requirement to [not believe that one ought to  $\phi$  *or* intend to  $\phi$ ]. Ordinarily, considerations that can guide agents in *adopting* the belief that one ought to  $\phi$  can also provide guidance for forming the intention to  $\phi$ . The considerations guiding the *dismissal* of a belief that one ought to  $\phi$ , however, would normally be different, if not contrary, to the considerations that can guide the formation of an intention to  $\phi$ . Any such case would be atypical.

Let us consider such an atypical case, in which a consideration that provides evidence *against* believing that one ought to  $\phi$  at the same time counts in *favour* of intending to  $\phi$ . Sebastian is going for lunch with his colleagues, one of whom told him that the soup of the day is a bouillabaisse. Sebastian believes that he ought to eat more fish, because it contains a lot of healthy unsaturated fatty acids. For this reason he comes to believe that he ought to order the soup of the day. But since Sebastian does not like fish very much, he has not yet formed the intention to order the soup of the day. Arriving at the restaurant, Sebastian learns that the soup of the day is not a bouillabaisse, but a delicious lentil soup that he has had before. This consideration weakens his evidence that he ought to eat the soup of the day and can thus guide him in dropping his belief. At the same time, it can guide him in adopting the intention to order the soup of the day.

So here we have an example of a consideration that could guide both the disbelief that one ought to  $\phi$  and the intention to  $\phi$ . It seems to me, however, that this is not plausibly a case in which an agent should be seen as being guided by a *disjunctive* requirement. If Sebastian's disbelief is guided by a requirement in virtue of a consideration that evidentially supports his disbelief, it seems absurd to think that this very requirement could *alternatively* be satisfied by the formation of an intention. What is going on in Sebastian's case is that one consideration independently supports two different responses rather than one disjunctive response.

Again, this point seems to generalise. Whenever an agent's belief-state is guided by a requirement in virtue of a consideration that evidentially supports this particular belief-state, it seems very odd to think that the agent could alternatively satisfy this very requirement by adopting some other attitude-state. And if this is true, then *agents can never be guided by a requirement on a disjunctive response, one part of which is a belief-state*. For in order to be guided by such a requirement they have to be guided by it in virtue of a consideration that can guide each of the responses. In order to guide the belief-state, this consideration has to evidentially support the belief-state. And if an agent's belief-state is guided by a requirement in virtue of a consideration that evidentially supports this belief-state, this requirement cannot plausibly be satisfied *alternatively* by some other attitude-state. We can thus conclude that, since *all* wide-scope require-

ments of rationality considered here can be satisfied by a belief-state, none of them can be guiding.

### *Conclusion*

This discussion, I think, casts light on how awkward structural wide-scope requirements of rationality really are. Outside the realm of structural rationality, where wide-scope principles have been introduced to solve certain theoretical problems, there seems to be no example of a requirement on a disjunctive response, the parts of which can possibly (or at least apart from very atypical cases) be supported only by different, if not contrary, considerations. Nor is there any plausible example of any other requirement that can be satisfied by adopting a belief-state or *alternatively* some other attitude-state.

Another point needs emphasis: We think of weakness of will or means/end-incoherence as *practical* forms of irrationality; it is thus natural to think of the corresponding requirements of rationality as *practical* requirements as well. As it turns out, however, on the wide-scope account these requirements are not really practical requirements. Genuinely practical requirements cannot be satisfied by adjusting belief-states only, but the wide-scope requirements at issue *can* be satisfied by adjusting belief-states only. I think this poses the question of whether we really understand such requirements in the first place. I have argued that there is no reason to doubt the existence of disjunctive requirements in general, but I think it is difficult to understand a requirement on a disjunctive response, one part of which concerns beliefs and the other intentions. Such requirements are very peculiar hybrids of theoretical and practical requirements, which plough a lonely furrow: again, there seems to be no plausible example of such a hybrid requirement outside the realm of structural rationality. We can understand what it means to say that an agent is either rationally required to intend some action *or* rationally required to hold or withdraw a belief, and we can understand that an agent is rationally required to give the disjunctive response of intending one thing *or* intending another, but the idea that rationality requires us to give the disjunc-

tive response of intending an action *or* holding or withdrawing a belief is, on reflection, very difficult to make sense of.

In any case, I have argued that no such requirement would be capable of guiding an agent in giving the responses it calls for. This alone gives us sufficient reasons to doubt the existence of such requirements.

## 6.7 The myth of structural requirements of rationality

Let me take stock of this discussion of structural requirements of rationality. I have considered various objections to both narrow- and wide-scope versions of such requirements. As we have seen, the narrow-scope account licenses implausible bootstrapping of rational requirements, and it allows for rational requirements that are inconsistent and undermine each other in counterintuitive ways. These, I think, are more than sufficient reasons to reject the narrow-scope account. The wide-scope account avoids these problems, but it faces others. First, wide-scope requirements pose a symmetry between two ways of satisfying them, whereas the rational processes that correspond to the requirements in question are asymmetrical. The wide-scope account thus cannot account for a plausible connection between the satisfaction states that a rational state-requirement predicts and the rational processes an agent can undertake to reach these states. Second, and perhaps more significantly, due to their peculiar disjunctive form, wide-scope requirements do not meet the plausible condition for rational requirements that they are capable of guiding agents and have other awkward properties that are difficult to make sense of.

Where does this leave us? The short history of the debate about the scope of structural requirements of rationality features two general kind of arguments. The first is that since the narrow-scope account is false, the wide-scope account must be correct.<sup>61</sup> The second is that since the wide-scope account is false, the narrow-scope account is correct.<sup>62</sup> The discussion of this chapter suggests that the arguments against both accounts are quite strong. So perhaps it is time to conclude that neither is correct.

---

<sup>61</sup> See e.g. Broome (1999) and Brunero (2010a).

<sup>62</sup> See e.g. Kolodny (2005, 518–530) and Schroeder (2004).

Thus, my conjecture at this point is that since no plausible interpretation of structural requirements of rationality can be given, we should reject the assumption that there are any such requirements – the idea that there are is a myth.

The problems that occurred with each of the accounts themselves suggest an alternative picture. Narrow-scope requirement claims state that in every possible world in which an agent is in the antecedent state, he is rationally required to be in the consequent state. This means, at least on the diachronic account, that the agent's subsequent refusal to adopt the consequent state constitutes the violation of a rational requirement. This violation claim seems too strong; it seems that there are possible worlds in which an agent's rationality is preserved despite his refusal to adopt the consequent state, provided he gives up the antecedent state instead. This is what the wide-scope account gets right. However, the wide-scope advocates avoid this implausible violation claim only at the cost of maintaining an implausible satisfaction claim. The wide-scope account entails that in all possible worlds in which agents drop the antecedent state, these agents thereby come to satisfy the requirement, even though in many of these worlds, there are no rational processes available that could lead them to drop this attitude, and even though in those cases in which there *is* such a process available, this process is not a satisfaction process of the requirement in question, but the satisfaction process of another requirement that is concerned with quite a different potential conflict among attitude-states. Moreover, the wide-scope account posits requirements of a peculiar disjunctive structure that deprives them of any opportunity to guide agents.

It thus seems that both narrow- and wide-scope requirements over-generalise. Not in every world in which an agent is in the antecedent state and refuses to adopt the consequent state is he violating a rational requirement (as the narrow-scope account entails); but nor does he satisfy a rational requirement in every world in which he gives up the antecedent attitude (which is entailed by the wide-scope account). Rather, in some worlds, agents are rationally required to adopt the consequent attitude, and in others they are required to drop the antecedent attitude. This suggests that whatever explains why combinations of attitude-states such as the ones depicted in (A)-(D) are necessarily irrational, is neither a non-disjunctive requirement to be in the consequent

state (if one is in the antecedent state), nor is it a disjunctive requirement to [be in the consequent state *or* not be in the antecedent state]. What explains the appearance of structural irrationality might instead be a disjunction *of* requirements, such that necessarily, agents are rationally required to be in the consequent state *or* rationally required not to be in the antecedent state.<sup>63</sup> These requirements would plausibly be ordinary non-disjunctive requirements, which would avoid the peculiarities of the wide-scope account and ensure that the requirements in question can be guiding. The remainder of this thesis is concerned with the exploration and development of this general idea.

---

<sup>63</sup> See Kolodny (2007a) and (2008a) for an account of *first-order* structural requirements along these lines. My aim is to defend a more general approach, which covers both first- and second-order structural requirements, and to do this in a way that vindicates rather than undermines the normativity of rationality. See Chapter 9.1 for how my account relates to Kolodny's.

## Rationality as responding correctly to reasons

---

The discussion of the relation between structural rationality and normative reasons in the last six chapters motivates the conjecture that there are no structural requirements of rationality (7.1). The aim of the next four chapters is to outline an alternative account of rationality and explain the appearance of structural irrationality without adherence to such requirements. This chapter is concerned with the first of these tasks; it provides the sketch of an account of rationality according to which rationality requires us to respond to available reasons, where a reason is available if it is part of an agent's body of evidence (7.2-7.3). I argue that this response conception of rationality is compatible both with the factivity of evidence and the thesis that rationality supervenes on the mind (7.4-7.5), and defend the notion of a practical evidence-based reason against the objection that all practical reasons must be good-making features (7.6). The remainder of the chapter is concerned with defending some more specific response requirements of rationality, such as the requirement to believe what one's evidence supports as true (7.7), and the requirement to intend what one's evidence suggests one ought to do (7.8-7.9).

### 7.1 Taking stock: structural rationality and normative reasons

Let me begin this chapter by summarising where the discussion of the relation between structural rationality and normative reasons in the last six chapters has left us. I began, in Chapter 1, by describing the basic problem as a tension between three claims:

- (1) There are structural requirements of rationality.
- (2) Rational requirements are normative.
- (3) Structural requirements of rationality (if there are any) are not normative.

I also described three possible views that one can take with respect to this tension: the *structuralist view*, which denies the arguments in favour of (3); the *mixed view*, which denies the normativity of rationality (2); and the *anti-structuralist view*, which denies the existence of structural requirements of rationality (1).

In Chapter 2, I argued that our practice of rational criticism essentially involves the assumption that rationality is normative, and that giving up (2) therefore amounts to a radical error theory that undermines this practice. Proponents of the mixed view thus have to bear considerable theoretical costs. After an examination of the logical form and content of structural requirements in Chapter 3, I turned to claim (3) in Chapters 4 and 5, arguing that the normativity of structural requirements of rationality is untenable. This is why I think that the structuralist view fails. Over and above all that, I argued in Chapter 6 that the assumption of structural requirements of rationality itself poses several severe problems quite independently of whether they are understood normatively or not. This provides additional reasons for thinking that the structuralist view and the mixed view, which are both committed to such requirements, are false.

I conclude that the anti-structuralist view, which has not yet been pursued as a solution to our problem, should be taken very seriously. The anti-structuralist view embraces the normativity of rationality (2) at the cost of denying that there are any structural requirements of rationality (1). Defending this view involves two tasks. The first of these tasks is to present a positive account of what rationality consists in that is compatible with the normativity of rationality. The second task is to explain the phenomenon of structural rationality in terms of this account rather than in terms of structural requirements of rationality. This chapter takes up the first of these tasks, by providing the sketch of an alternative account of rationality, defending it against objections, and discussing some of its implications.

## 7.2 A fresh start: rationality as responding to reasons

The slogan for the positive account is simple: rationality consists in responding correctly to reasons. That is, the capacity of rationality consists in the capacity to respond to reasons, and requirements of rationality consist in requirements of reasons. This idea

is not new; it is present, more or less explicitly, in many philosophical approaches to rationality.<sup>1</sup> The slogan needs to be filled in, of course, with a more detailed account of the ways in which responding to reasons is a matter of rationality. But for the moment, it provides an intuitive starting point that promises a satisfying response to several desiderata that the discussion so far has revealed.

First, the account could provide an answer to both the question of the *content* and the question of the *authority* of rational requirements. If rationality consists in responding correctly to reasons, then rational requirements could be understood as inheriting both their content and their authority from the content and authority of the relevant reasons. Second, this account of the authority of rationality would provide an elegant response to the why-be-rational challenge. As we have seen, the prospects of finding a particular reason to satisfy any rational requirements look dim. But if rationality consists in responding to reasons, then no reason to be rational is needed in order for rationality to be normative. Recall the attraction of a *buck-passing account* of requirements of morality or prudence that I have emphasised in Chapter 5. Such requirements, I have claimed, are best understood as statements to the effect that there are reasons of a particular kind, and are thus in no need of a *further* reason that counts in favour of satisfying the requirement. I have argued above that rational requirements cannot be understood as statements about the existence of reasons of a particular kind, because there are no reasons of a particular kind that merit the name “rational reasons”. However, if rationality consists in responding to reasons, we might be able to understand rational requirements as verdictive statements not about reasons of a particular *kind*, but about reasons of *all* kinds, or reasons of all kinds that stand in a certain relation to an agent. In this way, the why-be-rational challenge could be elegantly dismissed.

Last but not least, understanding rationality as consisting in responding to reasons, and explaining the phenomenon of structural irrationality in these terms, would provide a *unified* account of rationality. It would tell us what holds different

---

<sup>1</sup> “To say it is rational to do something is to say that doing it is supported by the preponderance of reasons” (Gibbard 1990, 161). “Rationality involves taking account of reasons” (Nozick 1993, 72). “An account of rationality is an account of the capacity to perceive reasons and to conform to them” (Raz 1999a, 68).

aspects of rationality together, and it is not clear whether proponents of the structuralist or the mixed view are able to do that. To the contrary, they seem committed to maintaining a deep division between two very different kinds of rationality (or, alternatively, to an implausible denial of all non-structural requirements of rationality).<sup>2</sup>

In his article “Does rationality consist in responding correctly to reasons?”, John Broome presents an objection to the rationality conception at issue. His argument is that you could have all things considered most reason to do something, while being inculpably ignorant of these reasons because you do not have any evidence for them. In such a case you could be rational without responding correctly to your reasons. His example is this:

The fish on the plate in front of you contains salmonella. This is a reason for you not to eat it, and let us assume all your reasons together require you not to eat it. But you have no evidence that the fish contains salmonella. Then you might eat it even though your reasons require you not to, and nevertheless you might be rational.<sup>3</sup>

In other words, since you can fail to respond correctly to your reasons and still be fully rational, rationality cannot consist in responding correctly to reasons. Broome calls this the “quick objection”, and he mentions no further reasons for dismissing the conception of rationality at issue.<sup>4</sup> I believe his objection is in fact too quick.

First, the idea that rationality consists in responding to reasons need not be interpreted as strongly as Broome presupposes. In order to provide an answer to the question of the content and authority of rationality that evades the why-be-rational challenge, we need only assume that all rational requirements are, or derive from, requirements of reasons, not necessarily that all requirements of reasons issue requirements of rationality. To begin with, we might say that rationality requires us to respond to only those reasons that are *available* to an agent.

---

<sup>2</sup> See e.g. Kolodny’s distinction between “objective” and “subjective” rationality (2005, 510).

<sup>3</sup> Broome (2007d, 352)

<sup>4</sup> All further objections in Broome (2007d) concern related but importantly different views, such as the view that rationality consists in responding to *beliefs* about reasons.

What does it mean to say that a reason is available? I shall take for granted the idea, common in philosophical debates about theoretical rationality, that agents have at their disposal a body of evidence that provides reasons for beliefs.<sup>5</sup> Further, I shall assume that for something to be part of an agent's evidence, it is a sufficient condition that the agent knows it.<sup>6</sup> If you know that it is raining, then the fact that it is raining is among your evidence; it might be your evidence for believing that the streets are slippery, for example. It follows that facts or true propositions can be evidence (I remain neutral with respect to the question of whether phenomenal experiences or other entities that are not propositionally structured can constitute evidence as well). As propositions or facts that are part of an agent's body of evidence can figure as reasons for belief that are available in the relevant sense, the very same propositions or facts can also figure as practical reasons that are available in the relevant sense. If the fact that it is raining is among your evidence, it might provide you with an available reason to believe that the streets are slippery, but it might just as well provide you with an available reason to take an umbrella with you.

I shall call reasons that are available in this sense "evidence-based reasons". Against this background, the idea that rationality consists in responding correctly to reasons can be expressed as follows:

*Rationality as responding to reasons (RRR):*

*Intention:* If A has decisive evidence-based reason (not) to intend to  $\phi$ , then A is rationally required (not) to intend to  $\phi$ .

*Belief:* If A has decisive evidence-based reason (not) to believe that p, then A is rationally required (not) to believe that p.<sup>7</sup>

<sup>5</sup> See e.g. Kelly (2006, §1), Williamson (2000, 186–190).

<sup>6</sup> Famously, Williamson (2000, ch. 9) claims that this is not only a sufficient, but also a necessary condition.

<sup>7</sup> Some philosophers hold that requirements against beliefs are not based on reasons against beliefs, but instead on the *absence* of (sufficient) reasons *for* these beliefs; see e.g. Skorupski (2010, 39, n. 9). Others think differently; see e.g. Schroeder (unpublished). Though (RRR) would have to be modified if the former view were correct, this modification would provide no difficulty for the general account I am proposing.

R is an evidence-based reason for A to  $\phi$  if, and only if, R is a reason for A to  $\phi$  and A's body of evidence includes R. Whether A has decisive evidence-based reason to  $\phi$  is determined by her evidence-based reasons for and against  $\phi$ -ing (and possibly other evidence-based factors).<sup>8</sup> Cases in which we are rationally permitted to do what we have decisive *overall* reason not to do, such as Broome's salmonella example, are no counterexamples to a rationality conception along the lines of RRR. Yet such a conception is a meaningful interpretation of the view that rationality consists in responding to reasons.

Second, Broome's objection relies on the crucial assumption that what we have decisive overall reason, or ought all things considered, to do can depend on facts to which we have no access. As we shall see in Chapter 8, however, there are good reasons to doubt this assumption. On an alternative view, the notion of 'ought' or 'decisive reason' relevant in deliberation is itself constrained by an agent's epistemic perspective.<sup>9</sup> One way to spell out this alternative view is as follows:

*The evidence-relative view:* A ought (or has overall decisive reason) to  $\phi$  if, and only if, A has decisive evidence-based reason to  $\phi$ .

I shall defend this view in Chapter 8. For now, it is important to note that Broome's argument essentially relies on its being false. If what you have decisive reason to do is a function only of your evidence-based reasons, then you cannot have decisive reason to refrain from eating the fish because of some fact that you have no evidence for.

There is an important lesson to draw from Broome's objection. Debates on the normativity of rationality are often based on the assumption that there is a fundamental divide between what one might call the "subjectivity" of rationality and the "objectivity" of reasons. According to this dualistic picture, being rational is a matter of

---

<sup>8</sup> Here, I remain neutral as regards the question whether factors which are not reasons against  $\phi$ -ing, such as reasons to disregard reasons for  $\phi$ -ing or reasons not to act for certain reasons for  $\phi$ -ing, can defeat reasons to  $\phi$ , as Raz (1975a, ch.1; 1975b; 1990) claims. Note that the phrase "A has decisive evidence-based reason" is ambiguous between the reading I explicate above and a different reading, according to which it implies that A's evidence-based reasons are overall decisive.

<sup>9</sup> Ironically, Broome (2013, ch. 3) himself argues that the deliberative 'ought' is not to be understood independently of the epistemic limitations of the agent.

establishing coherence among subjective mental states, while conforming to reasons is a matter of establishing correspondence to objective facts.<sup>10</sup> Once this picture is accepted, it is indeed difficult to see how rationality could be normative.<sup>11</sup> But the picture is itself questionable. For one, rationality amounts to more than establishing coherence among our beliefs and intentions; it is also concerned with responding to available evidence. The remainder of this chapter is concerned with exploring and defending this intuitively plausible idea. For another, our reasons are not simply determined by the objective facts independently of our epistemic situation, but crucially depend on our evidence. This claim will be the focus of the next chapter. The hypothesis I am going to explore in the remainder of this thesis is as follows: *Once we see that rationality can take us beyond subjective coherence and that reasons are sensitive to the available evidence, we will also see that requirements of rationality and reasons coincide in important ways.* The remaining task is to show how the appearance of structural irrationality can be explained in terms of non-structural rational requirements, i.e. requirements to respond to available reasons, which will be taken up in Chapters 9 and 10.

### 7.3 Response rationality

The account of rationality that I would like to propose draws on a natural and widespread way of thinking about the relation between reasons and rationality. We human beings have a capacity to respond to available reasons. For example, we are capable of weighing our evidence for and against a certain hypothesis, forming beliefs that our evidence supports as true and refraining from forming beliefs that our evidence does not sufficiently support as true. We are also capable of perceiving available practical reasons, weighing them in deliberation and forming intentions that these reasons support. This capacity to respond to reasons is, or belongs to, our capacity of rationality. If we form beliefs that are not supported by our evidence, this is an error of our rational capacity. If we form intentions against our available practical reasons, this, too,

---

<sup>10</sup> “Rationality ... is a matter of whether our beliefs *cohere* with one another, not whether they *correspond* to the world” (Brunero 2009a, 322). “While reasons are given by the facts, what we can rationally want or do depends on our beliefs” (Parfit 2011a, 111). Cf. also Bridges (2009, 354).

<sup>11</sup> See also Way (2009, 2).

is an error of our rational capacity. Having beliefs and intentions that appropriately reflect the available reasons is something that can be expected of us *qua* beings with rational capacities.

These statements are close to being platitudes. In particular, that rational persons adjust their beliefs to the evidence is a commonplace in the philosophical literature.<sup>12</sup> Yet some authors prefer to reserve the term “rationality” for matters of subjective coherence only.<sup>13</sup> On one reading of this claim, it involves a fundamental form of scepticism: whether we conform to reasons is a matter of luck rather than capacity; all that we are capable of *qua* rational beings is establishing internal coherence. This view has far-reaching implications. If we do not have a capacity to respond to reasons, then we cannot be responsible for whether we conform to our reasons, and consequently, we cannot be worthy of praise or criticism for conforming or not conforming to reasons. Arguably, our having this capacity is a precondition of our being subject to reasons in the first place.<sup>14</sup> As with other fundamental kinds of scepticism, this view is difficult to refute (if such a refutation is possible at all). However, the next section addresses a main motivation for this view, which consists in the general idea that what is rational, or what we are responsible for, depends only on our mind, while reasons must be understood as depending in part on how the world is independently of our mind. If what I argue in the next section is correct, this general idea provides no objection to the response conception of rationality.

On a different reading of the claim that rationality is concerned with coherence only, we have a capacity to respond to available reasons, but this capacity is not part of

---

<sup>12</sup> See e.g. the entry on “evidence” in the *Stanford Encyclopedia of Philosophy*: “insofar as one is rational, one is disposed to respond appropriately to one’s evidence” (Kelly 2006, §2). Similar claims are made by other authors: “If you have no clear evidence that *p* is true [...], you are rationally required not to believe *p*” (Cullity 2008, 79). “It is irrational for one to believe that *p* just in case one lacks sufficient evidence for *p*” (J. Ross 2012a, 173). “Rational subjects have the capacity to grasp the evidence available to them and to form beliefs that accord with that evidence, the justification of subjects’ beliefs turns on whether or not they have exercised such evidence-grasping and belief-forming capacities as they have” (Smith 2004, 167). “Rationality requires one to conform one’s beliefs to one’s evidence” (Williamson 2000, 12).

<sup>13</sup> See e.g. Davidson (1985, 190), who holds that believing against one’s evidence is not irrational if one has false beliefs about one’s evidence.

<sup>14</sup> Cf. Raz, who maintains that reasons “are reasons because rational creatures can recognise and respond to them with the use of Reason” (2010a, 7). Kolodny (2011b, 333–341) seems to agree.

the capacity for rationality “narrowly construed”.<sup>15</sup> This is no objection to the view that the capacity to respond to available reasons is part of the capacity of rationality broadly construed. Call the narrow conception ‘coherence rationality’ and the broad conception ‘response rationality’. First, it is not clear whether we should think of the involved capacities involved as belonging to two different faculties rather than one. Plausibly, response rationality includes coherence rationality: responding to reasons involves not only forming beliefs and intentions that are directly supported by reasons, but also reasoning accordingly, forming and revising further attitudes such that they cohere with the supported attitudes, etc.<sup>16</sup> Second, we have seen the various problems connected with the idea that coherence rationality issues requirements on its own. Hence it is worth exploring whether the appearance of such requirements can be understood in terms of response rationality rather than coherence rationality. This is one way to understand the project I am pursuing.

It is sometimes argued that the verdict of irrationality applies only to instances of incoherence and is inappropriate in cases in which agents fail to respond correctly to available reasons while remaining internally coherent.<sup>17</sup> But this is questionable. Dogmatic people may be very eager to maintain their internal consistency and yet display drastic forms of irrationality in their refusal to take into account conspicuous evidence. Derek Parfit’s famous example of a man who is indifferent to pain on future Tuesdays may serve as a further illustration of this point with regard to practical reasons.<sup>18</sup> Preferring the prospect of an excruciating pain on a future Tuesday to a mild pain on any other day constitutes a severe form of irrationality, even though it might not involve any incoherent attitudes at all. The irrationality lies in nothing but a failure to respond to a particularly evident reason.

---

<sup>15</sup> See esp. Scanlon (1998, 25–30). But see also Scanlon (1998, 23): “A rational creature is, first of all, a reasoning creature – one that has the capacity to recognize, assess, and be moved by reasons”. The narrow construal of rationality is sometimes contrasted with “reason” (the mass term) or “reasonableness”.

<sup>16</sup> This does not presuppose the existence of coherence requirements, which the anti-structuralist view denies. Chapter 9 provides an explanation of how response rationality includes coherence rationality without presupposing the existence of coherence requirements.

<sup>17</sup> See e.g. Davidson (1985, 193); Foot (1972, 310); Scanlon (1998, 25–26).

<sup>18</sup> See Parfit (1984, 123–124). For a defence of the irrationality of future Tuesday indifference against Scanlon, see also Parfit (2011a, 118–125).

The point remains that not all cases of response failures are naturally described as cases of irrationality. The explanation for this is that rationality is a matter of degree, and that we tend to describe agents as irrational only if the rational failure is particularly salient.<sup>19</sup> In order to maintain the assumption of response requirements of rationality, we need, however, only to suppose that the violation of such requirements shows the agent to be *less than fully rational*, or open to rational criticism. And this strikes me as a plausible verdict when the neglected reason is part of the agent's evidence (as we assume it is).

In sum, it seems plausible to say that we have a capacity to respond to available reasons, which belongs to the rational faculty. Rationality is not only concerned with internal coherence, but also with responding to reasons, provided that the reasons are available. This justifies, at least *prima facie*, the assumption of response requirements of rationality along the lines of (RRR). Let us see whether this assumption stands to objections.

#### 7.4 The supervenience argument

This section is concerned with what I take to be the most important challenge for the very idea that we are capable of responding to reasons in virtue of possessing the rational faculty, and thus also to the assumption of response requirements of rationality. The challenge might be described as a conflict between two independently plausible assumptions about rationality and evidence. On the one hand, it seems a plausible principle that “rationality supervenes on the mind”, as John Broome puts it.<sup>20</sup> If A and B have exactly the same mental properties, and run through exactly the same mental processes, then they also do not differ with respect to whether their beliefs and intentions are rational or irrational.<sup>21</sup> On the other hand, evidence seems to be factive.

---

<sup>19</sup> See also Parfit (2011a, 56). One important reason why the rational failure is particularly salient in obvious kinds of internal incoherence, such as those involved in structural irrationality, is that such incoherence is particularly salient “second-order” evidence that at least one attitude of a particular group of attitudes cannot be sufficiently supported by reasons. See Chapter 9.3.

<sup>20</sup> Broome (2007d, 352). See also Wedgwood (2002).

<sup>21</sup> Since the rationality of an attitude can depend on the process by which it is formed, the supervenience base for rationality must include historical mental properties as well. I take this for granted in the following discussion.

If E is evidence for A that q, then E is true, or E is the case, or E occurs (in case E is not propositionally structured). Needless to say, in order for E to be evidence that q, q need not be true: one can have evidence for a false proposition. But nothing can be evidence for anything unless it occurs, is true, or is the case.<sup>22</sup> Moreover, it seems plausible that not all facts that are evidence are facts about the agent's internal state; ordinary external facts such as the fact that it is raining can be part of an agent's body of evidence. This raises a difficult question: How can rationality depend entirely on one's mental properties and at the same time require us to respond to our evidence, if what our evidence is partly depends on facts that obtain independently of one's internal state?

The challenge can be made more precise. If the property of rationality supervenes on the mind, then whether a person is rationally *required* to do something, too, must supervene on the mind; it must be insensitive to factors that make no difference to this person's mental properties:

- (1) *Rationality supervenes on the mind*: Whether A is rationally required to  $\phi$  depends only on A's mental properties.

The factivity of evidence in turn seems to support some kind of externalism about evidence, according to which whether something is part of an agent's body of evidence may depend on factors that make no difference to this agent's internal state.<sup>23</sup> Consider a typical sceptical scenario, which involves a 'good case' and a 'bad case'. Goodsmith is looking for a murderer in the forest and sees someone disappearing behind a tree. It is natural to assume that his evidence includes the fact that someone disappeared behind

<sup>22</sup> See e.g. Williamson (2000, 201).

<sup>23</sup> By an agent's 'internal state', I mean the totality of this agent's *non-factive* mental states, such as e.g. beliefs. Internalism and externalism might also be defined, presupposing a modest form of physicalism, in terms of supervenience on the internal physical state (cf. Williamson 2000, 49). Note that externalism about evidence in the defined sense does not by itself entail externalism about justification. For this conclusion, the following bridge principle is necessary: If whether A's evidence includes E may depend on factors that make no difference to A's internal states, then whether A is in a position to be justified in believing p may also depend on such factors. This principle is analogous to the one mentioned below in (5), substituting "internal state" for "mental properties" and "being in a position to be justified in believing p" for "having decisive evidence-based reason to  $\phi$ ". As I argue below, such a bridge principle can reasonably be questioned.

the tree. If he believes that the murderer is hiding behind the tree, we can correctly cite the fact that someone disappeared behind the tree as the evidence he has for his belief. Badsmith is in exactly the same internal state as Goodsmith, but while he has the same *appearance* of someone disappearing behind the tree, this appearance is non-veridical. Since evidence is factive, Badsmith's evidence does not include the proposition that someone disappeared behind the tree. Hence:

- (2) *Externalism about evidence*: Whether A's evidence includes E may depend on factors that make no difference to A's internal state.

Next, it is a common view that a person's mental properties depend only on her internal states:

- (3) *Internalism about the mind*: A's mental properties depend only on A's internal state.

Granting these three assumptions, the following argument can be made against the view that rationality requires us to respond to evidence-based reasons:

- (4) Whether A's evidence includes E might depend on factors that make no difference to A's mental properties (*from 2 and 3*).
- (5) If whether A's evidence includes E might depend on factors that make no difference to A's mental properties, then whether A has decisive evidence-based reason to  $\phi$  might also depend on such factors (*bridge principle*).
- (6) Therefore, whether A has decisive evidence-based reason to  $\phi$  might depend on factors that make no difference to A's mental properties (*from 4 and 5*).
- (7) Therefore, A might have decisive evidence-based reason to  $\phi$  and yet not be rationally required to  $\phi$  (*from 1 and 6*).<sup>24</sup>

---

<sup>24</sup> Note that 'ϕ-ing' in this argument refers only to mental actions or states.

Let us call this the *supervenience argument*. I have framed it in terms of evidence-based reasons, but it seems to me that a similar argument can be made against any particular conception of rationality according to which rationality is concerned with responding to reasons. The supervenience argument touches on some fundamental issues in both epistemology and the philosophy of mind which I cannot address here in a fully satisfying way. I will confine myself to pointing towards some possible replies and defending in some more detail the one that seems to me most reasonable. Given that the argument threatens the response conception of rationality as a whole, and given the failure of other conceptions of rationality discussed in earlier chapters, the alternative may just be to accept a fundamental scepticism about rationality. And granting that the existence of reasons is bound to the existence of a rational capacity to respond to them, this scepticism might ultimately call into question the existence of reasons itself. So those who do not find such a fundamental scepticism an attractive theoretical option should be willing to consider seriously the possibility that – perhaps despite first appearances – this argument is based on at least one false premise.

I shall not question the assumption that rationality supervenes on the mind (1). Externalism about evidence (2) is denied by some internalists about justification, but the argument for it is straightforward. The internalist about evidence must deny either that evidence is factive, or that one's evidence can ever include a proposition about the external world.<sup>25</sup> In the first case, she confuses evidence with apparent evidence, with

---

<sup>25</sup> For example, R. Feldman's view that "S has p as available evidence iff S is currently thinking of p" (1988a, 232) denies externalism about evidence by denying the factivity of evidence. Similarly, according to the view that Conee calls "seeming evidentialism" (and which he thinks is a "credible version of evidentialism"), "positive evidence is supplied by seeming truth. The general idea is that someone's evidence about a proposition includes all that seems to the person to bear on the truth of the proposition. ... our initial evidence ... consists in how things initially seem to us" (2004, 15–16). Since according to this view, the evidence is not the seeming itself, but what seems to be the case (even if in fact it is not), it is incompatible with the factivity of evidence. See also Schroeder (2011a), who maintains that one has p as evidence if one has a presentational attitude towards p. Huemer's preferred version of justificational internalism, according to which "all of the conditions which confer justification supervene on how things seem to the subject" (2006, 148) might also be taken to deny externalism about evidence. This is because the condition that A's evidence includes E is a *pro tanto* justification-conferring condition. It should be noted, however, that Huemer counts as justification-conferring conditions only non-redundant parts of a sufficient condition for justification, whereas the condition that A's evidence includes E might be both redundant and insufficient for justification. Consequently, his thesis does not by itself entail that externalism about evidence is false. (As noted above, the step from justificational internalism to evidence internalism requires a bridge principle analogous to [5], which might reasonably be questioned.)

what the subject takes to be evidence. Evidence raises the probability of hypotheses, but nothing can raise the probability of a hypothesis unless it occurs or is the case.<sup>26</sup> A subject's evidence might include her own non-veridical mental states or phenomenal experiences. In this case, the evidence consists in the fact that the subject is in this state, or the evidence consists in the state itself. However, a false proposition itself cannot be evidence, it can only mistakenly appear as evidence or falsely taken to be evidence.

In the second case, the evidence internalist "drives evidence towards the purely phenomenal", as Timothy Williamson puts it.<sup>27</sup> Contrary to common practice, she allows as evidence only phenomenal experiences or propositions about such experiences. I shall continue to assume that ordinary, known propositions about the external world, such as those cited in courtrooms or in defence of scientific theories, can figure as an agent's evidence. Once it is accepted that evidence can include propositions about the external world, and that evidence is factive, we can construct a bad case which proves externalism about evidence (2) to be correct.

A more promising option is to follow Williamson in questioning internalism about the mind (3). Drawing on semantic externalists such as Hilary Putnam, who has famously argued that the *content* of mental attitudes is in part determined by conditions that are independent of an agent's internal state, Williamson argues that the same is true for some mental attitudes themselves, namely *factive* mental attitudes such as knowing or remembering.<sup>28</sup> Accordingly, not only whether an agent's mental attitude has a particular content, but also whether the agent has the mental attitude in the first place might depend on external factors. Williamson's account, which has gained much attention and appreciation in the last years, fits well with the conception of rationality defended here. According to Williamson, evidence is knowledge, and knowledge is a genuine state of mind.<sup>29</sup> It follows that 'A's evidence includes p' is both factive and

---

<sup>26</sup> A statement to the effect that the probability of p, conditional on h, is higher than the unconditional probability of p – i.e.  $P(p|h) > P(p)$  – can of course be true even if h is false. But 'raising' is a success term; a statement to the effect that h raises the probability of p thus entails that the new probability of p is higher than its initial probability, and this follows only if h is assumed to be true.

<sup>27</sup> Williamson (2000, 173).

<sup>28</sup> Cf. Putnam (1973), Williamson (2000, ch. 1-2).

<sup>29</sup> Williamson (2000, ch. 1 and ch. 9).

supervenes on the mind. This provides a straightforward explanation for why rationality can at the same time supervene on the mind and require us to respond to evidence-based reasons. If assumption (3) is denied, then the step to (4) is unsupported and the supervenience argument fails. (4) also cannot be upheld for independent reasons. If having evidence itself amounts to a mental state, then clearly whether A's evidence includes E cannot depend on factors that make no difference to A's mental properties.

However, it might now be argued that on the assumption of an externalist conception of the mind, premise (1) is too weak; rationality must supervene on non-factive, or purely internal, mental states.<sup>30</sup> Given this reformulation of (1), an argument to the same conclusion (7) can be produced which does not require the assumption of internalism about the mind (3). But should we accept a reformulation of (1) in terms of internal states? Although this internal supervenience thesis is more controversial than the general mental supervenience thesis, it is, I think, still a plausible thesis. Consider Goodsmith and Badsmith again, who have the same impression of someone disappearing behind the tree. Since evidence is factive, Goodsmith's but not Badsmith's evidence includes the fact that someone disappeared behind the tree. According to Williamson, this means that they are not in the same state of mind. Therefore the mental supervenience thesis does not guarantee that Goodsmith and Badsmith do not differ with respect to whether they are rational. It is compatible with the verdict that it is rational for Goodsmith to believe that the murderer is hiding behind the tree, while it is not rational for Badsmith to believe this. But intuitively, as things appear to them in exactly the same manner, Goodsmith and Badsmith do not differ with respect to whether they are rational.<sup>31</sup> Hence given an externalist picture of the mind, a more restrictive interpretation of the supervenience thesis is plausible; one that guarantees that Goodsmith and Badsmith do not differ with respect to whether they are rational. Any such principle will suffice for a version of the supervenience argument that does not require internalism about the mind (3) to be true.

---

<sup>30</sup> This is explicit in Wedgwood (2002, 349).

<sup>31</sup> See also Huemer (2006, 149–152) for an interesting and, to my mind, convincing argument for this view.

Now, as I have said, this is not uncontroversial; some authors explicitly reject the thesis that rationality supervenes on the internal.<sup>32</sup> If they are right, then the supervenience argument fails because it is ultimately based on this thesis. So much the better for my account of rationality. However, if rationality supervenes on the internal, then the argument still stands. Since the latter seems to be at least a plausible possibility, we should take a look at the last remaining premise of the supervenience argument: the bridge principle (5). In the next section, I propose a view that denies this principle. If this view is correct, then the response conception of rationality is compatible with both the factivity of evidence and the internal supervenience thesis.

### 7.5 The backup view

To keep things as simple as possible, I will suppose that the original argument up until (4) is sound: whether A is rational or not, but not whether A's evidence includes E, supervenes on the mind, which is understood in terms of purely internal states. Accordingly, whether A's evidence includes E might depend on factors that make no difference to the supervenience basis of rationality. The bridge principle allows us to conclude from this that whether A has decisive evidence-based reason to  $\phi$  might also depend on such factors. Though initially plausible, this step is in fact questionable. In this section, I shall try to make plausible the contrary view, according to which whether A has decisive evidence-based reason depends on the mental, even though what these reasons *are* does not. For reasons that will become apparent in the course of the discussion, I shall call this *the backup view*. The backup view holds that even though our evidence in the bad case is different from our evidence in the good case, our evidence in both cases support the same responses.

---

<sup>32</sup> Among those who deny that rationality supervenes on the internal are Gibbons (2006, §1; 2010a) and Williamson (2000, 178–180). (Gibbons states his argument in terms of justification, but he intends it to apply to rationality as well. In personal correspondence, he wrote to me that he takes 'being in a position to know' to be a factive mental state; his view is thus compatible with the general mental supervenience thesis.) Gibbons puts forward an interesting argument against internalism: sometimes we are rationally required to believe that p because p is supported by something that is obvious to us, even if we fail to take account of what is obvious to us. However, it seems to me more plausible to say that in such cases (i) we are rationally required to take account of what is obvious to us, and (ii) once we have done this, we are rationally required to believe what this supports. So understood, Gibbons' examples do not seem to pose a problem for the internalist.

Recall Goodsmith and Badsmith on their search for the murderer. Since they share all (internal) mental properties, both have the same impression of someone disappearing behind the tree. However, since Badsmith's appearance is non-veridical, Goodsmith's but not Badsmith's evidence includes the proposition that someone has disappeared behind the tree. Let us say that the fact that someone has disappeared behind the tree is in the situation sufficient evidence for believing that the murderer is hiding behind the tree. Since Goodsmith's evidence includes that proposition, Goodsmith has decisive evidence-based reason to believe that the murderer is hiding behind the tree.<sup>33</sup> If whether A has decisive evidence-based reason to  $\phi$  depends only the mind, as the backup view maintains, then Badsmith too must have decisive evidence-based reason for this belief. But since evidence is factive, the reason cannot be the same as Goodsmith's. What theoretical assumptions do we have to make in order to maintain that Badsmith has decisive evidence-based reason for believing that the murderer is hiding behind the tree?

To begin with, we have to assume that appearances or seemings do themselves provide evidence, at least as long as one does not have sufficient evidence that the appearance is non-veridical. As Michael Huemer puts it:

*Phenomenological Conservatism (PC)*: "If it seems to S that p, then, in the absence of defeaters, S thereby has at least some degree of justification for believing that p."<sup>34</sup>

Huemer puts forward some interesting arguments for this claim; I shall here assume it simply on the basis of its intuitive plausibility. Note that (PC) is not in conflict with the factivity of evidence, as long as we take the evidence to consist in the seeming or the fact that something seems to be the case (rather than in what seems to be the case).<sup>35</sup> Moreover, since (PC) only states a sufficient and not a necessary condition, it is

---

<sup>33</sup> As I shall argue below, sufficient evidence for p provides a decisive (evidence-based) reason to believe p only if one attends to p. I ignore this complication here for the sake of simplicity.

<sup>34</sup> Huemer (2007, 30). See also Huemer (2001, 98–115).

<sup>35</sup> Huemer's formula seems compatible with both readings, but contrast Connee (2004, 15–16), as quoted above. Externalists of all kinds can agree with this claim; see e.g. Williamson (2000, 198–199).

entirely compatible with the assumption that ordinary, known propositions about the external world can be evidence.

Other things equal, if  $p$  is evidence for  $q$ , and  $q$  would be evidence for  $r$ , then  $p$  is also evidence for  $r$ .<sup>36</sup> Hence, since a seeming that  $p$  provides evidence for  $p$ , a seeming that  $p$  also provides evidence for  $q$  in the case that  $p$  would be evidence for  $q$  if  $p$  were true. We might thus argue as follows: Badsmith has decisive evidence-based reason to believe that the murderer is hiding behind the tree because it seems to him that someone disappeared behind the tree, and he has no evidence that this seeming is non-veridical. Even though Goodsmith and Badsmith have different bodies of evidence, their evidence-based reasons require them to have the same beliefs.

This does not by itself show the bridge principle to be false; the question is whether the story I have told can be generalised. Perhaps the example has been described in a one-sided way. Can we create a coherent variation in which Goodsmith has evidence that the murderer is hiding behind the tree which Badsmith lacks because it does not even seem to him that someone disappeared behind the tree? Since Goodsmith and Badsmith share all mental properties, it would follow that it also does not seem to Goodsmith that someone is hiding behind the tree. This can be excluded on the following assumption:

*Evidence entails seeming (EES):* If  $A$ 's evidence includes the proposition that  $p$ , then it seems to  $A$  that  $p$ .<sup>37</sup>

A seeming in the sense intended here need not presently occur in your mind, it can be implicit or dispositional. It is simply the sense of seeming in which it is true that it seems to you that  $p$  whenever you know that  $p$ . A seeming that  $p$  can also coexist with a seeming that not- $p$ , as when it seems to you that a stick immersed in water bends even though you know that it does not in fact bend. In such a case, it seems to you that

---

<sup>36</sup> The qualification "other things equal" allows for exceptions in cases in which  $p$  is strong independent evidence against  $r$ .

<sup>37</sup> Note that this is a claim only about propositional evidence. If there is also non-propositional evidence (perhaps consisting in phenomenal experiences themselves rather than the fact that one has such experiences), EES does not require that it must seem to one that one has such an experience.

what seems to you to be the case is an illusion, and so it seems to you that the stick does not bend, even though it also seems to you that it does.

According to EES, A's evidence includes a proposition only if this proposition is in A's 'actual possession', i.e. if p is actually known, or if it at least seems to A that p. Williamson gives a convincing argument against the contrary view, according to which A's evidence can include propositions that are only in one's 'potential possession'.<sup>38</sup> Suppose that I am in the position to know each proposition  $p_1, \dots, p_n$ , but since my capacities are limited, I am in no position to know or even believe all of them. Suppose that q is highly probable given  $p_1, \dots, p_n$  together, but highly improbable given any subset of them. If my evidence includes  $p_1, \dots, p_n$ , then I have sufficient evidence to believe that q. But this is intuitively false, for what I have sufficient evidence to believe should not depend on factors to which I have no access, and I do not have access to  $p_1, \dots, p_n$  as a whole. I shall therefore assume that the relevant notion of evidence entails actual possession, which renders (EES) true.

It follows that if Goodsmith's evidence includes the proposition that someone disappeared behind the tree, then it seems to Badsmith that someone disappeared behind the tree. Could Badsmith fail to have decisive evidence-based reason to believe that the murderer is hiding behind the tree in other ways? He could if he had evidence that his seeming is non-veridical. So let us suppose that Badsmith's evidence includes the proposition that he regularly suffers from blurred vision, and that this evidence suffices to undermine the evidential status of his appearance in such a way that the claim that he has decisive evidence-based reason to believe that the murderer is hiding behind the tree is rendered false. Could it still be the case that *Goodsmith* has decisive evidence-based reason for this belief? Well, if Badsmith's evidence includes the proposition that he regularly suffers from blurred vision, then according to (EES) this must seem to him to be the case, and so it must seem to Goodsmith to be the case that he regularly suffers from blurred vision as well. As seemings can themselves have a positive evidential status, they can also undermine the evidential status of other factors. Hence

---

<sup>38</sup> Williamson (2000, 202–203). Note that this argument is independent of Williamson's view that having evidence amounts to a mental state and thus supervenes on the mind. It is therefore compatible with accepting steps (1)–(4) of the supervenience argument.

it is natural to suppose that unless Goodsmith has evidence that his appearance that he regularly suffers from blurred vision is non-veridical, the fact that he has this appearance undermines the evidential status of his appearance that someone disappeared behind the tree. In this case, we can plausibly deny that Goodsmith has decisive evidence-based reason to believe that the murderer is hiding behind the tree. Treating propositional evidence as knowledge, we can deny that Goodsmith's evidence includes the proposition that someone disappeared behind the tree, for we can plausibly deny that Goodsmith knows that  $p$  if what seems to him to be the case undermines the veridicality of his seeming that  $p$ . On a different conception of evidence, we can still plausibly deny that his evidence-based reason is decisive under the circumstances. If, on the other hand, Goodsmith has additional evidence,  $q$ , which suggests that the appearance that he regularly suffers from blurred vision is non-veridical, then this entails again that it must seem to Badsmith that  $q$ , and so forth.

Here is a more general form of this argument. If Goodsmith's evidence includes  $p$ , it must seem to Badsmith that  $p$ . Unless Badsmith's evidence includes some  $q$  which suggests that his seeming that  $p$  is non-veridical, Badsmith's seeming that  $p$  provides a reason for him that is equally strong as the reason that  $p$  provides for Goodsmith. In this case, what Goodsmith and Badsmith have decisive evidence-based reason to believe coincides. If, on the other hand, Badsmith's evidence includes  $q$ , then it must seem to Goodsmith that  $q$ . Unless Goodsmith's evidence includes some  $r$  which suggests that his seeming that  $q$  is non-veridical, Goodsmith's seeming that  $q$  undermines the evidential status of his seeming that  $p$  to the same degree as  $q$  undermines the evidential status of Badsmith's seeming that  $p$ . In such a case,  $p$ 's evidential status is itself undermined for Goodsmith, and what Goodsmith and Badsmith have decisive evidence-based reason to believe coincides again. If, on the other hand, Goodsmith's evidence includes  $r$ , then it must seem to Badsmith that  $r$ , and so forth.

Besides (EES), this argument requires that something along the following lines is true:

*The evidential impact of seemings (EIS):* If it seems to A that p, and A lacks any evidence that this seeming is non-veridical, then

- (i) if p would be an evidence-based reason for A to believe q if p were true, then A's seeming that p provides an evidence-based reason to believe that q that is equally strong as p would have been, and
- (ii) if p would undermine the evidential status of r if p were true, then A's seeming that p undermines the evidential status of r to the same degree as p would have.

This is stronger than phenomenological conservatism (PC) on its own, but it is a plausible extension. The alternative is to maintain that an undefeated seeming that p has a different evidential impact for Badsmith as p has for Goodsmith, even though the seeming provides *some* evidence for Badsmith (PC), and even though p has its evidential impact for Goodsmith only in virtue of an undefeated seeming that p of exactly the same quality (EES). But this alternative view strikes me as implausible. No person could ever detect the alleged difference in the evidential status of her undefeated seeming that p, on the one hand, and p, as mediated through an undefeated seeming of the same quality, on the other. No person could ever be responsible for not conforming to a decisive evidence-based reason that is provided by p but not by an undefeated seeming that p. A person's reasons should not put her under requirements that she is not even in principle capable of detecting and responding to at the time at which the response is required. (EIS) therefore seems to be on the right track. Together, (EES) and (EIS) entail what can be called *the backup view*: whenever some facts provide evidence for us, our appearances provide sufficient 'backup evidence' that would support the same beliefs if we were in the bad case. Hence, even though we have different evidence in the good and the bad case, our evidence supports the same beliefs in the good and the bad case.

If the backup view is true, then the bridge principle (5) is false. Goodsmith and Badsmith have different bodies of evidence, but if they share all mental properties, their evidence support the same beliefs. As a result, the supervenience argument fails even if premises (1)-(3) are granted. Though I have not presented a conclusive argument in favour of the backup view, I hope at least to have shown that it is a plausible theoretic-

cal option that should be taken seriously. The onus of proof is thus on the defender of the supervenience argument; he cannot simply take the bridge principle for granted.

## 7.6 Evidence, practical reasons, and goodness

The backup view states that even though our evidence in the good case is different from our evidence in the bad case, the same responses are required by our evidence-based reasons in the good and the bad case. If this view is correct, the supervenience argument can be dismissed. In the last section, I argued that the backup view is a plausible theoretical option with respect to epistemic reasons. But if the supervenience argument is to be dismissed by a rejection of the bridge principle, the backup view must hold not only in the epistemic but also in the practical case. According to a practical version of the backup view, even though we have different practical evidence-based reasons in the good and the bad case, these reasons favour the same actions and intentions in both of these cases.<sup>39</sup> This section discusses an objection to the practical version of the backup view. The objection is also of general interest for the approach of rationality proposed here, since it poses a general question concerning the relation between evidence and practical reasons.

Let us grant that the backup view is correct as far as epistemic reasons are concerned. The view is based on the assumption that if  $p$  provides a reason to believe that  $q$ , then an undefeated seeming that  $p$  also provides a reason to believe that  $q$ . This assumption seems plausible because an undefeated seeming that  $p$  is evidence for  $p$ , and (other things equal) if  $p$  provides a reason to believe that  $q$ , then evidence for  $p$  also provides a reason to believe that  $q$ . The objection is that this is not true for practical reasons. If ' $\phi$ ' refers to an action or intention, ' $p$  provides a reason to  $\phi$ ' does not entail that evidence for  $p$  also provides a reason to  $\phi$ ; and so it does not follow that a seeming that  $p$  provides a reason to  $\phi$ . If this is so, then the practical version of the backup view fails: the responses favoured by our practical reasons in the good and the bad case might

---

<sup>39</sup> I assume that the same actions are available in the good and the bad case. If the bad case is construed in such a way that the action is not really available (e.g. because one is only a brain in a vat), then since 'ought' implies 'can', one does not have the same reasons for action. One still has the same reasons for intentions, though. This is compatible with the supervenience thesis as long as we restrict practical requirements of rationality to requirements on intentions.

not be the same. My aim in this section is to reject the fundamental assumption on which this objection is built. The rejection of this assumption establishes that the backup view is a viable option not only in the epistemic but also in the practical case.

So what is wrong with the claim that if  $p$  provides a practical reason to  $\phi$ , then evidence for  $p$  also provides a practical reason to  $\phi$ ? I take it that opposition to this claim stems from the assumption that practical reasons, in contrast to epistemic reasons, are essentially related to the good. A fact is an epistemic reason for a belief that  $p$  in virtue of making  $p$  more likely to be true. A fact is a practical reason to  $\phi$  in virtue of showing the action to be valuable in some respect; only good-making features can be practical reasons.<sup>40</sup> But evidence that a good-making feature occurs is not necessarily itself a good-making feature, and so it does not guarantee that a good-making feature occurs. For example, that bungee jumping would be fun makes bungee jumping good in some respect. But that it seems to you that bungee jumping would be fun does not make bungee jumping good; it only suggests that it is good. If practical reasons must be provided by good-making features, then the story I have told for epistemic reasons does not work for practical reasons. I shall call this the ‘goodness objection’.

The conception of practical reasons on which the goodness objection relies might be expressed by the following principle:

*Loss principle:* There is always a loss involved in not conforming to a practical reason, no matter whether it is outweighed or defeated.<sup>41</sup>

---

<sup>40</sup> A prominent proponent of this view is Raz. Compare: “Reasons [for action] are facts in virtue of which those actions are good in some respect and to some degree.” (Raz 1999b, 23). “Reasons for action are such reasons by being facts that establish that the action has some value” (Raz 2010b, 59). “Reasons for action consist of the fact that the action has some value. [...] The reason must be something that makes the action [...] one that it would be good to perform, and that means that it must be something that renders the action desirable, namely a fact that shows some good in the action” (Raz 2010b, 75). Insofar as this claim is to be understood in such a way that it supports what I call the *goodness objection* above, because it denies that facts which merely *suggest* goodness can be reasons, then (at least as far as I can see) Raz gives no argument for it. In other passages, Raz expresses the nexus between values in reasons in such a way that it is compatible with the view that goodness-indicating facts can be reasons: “Values ‘control’ reasons in that one can have reasons for an action only if its performance is, *or is likely* to produce, or contribute to producing, good, or if it is likely to contribute towards averting something bad” (Raz 1998, 47, my emphasis). This weaker claim does not support the goodness objection, however.

<sup>41</sup> Defenders of a liberal transmission principle need to restrict this claim to reasons that are not derived from it, for there is obviously no loss involved in not taking a means to an action that one has intrinsic reason to

It should be clear that nothing like this is true for epistemic reasons. There is no loss involved in not believing what a clearly outweighed piece of evidence suggests as true. In practical conflicts, things seem different. If you have some reason to go to the movies, and you have a stronger reason to go to the theatre (perhaps because your friend is part of the cast), and you cannot do both, then even if you conform to the reason that is clearly stronger, there is some loss involved. Or so it seems plausible to say.

The goodness objection has severe consequences for the view that rationality requires us to respond to evidence-based reasons. Consider first the epistemic case. Suppose that if *p* were true, *p* would be a good evidence-based reason for *A* to believe *q*, but *p* is false. If *A* has good evidence for *p*, *A* will still have good evidence-based reason to believe *q*. Now contrast this with the practical case. Suppose that if *p* were true, *p* would be a good evidence-based reason for *A* to perform an action or intend it, and *p* is false. If the goodness objection is right, then *A* might have overwhelming evidence for *p*, and yet no reason at all (and hence no evidence-based reason) for this action or intention. It would follow that the set of practical evidence-based reasons is quite often very small.

Consider an example by Derek Parfit, slightly modified for our purposes.<sup>42</sup> Suppose that you stand on the balcony of your hotel and you have very good evidence that the hotel is on fire. There has been a fire alarm, the halls are full of smoke, people are crying for help, etc. In fact, there is no fire; the receptionist has just forgotten to tell you that a movie about a fire in a hotel is being shot in the hall beneath your room. Do you have a reason to jump from the balcony? Parfit says that you do not. You have reason to believe that the hotel is on fire, and if you believe this, you have an “apparent reason” to jump; it would then be rational for you to jump, but no reason actually counts in favour of jumping.<sup>43</sup> Apparently, he assumes that practical reasons must be

---

perform if one takes an alternative and equally good means. For the loss principle, see e.g. Raz (2011, 41): “when we have reason to perform several incompatible actions, even when it is clear which action is supported by the better reason, there is some loss, consisting in the fact that even when we successfully do our best there are concerns that we have reason to meet or satisfy which were left unsatisfied.”

<sup>42</sup> Cf. Parfit (2001, 17).

<sup>43</sup> See Chapter 5.6 for Parfit’s apparent reasons account of rationality.

good-making features. Jumping from the balcony does not in fact serve any value; there would be no loss if you did not jump.

I shall not repeat the arguments against the implication of Parfit's view, according to which we may have no reason at all to do what we are rationally required to do. My aim in this chapter is to outline a different view, according to which rationality is concerned with responding to reasons. The rational thing to do in the situation seems to be to jump, even though there may be nothing that makes jumping good. Thus to defend my view, I seem committed to denying the claim that only good-making features can provide practical reasons, and with it the loss principle.

The view that all practical reasons are provided by good-making features presupposes a teleological conception of practical reasons, which is controversial. Some deontologists would claim that the fact that an action would constitute a lie or the breaking of a promise is a reason against the action independently of whether the action serves any value; the right is not, or at least not entirely, to be determined by the good.<sup>44</sup> Once we allow that the question of whether something provides a reason to act can be answered independently of evaluative considerations, the goodness objection loses its bite. It is no reply to say that just as evidence that indicates the occurrence of a good-making feature is not itself a good-making feature, evidence that indicates the occurrence of a (potential) *right*-making feature is not itself a (potential) right-making feature, for this would just beg the question against the view at issue. To say that evidence for a potential right-maker is not itself a potential right-maker amounts to no more than a negation of the view that evidence for  $p$  provides a reason to  $\phi$  if  $p$  provides a reason to  $\phi$ .

Another worry with the view that only good-making features can be practical reasons is that goodness-indicating facts can naturally be cited as motivating reasons, even if no good-making features are present. Suppose that you decide to do bungee jumping because your experience with sports suggests that bungee jumping would be fun. As turns out, bungee jumping is not actually much fun for you. If the loss principle is correct, then you did not have a reason to go bungee jumping, since there would not

---

<sup>44</sup> See especially Prichard (1949) and Ross (1930). See also Owens (2012, 14–17) for a recent treatment of this question.

have been a loss involved in deciding against it. Yet it seems natural to say that you decided to do it for a reason; there is a perfectly natural and intelligible reason explanation of your action. Intuitively, if someone acted for a reason, then there was a reason to perform the action. It is often claimed that this is not true if the action performed was based on a false belief. According to this view, you can act for the reason that  $p$  even if  $p$  is false and there was not actually a reason to perform the action.<sup>45</sup> But in our case, we need not even assume that you falsely believe that bungee jumping is fun; you need only believe that there is a good chance that it is. If you acted for a reason and you did not have any false beliefs, should we not conclude that you must have had a reason to perform the action?

These are controversial issues. I shall now argue that *even if* we deny the suggested connection between motivating and normative reasons, and even if we grant a teleological conception of practical reasons, the loss principle needs to be rejected. Consider the following example. A doctor is asked for a treatment for a severe disease that significantly affects the life quality of a patient. There is only one possible treatment. The information that is available from other cases in which people have been treated in this way suggests that the chances of cure are very high, but there is a small risk that the treatment will deteriorate the patient's condition considerably. It seems clear that in such a situation, the doctor has some reason to refrain from giving the treatment, which is provided by the risk involved in giving it or by the facts that indicate that this risk exists. Let us say, however, that in our case, the chances of success outweigh the risk of deterioration. Doctor and patient decide in favour of the treatment, and the patient gets cured. There clearly was a reason to act as they did, but there also was a reason to refrain from giving the treatment. Intuitively, however, there was no loss involved in the way they acted. There is not the slightest reason to regret that the doctor did not conform to this reason. Therefore, the loss principle is false.

---

<sup>45</sup> See e.g. Dancy (2000a, 131–137). Dancy's position relies on the controversial claim that something that is not the case can explain an action. Alvarez (2010, ch. 5-6) and Bittner (2001, §199-§213), in contrast, maintain a factive conception of 'acting for a reason'.

The view that only good-making features can be practical reasons ignores the normative significance of risk, or more generally the normative significance of probabilities.<sup>46</sup>

Can this argument help us with the hotel case? We said that you have good evidence that the hotel is on fire. This evidence makes it likely that jumping from the balcony will save your life. The fact that jumping is likely to save your life, I submit, is a good reason to jump. True, if the hotel is not in fact on fire, then there is no loss involved in not conforming to this reason. But as we have seen, this is no reason to deny that this reason exists. And intuitively, the fact that an action is likely to save your life is a good reason to perform this action.

Could one reject the view that the probabilities involved in the hotel case provide reasons once one accepts that the probabilities involved in the doctor case provide reasons? It might be objected that the probabilities in the hotel case are “subjective” while the ones involved in the doctor case are “objective”. Three remarks are in order. First, the relevant notion of likeliness in the hotel case is not subjective in the sense that it is given by the agent’s credences, or by what he believes or believes to be likely. It is an epistemic or evidential notion of probability: jumping is likely to save your life, given your evidence.<sup>47</sup> Second, the relevant notion of risk in the doctor case might well be an epistemic notion in the very same sense. Suppose that there is a fact of the matter, which has not yet been discovered, that the relevant treatment has the deteriorating effects only for people who have a particular genetic disposition. Since the patient in question does not have this disposition, there is no objective risk that the

---

<sup>46</sup> The proponent of the good-making features view might want to describe the situation in the following way: even though the risk of side-effects alone provides no reason to refrain from giving the treatment, it constitutes a normative risk, i.e. a risk that there is a reason to refrain from giving the treatment. However, this reply does not change the fact that the proponent of the good-making feature view has to deny the intuitively plausible claim that there was a reason to refrain from giving the treatment. Moreover, the proposal faces a dilemma. Either the normative risk provides a reason against the action or not. If it does, then the proposal entails the claim that it was supposed to avoid, namely that there was a reason against giving the treatment. If it does not provide such a reason, then it cannot guide us in deliberation about what to do. Hence, the response either collapses into the claim that risks of harm provide reasons or it fails to account for the fact that it makes perfect sense to deliberate about what to do in situations of uncertainty and that we can reasonably be guided by probability considerations.

<sup>47</sup> Hájek (2002, §3) distinguishes three main concepts of probability: (1) “a quasi-logical concept, which is meant to measure objective evidential support relations”, (2) “the concept of an agent’s degree of confidence”, (3) “an objective concept that applies to various systems in the world, independently of what anyone thinks”. Of these three concepts, the kind of probability I have in mind belongs to the first.

treatment deteriorates his condition. Yet given the doctor's evidence, there is a risk, and this risk is normatively relevant. It provides a reason against giving the treatment. Third, even if the relevant notion of probability in the doctor case were objective, the doctor case falsifies the loss principle. So one cannot argue on the basis of the loss principle, or the view that only good-making features are reasons, against the claim that one has an evidence-based reason to jump in the hotel case. But as I said, intuitively the fact that jumping is likely to save one's life is a good reason to jump, and that is so even if the likeliness is explicitly understood in terms of evidential probability.

The defender of the loss principle has two options. Either he flatly denies that probability considerations can provide reasons for actions. This amounts to no less than denying an important part of ethical reality. I shall continue to assume that the risks and chances involved in prescribing pharmaceuticals, distributing organs, investing money, or making political decisions about nuclear energy or climate change can provide reasons for action. Alternatively, the defender of the loss principle must claim that probability considerations are themselves good-making (or bad-making) features in all the cases in which they provide reasons. For example, he must claim that risking the deterioration of the disease does itself have a disvalue, which shows giving the treatment to be bad in some respect. This view insists, somewhat artificially to my mind, that there *is* a loss involved if the doctor and the patient decide in favour of the treatment, even though giving the treatment is favoured by the balance of reasons and turns out to be entirely successful.

As mentioned above, it does not strike me as plausible to say that there is something regrettable about the doctor's conformity with the risk-related reason. More importantly in our context, if the defender of the loss principle allows considerations about the probability of the value of an action's outcome to figure as good- or bad-making features themselves, then the loss principle cannot any longer be used for the goodness objection. Recall that we are considering the question whether evidence for  $p$  provides a reason to  $\phi$  if  $p$  provides a reason to  $\phi$ . The objection was that this is not the case, since only good-making features are reasons, and the existence of evidence indicating the occurrence of a good-making feature does not guarantee that a good-making feature obtains. However, if chances have value and risks have disvalue, then it seems

that evidence indicating the occurrence of a good-making or bad-making feature *does* indeed guarantee that a good-making or bad-making feature obtains. For if one has some evidence that a good-making feature occurs, then there is some chance involved in the relevant action; and if one has some evidence that a bad-making feature occurs, then there is some risk involved in the relevant action.

The goodness objection thus fails either way. If evidence for the occurrence of a good- or bad-making feature does not guarantee the existence of such a feature, then the view that only good- or bad-making features can provide reasons is false – it ignores the fact that risks and chances can have normative relevance. If, on the other hand, evidence for the occurrence of a good- or bad-making feature does guarantee the existence of such a feature (because chances and risks are themselves good- and bad-making features), then the view that only good- or bad-making features can provide reasons for action is compatible with a practical version of the backup view.

Two lessons can be drawn from this discussion. First, the worry that there are often too few practical evidence-based reasons for us to understand practical rationality as a matter of responding to these reasons is unjustified. Evidence for the occurrence of a reason-giving fact itself provides a reason for action, at least very often, namely by constituting a chance or a risk. Second, insofar as the supervenience argument is to be rejected by way of rejecting the bridge principle, there is at least no principled reason to think that a practical version of the backup view is not available. According to such a view, even if two people have different evidence (and thus different evidence-based reasons) in the good and the bad case, the same actions and intentions are supported by their practical evidence-based reasons. For example, if the fact that someone disappeared behind the tree provides a decisive evidence-based reason for Goodsmith to creep up on the tree and look whether the murderer is hiding there, then Badsmith's undefeated seeming that someone disappeared behind the tree also provides such a reason, presumably in virtue of making it likely that doing so would lead to catching the murderer.

Again, I do not claim to have given a conclusive argument for this view; I only have dismissed an important objection to it. I think it is an attractive view, because it is able to capture intuitions about the connections between reasons and responsibility

while at the same time upholding the factivity of reasons. But it is clear that much more needs to be said for a proper defence, and I also do not want to preclude the possibility that there are alternative ways of dealing with the supervenience argument. My aim in this chapter is merely to provide the sketch of a conception of rationality according to which rationality is concerned with responding to reasons, and to offer what seem to me reasonable replies to some natural objections.

### 7.7 Theoretical rationality, uniqueness, and justificational fecundity

I have argued that rationality requires us to respond to our evidence-based reasons, and defended this general idea against an important argument. The remainder of this chapter aims to establish some more specific response requirements of rationality, which will also play a crucial role in the explanation of structural irrationality put forward in Chapters 9 and 10. This section starts with rational requirements on beliefs.

Recall the general response requirement on beliefs that I introduced above:

RRR *Belief*: If A has decisive evidence-based reason (not) to believe that p, then A is rationally required (not) to believe that p.

My aim in this section is to defend a more substantial theoretical requirement that is based on RRR and a substantial account of what our evidence-based reasons for belief are. I shall start with a simplified version, and then work my way through an objection to an important qualification. The simplified version is as follows:

TR\* *Theoretical rationality*: If A has sufficient evidence that p, then A is rationally required to believe p. If A lacks sufficient evidence that p, then A is rationally required not to believe p.

The general rationale for the step from RRR to TR\* is this: RRR tells us that we are rationally required to adjust our beliefs to our decisive evidence-based reasons for belief. What are these reasons? Throughout this thesis, I am assuming that reasons for

attitudes are generally provided only by properties of their contents; there are no state-given reasons for attitudes.<sup>48</sup> Reasons for belief, in particular, are provided by evidence for the truth of their content. If one has no evidence for *p*, or one's evidence for *p* is not sufficient to justify a belief in *p*, then the balance of one's evidence-based reasons count decisively against believing *p*.<sup>49</sup> Therefore, according to RRR, in such a case one is rationally required not to believe *p*. If, on the other hand, one's evidence makes *p* sufficiently likely to justify a belief in *p*, then the balance of one's evidence-based reasons count decisively in favour of believing it. Therefore, according to RRR, in such a case one is rationally required to believe *p*. So on these natural assumptions, TR\* follows from RRR.

There is, however, an objection to this view, which goes as follows. While it is plausible to assume that in every case in which one lacks sufficient evidence for *p*, the balance of one's evidence-based reasons counts against believing *p*, it is less plausible to assume that in every case in which one has sufficient evidence for *p*, one's evidence-based reasons count decisively in favour of believing it. Could we not sometimes be merely permitted to have a belief? The objection is that it would be overdemanding if we were required to have every belief that is sufficiently supported by our evidence.

TR\* does indeed have strong implications. It entails that for every proposition, we are either rationally required to believe it, or we are rationally required not to believe it. In other words, it entails that for every proposition, there is one and only one doxastic attitude that one is rationally permitted to take towards it. Drawing on Richard Feldman, Roger White calls this the "uniqueness thesis":

*The uniqueness thesis:* "Given one's total evidence, there is a unique rational doxastic attitude that one can take to any proposition."<sup>50</sup>

The uniqueness thesis implies that there is no such thing as a permitted doxastic state which is not required; every possible belief is either required or forbidden. For this

<sup>48</sup> See also (among others) Parfit (2011a, App. A), Skorupski (2010, 54–55), Way (2012a).

<sup>49</sup> As mentioned above, the claims I am making could in principle be expressed without supposing that requirements against beliefs are based on reasons against beliefs rather than the absence of reasons *for* beliefs.

<sup>50</sup> White (2005, 445).

reason, Niko Kolodny calls this claim “epistemic strictness”.<sup>51</sup> Kolodny and White both seem to accept the uniqueness thesis.<sup>52</sup> As White argues quite convincingly, epistemically permissive views, which deny the uniqueness thesis, introduce a kind of arbitrariness into rational belief-formation that is very much at odds with our first-personal experience of epistemic deliberation.

While I basically share these worries about arbitrariness, I also believe that the uniqueness thesis is far too demanding as it stands. This becomes clear when we look at a feature that Mark Nelson calls “the infinite justificational fecundity of evidence”:

*The fecundity thesis:* “[...] every single bit of evidence, whether experiential or propositional, potentially epistemically justifies an infinitely large array of different beliefs.”<sup>53</sup>

Nelson gives the example of a perceptual impression of dark, winged shapes moving across his visual field. This visual piece of evidence, he argues, potentially justifies an infinite number of beliefs, such as the belief that something is moving through the air in front of him, that birds are flying in front of him, that jackdaws are flying in front of him, or that at least three jackdaws exist. Yet it seems clear that we are not required to have all of these beliefs. First, as Gilbert Harman points out, there is no point in cluttering one’s mind with an endless array of uninteresting beliefs.<sup>54</sup> Second, since our cognitive capacities are limited, we simply cannot come to have all the beliefs that our evidence potentially justifies.

This provides a strong argument against the uniqueness thesis. According to the fecundity thesis, every piece of evidence potentially justifies infinitely many beliefs. So if, as the uniqueness thesis entails, epistemic permissions always amount to epistemic requirements, it would follow that every piece of evidence issues an infinite number of requirements on beliefs. Beings like us could not possibly satisfy all these requirements together. Therefore, the uniqueness thesis must be false.

---

<sup>51</sup> See Kolodny (2007a, 248).

<sup>52</sup> Indeed, Kolodny’s explanation of first-order structural requirements of rationality relies on the uniqueness thesis. See Kolodny (2007a, 248–250).

<sup>53</sup> Nelson (2010, 96).

<sup>54</sup> Harman (1986, 12).

This argument might be questioned as follows. In order to count as believing *p*, it is not necessary that we currently represent *p* as true in our mind, it is sufficient that we *would* represent *p* as true *if* we were presented with it. For example, even though I did not represent the proposition that 2 plus 2 equals 4 in my mind until just a moment ago, I believed this proposition all day long, since I would have endorsed its truth immediately once I had been presented with it. These latter kinds of beliefs are sometimes called *dispositional* or *implicit* beliefs and contrasted with *occurrent* beliefs. Most of our beliefs are dispositional. In particular, occurrent beliefs are usually accompanied by a huge number of dispositional beliefs. An occurrent belief in the proposition ‘jackdaws are flying in front of me’ is normally accompanied by a huge number of dispositional beliefs in propositions such as ‘birds are flying in front of me’, ‘something is moving through the air’, ‘at least three jackdaws exist’, and ‘at least two jackdaws exist’. These propositions might be believed even though they do not occur in our mind. For these reasons, it might be argued that the fecundity of evidence represents no problem for the uniqueness thesis. We are indeed required to have all the beliefs that our evidence supports as true – it is only that these beliefs need not be occurrent.

This reply fails, however. It is true that many of the propositions that our evidence supports as true are so obviously related to one another that one can count as believing all of them when one believes one of them. But this is not always the case, and as a consequence, the fecundity problem for the uniqueness thesis remains. In many cases, coming to have an occurrent belief about an evidentially supported proposition requires more than just being presented with the proposition; it requires thinking. In such cases, it is false to say that we already have the beliefs in question even just in a dispositional or implicit sense. And such cases are still ubiquitous enough to suggest that a requirement to have all these beliefs would be far too demanding.<sup>55</sup>

---

<sup>55</sup> I am here assuming an agglomeration principle, according to which a requirement to believe *p* and a requirement to believe *q* entail a requirement to [believe *p* and believe *q*]. As far as I can see, this principle is uncontroversial, but if it is disputed, the same conclusion can be reached via a principle of joint satisfiability, according to which *A*’s being required to believe *p* and *A*’s being required to believe *q* entail that it is possible for *A* to [believe *p* and believe *q*].

Here is an example. All day long I possessed clearly sufficient evidence that 16 times 7 equals 112.<sup>56</sup> But I did not believe it until just this minute, when I attended to this question and made up my mind about it. I also did not believe it dispositionally, for it is not true that I would have immediately endorsed the truth of this proposition after being presented with it. I would have had to calculate first, and thus did not have a belief about it in either the occurrent or the dispositional sense. So all day long, I did not believe a proposition I had sufficient evidence for. And there are many other propositions for which this is true as well. Now if the uniqueness thesis were true, I would have been required to have all these beliefs. But certainly such a requirement would seem both pointless and overly demanding. There are uncountably many propositions that are sufficiently supported by our evidence even though forming the respective beliefs would still require thinking, and we cannot reasonably be expected to make up our minds about all these propositions and form the respective beliefs.

I conclude that the uniqueness thesis, as it stands, cannot be true. Consequently, TR\* needs to be revised in such a way that it no longer entails the uniqueness thesis. As the title of his article makes clear, Nelson draws a much stronger conclusion, namely that “we have no positive epistemic duties” at all. It seems to follow from his position that TR\* needs to be rejected outright, since it essentially involves a claim about positive epistemic requirements.

Let us have a closer look at Nelson’s argument. According to the fecundity thesis, the evidence licenses an infinite variety of beliefs. If the evidence does not sufficiently support a proposition, then we are epistemically required to refrain from believing in this proposition. Such negative epistemic requirements exist. However, according to Nelson evidence never by itself grounds a positive requirement to adopt a belief. What proposition, out of the propositions the evidence permits us to believe, we *ought* to believe or are *required* to believe, depends on non-epistemic factors, such as our interests and desires:

---

<sup>56</sup> The point does not depend on there being evidence for mathematical propositions. Any other case in which one has sufficient evidence for an irrelevant proposition but coming to believe this proposition requires some thinking would do.

My thesis is that there is nothing we positively ought to believe simply in virtue of our epistemic circumstances, and nothing that we ‘ought’ to believe at all, except given some further interest, desire, duty, or such like.<sup>57</sup>

Part of my disagreement with Nelson is terminological. Many authors, among them Gilbert Harman and Robert Nozick, have argued that doxastic requirements (requirements to have or lack a belief) are conditional on non-epistemic factors.<sup>58</sup> Nelson seems to think that whenever a doxastic requirement is conditional on some non-epistemic factor, it cannot be an epistemic requirement. So if your being required to believe *p* is explained not only by the evidence you have for *p* but also by the fact that you are interested in whether *p* is the case, for Nelson this means that the requirement in question is not epistemic.<sup>59</sup> I believe this to be unnecessarily restrictive. Doxastic requirements can, I think, legitimately be understood as epistemic or theoretical requirements even if they are in place only when certain non-epistemic background conditions are satisfied. Things would look different if the requirements in question were assumed to involve prudential, moral, or any other *non-epistemic reasons for belief*, but they do not in the examples that Nelson discusses. In any case, Nelson does not deny that – given that certain background conditions are satisfied (which do not involve any non-epistemic reasons for belief) – we are often required to have particular beliefs that our evidence supports. This is all that is needed in order to vindicate a modified version of TR\* as a requirement of rationality or reasons. Nelson would insist that this would not be a purely epistemic or theoretical requirement, but that is a terminological choice that is both optional and irrelevant for the substantial question at issue.

As it happens, I also have a substantial disagreement with Nelson: positive doxastic requirements are not generally conditional on “some further interest, desire,

---

<sup>57</sup> Nelson (2010, 92).

<sup>58</sup> “One is to add a new proposition *P* to one’s beliefs only if one is interested in whether *P* is true (and it is otherwise reasonable for one to believe *P*)” (Harman 1986, 55). “Believe (an admissible) *h* [hypothesis] only if the expected utility of believing *h* is not less than the expected utility of having no belief about *h*” (Nozick 1993, 86).

<sup>59</sup> See esp. Nelson (2010, 92–93).

duty, or such like".<sup>60</sup> For agents like us to stand under a requirement to believe *p*, it is sufficient that we have sufficient evidence for *p* *and* pay attention to the question of whether *p* is the case. Plausibly, not just any kind of momentary attention will do; I take it that what kind of attention is relevant depends on the context. What is important here is only that some kind of attention will do. While I shall leave open whether other conditions can also be sufficient, my claim is that no other condition is necessary. No matter whether you or anyone else is interested in whether *p* is the case, if you in fact deliberate about whether *p* is the case, or you pay attention to *p* in some other relevant way, and your evidence sufficiently supports *p*, then you cannot rationally abstain from believing it.<sup>61</sup> I thus suggest the following revision of TR\*:

TR    *Theoretical rationality*: If *A* has sufficient evidence for *p*, and *A* attends to *p*, then *A* is rationally required to believe *p*. If *A* lacks sufficient evidence for *p*, then *A* is rationally required not to believe *p*.

TR follows from RRR on the assumption that one has decisive evidence-based reason to believe *p* if one has sufficient evidence for *p* and attends (in the relevant way) to *p*. These revised claims say, in effect, that *if* we make up our mind about something, we should do it in accordance with our evidence. But since they do not require us to make up our mind in the first place, the justificational fecundity of evidence is no objection to these claims. Since we do not in fact attend to most of the propositions that are supported by our evidence, TR does not imply that we are required to believe an infinite number of propositions.

---

<sup>60</sup> Here I also disagree with Harman and Nozick.

<sup>61</sup> I contend that this is true even if one irrationally attends to the question whether *p* is the case. The practical irrationality of attending to random or irrelevant questions does not excuse one from the epistemic irrationality that is involved in withholding a belief about an evidentially supported proposition one pays sufficient attention to. In other words, the mistake of a person who forms lots of irrelevant beliefs is not to be resolved by abstaining from believing evidentially supported answers to questions she attends to, but by abstaining from attending to irrelevant questions. A more difficult question is whether belief-abstention can be permitted if one compulsively, rather than irrationally, attends to a lot of random and irrelevant questions. But in this case, it seems to me plausible to say that insofar as a person lacks control about what questions she attends to, she also ceases to be subject to rational requirements on beliefs in general. Thanks to Thomas Kroedel for pressing me on this.

TR therefore no longer entails the uniqueness thesis. For all that TR says, there are many propositions that we may both rationally believe and rationally refrain from believing. At the same time, however, my proposal still preserves what seemed plausible about the uniqueness thesis. Recall White's point that the kind of rational arbitrariness between believing and not believing a proposition, which the denial of the uniqueness thesis seems to entail, is at odds with the experience of epistemic deliberation. The view that I am suggesting upholds this feature of epistemic deliberation. If we deliberate about whether  $p$  is the case, we attend to  $p$  – and once we attend to  $p$ , there *is* only one doxastic attitude that we can rationally hold with respect to  $p$ . If we may rationally refrain from believing something we could rationally believe, this is only because we may rationally not attend to it at all. This position strikes me as the best way to account for the fecundity of evidence while preserving what is plausible about the uniqueness thesis.

### 7.8 Evidence-based reasons and evidence for 'ought'-propositions

Before moving on to response requirements of rationality on intentions, my concern in this section is with the connection between evidence-based reasons and evidence for 'ought'-propositions. This connection will play a crucial role in the formulation of the practical requirement that I defend in the next section and the explanation of structural irrationality that I aim to provide in Chapter 9. The purpose of this section is to defend the following principle:

*Evidence principle:* If  $A$  has sufficient evidence that she ought to  $\phi$ , and  $A$  can  $\phi$ , then  $A$  has decisive evidence-based reason to  $\phi$ .

The 'can' clause is necessary since one cannot have a reason to do something that one cannot do. I take this to be an independent necessary condition on any reason.<sup>62</sup> Yet supposing that one can give a particular response, it strikes me as an interesting and

---

<sup>62</sup> For a forceful defence of the claim that 'reason' implies 'can', see Streumer (2007). Heuer (2010) puts forward some objections to this view, but I think they are effectively refuted by Streumer (2010).

important truth that sufficient evidence for an 'ought'-proposition goes along with decisive evidence-based reasons to give the response in question.<sup>63</sup>

Why should this be so? First, it is what we should intuitively expect. What A has sufficient evidence to believe and what A has decisive evidence-based reason to do feed on the very same body of facts, propositions, or pieces of evidence. If all the evidence taken together supports the belief that one ought to  $\phi$ , then it is quite natural to think that all the evidence taken together also supports  $\phi$ -ing itself (always provided that the independent condition on reasons, namely that one can  $\phi$ , is satisfied).

Second, there is also a more general rationale for the evidence principle. If you have sufficient evidence that you ought to  $\phi$ , then this is partly because you have sufficient evidence for the occurrence of an ought-making feature. This means that given your evidence, this feature is likely to obtain. As we have seen in the discussion of the goodness objection, the fact that an ought-making feature is likely to obtain is usually a good evidence-based reason to perform the action in question. And given that the evidence for the 'ought'-proposition is sufficient, it is difficult to see how this reason could be outweighed by other evidence-based reasons. Consider an example: you have sufficient evidence that you ought to go to the lake this afternoon. This must be because your evidence suggests that something makes it the case that you ought to go to the lake; e.g. you have sufficient evidence that swimming will provide you with an overdue recovery from all the work that you have done in the last days. This in turn means that given your evidence it is likely that swimming will provide this recovery. The fact that swimming is likely to provide the needed recovery is a good reason to go to the lake. Given that nothing in your evidence undermines the proposition that you ought to go to the lake (as we are assuming that you have sufficient evidence for it), it is difficult to see what evidence-based reason could outweigh this reason. Hence it is plausible to say that the facts that provide sufficient evidence for the proposition that you ought to go to the lake also provide decisive evidence-based reason to go to the lake.

---

<sup>63</sup> Gibbons (2009, 173) defends the stronger principle "If you're justified in believing that you ought to  $\phi$ , then you ought to  $\phi$ ". I believe this claim needs to be complemented with the 'can' condition, but once we do this, the evidence principle above and the evidence-relative view defended in the next chapter together entail Gibbon's JO-principle.

A third reason to accept the evidence principle is that negating it has awkward implications. Suppose that you ask yourself whether you ought to  $\phi$  at  $t$ , you consider all your evidence, and this evidence is sufficient to justify a belief that you ought to  $\phi$  at  $t$ . If you respond correctly to your evidence-based reasons in such a case, then you will believe that you ought to  $\phi$  at  $t$ . But now suppose that despite all of this, you do not have decisive evidence-based reason to  $\phi$  at  $t$ , as someone who denies the evidence principle must think is possible. It follows that you might respond correctly to your evidence-based reasons while not  $\phi$ -ing at  $t$ . So if one denies the evidence principle, one is committed to the following possibility: You respond correctly to all your available reasons, and at  $t$  you believe that you ought to  $\phi$  at  $t$ , but you do not  $\phi$ , even though you can. I submit that I find this implication quite disturbing. It is, to say the least, counterintuitive that someone who is responsive to all his available reasons could end up in this situation. We can go further. If despite your evidence that you ought to  $\phi$ , you may not have a decisive evidence-based reason to  $\phi$ , then it is difficult to see why you may not even have a decisive evidence-based reason *against*  $\phi$ -ing. It would follow that responding correctly to all your evidence-based reasons requires you to believe that you ought to  $\phi$ , and at the same time requires you not to  $\phi$ . This is an even more disturbing implication. It is quite implausible to assume that one could be required to be in such a state by responding correctly to all one's available reasons.

So the evidence principle is intuitively plausible, it is supported by natural assumptions about reasons that are provided by probability considerations, and denying it leads to very strange results. Resistance to it might stem from resistance to a *different* claim, according to which reasons can be *analysed* in terms of evidence for ought-propositions:

*Evidence analysis:* For R to be a reason for A to  $\phi$  is for R to be evidence that A ought to  $\phi$ .<sup>64</sup>

---

<sup>64</sup> See Kearns and Star (2008; 2009) and Thomson (2008, ch. 8-9) for an analysis of reasons along these lines.

If this claim were true, then the evidence principle would plausibly also be true, but the evidence principle does not require the truth of this analysis. For the record, I do not subscribe to the evidence analysis. There are a number of pressing objections to it, but as I shall argue now, none of them provides an objection to the evidence principle.

First, there are plausible counterexamples to the claim (entailed by the evidence analysis) that every reason to  $\phi$  is also evidence that one ought to  $\phi$ . For example, John Brunero has argued that the fact that Dad suggested a certain present for Mom could be a reason to buy this present (because this would make Dad happy), while at the same time *lowering* rather than *raising* the probability that one ought to buy this present (because Dad's suggestions have so far proved to be completely unsuitable for Mom).<sup>65</sup> The evidence principle does not entail that all reasons are evidence for ought-propositions, so this is not an objection against the evidence principle.

Second, there are also plausible counterexamples to the claim (entailed by the evidence analysis) that every fact that is evidence for an ought-proposition is also a reason. For one, R can be evidence that A ought to  $\phi$  even if A cannot  $\phi$ , but as mentioned above, one cannot have a reason to perform an action that one cannot perform.<sup>66</sup> While the evidence principle explicitly allows that there is an independent ability-constraint on reasons, this does not seem to be open to someone who wants to analyse the notion of a 'reason' in terms of evidence. For another, some facts that are evidence that one ought to  $\phi$  are better regarded as background or 'enabling' conditions for reasons or 'oughts' rather than reasons themselves. For example, given that you have promised to  $\phi$ , the fact that you are able to  $\phi$  is evidence that you ought to  $\phi$ , but it is not plausibly itself a reason to  $\phi$ . Similarly, given that you have promised to  $\phi$ , the fact that you have no reason not to  $\phi$  is evidence that you ought to  $\phi$ , but not plausibly itself a reason to  $\phi$ .<sup>67</sup> Again, this is an objection to an analysis of reasons in

---

<sup>65</sup> See Brunero (2009b, §1). Kearns and Star (2013, §4) discuss some replies to this objection, the most promising of which involves the suggestion that the probability is raised relative to some salient subset of one's evidence. However, Kearns and Star have not yet taken up the challenge to explain how the relevant subsets can be specified.

<sup>66</sup> See above for further literature. Kearns and Star (2009, 235–236) simply bite the bullet on this point and deny that 'reason' implies 'can'.

<sup>67</sup> See Dancy (2004a, 38–41) for this example and the distinction between reasons and enablers, and Brunero (2009b, §2) for applying the example to Kearns's and Star's analysis. Kearns and Star (2013, §5) deny that the

terms of evidence for ‘ought’-propositions, but it is not an objection to the evidence principle. Assuming that the facts that you have promised to  $\phi$  and that you have no reason not to  $\phi$  provide sufficient evidence that you ought to  $\phi$ , the evidence principle forces us to conclude that you also have decisive evidence-based reason to  $\phi$ . This is the natural verdict. But there is no implication that the fact that you have no reason not to  $\phi$  itself provides an evidence-based reason to intend to  $\phi$ . Much more plausibly, this fact is an evidence-based background condition for the fact that you have promised to  $\phi$  to be a *decisive* evidence-based reason to intend to  $\phi$ . While the evidence analysis seems to collapse the distinction between reasons and background conditions, the evidence principle is entirely compatible with this distinction.

There is one kind of objection to the evidence analysis that also applies to the evidence principle. Some authors have argued that there are facts that can provide sufficient evidence that one ought to  $\phi$ , even though they are not reasons to  $\phi$ , and even though there are no other reasons to  $\phi$ . This is a possibility that the defender of the evidence principle has to deny. Let us have a look at the examples. In the first kind of case, the newspaper says that people are starving in Africa, and this is evidence that one ought to donate money to Oxfam. Suppose that one has no other evidence-based reason to donate money to Oxfam. The objection is that while the fact that people are starving in Africa is a reason to donate money, the fact that *the newspaper says* that people are starving in Africa is no reason to donate money. After all, the newspaper could have it wrong, and if there are no people starving, then there might be no reason to donate the money.<sup>68</sup> If this is true, then persons could have sufficient evidence that they ought to  $\phi$  even if they do not have decisive evidence-based reason to  $\phi$ .

The objection seems to presuppose that only good-making features can be reasons, and that evidence that a good-making feature occurs is not itself a good-maker.

---

example provides an objection. They hold that *given that you have promised to  $\phi$* , the fact that you have no reason not to  $\phi$  is indeed a reason to  $\phi$ . Similarly, they would presumably have to claim that the fact that you have no greater reason not to  $\phi$  is a reason to  $\phi$ . Here it might be objected that taking this fact to be a reason involves some kind of double counting: verdictive statements about the overall balance of reasons cannot themselves add to one’s reasons (cf. Dancy 2004a, 40). However, one might reply that such facts are ‘derivative reasons’ that do not have independent normative force (cf. Parfit 2011a, 39, for a suggestion along these lines, though in a different context). Be that as it may, it remains intuitively implausible to say that the fact that one does not have a reason to  $\phi$  itself counts in favour of  $\phi$ -ing.

<sup>68</sup> Kearns and Star (2009, 233) attribute this objection to James Morauta.

I have argued above that the conjunction of these claims forms an untenable view that ignores the normative significance of risks and chances. Why does the newspaper report provide sufficient evidence that one ought to donate money? At least partly because it makes it likely that donating money will help people in desperate straits. The fact that an action is likely to help people in desperate straits is a good reason to perform this action. The assumption that the newspaper report does not provide an evidence-based reason to donate money is simply unjustified.

A trickier counterexample has been presented by John Broome.<sup>69</sup> Suppose that a reliable book says that you ought to eat plenty of cabbage, but it does not say why. This seems to be evidence (perhaps even sufficient evidence) that you ought to eat cabbage, but it does not seem to provide a reason itself. Given that you do not have a further evidence-based reason to eat cabbage, the example might seem to falsify the evidence principle. So why is the fact that a reliable book says that you ought to eat cabbage not a reason to eat cabbage? Because reasons are, according to Broome, potential right-makers, they must be able to figure in an explanation of why you ought to do something.<sup>70</sup> The fact that a reliable book says that you ought to eat cabbage, however, seems only to *indicate* that you ought to eat cabbage, not to be part of what *makes it the case* that you ought to eat cabbage.

In order for this to be a counterexample to the evidence principle, we need to assume that the book's assertion provides sufficient evidence that you ought to eat cabbage. This can only be so if the book's assertion also provides sufficient evidence that there is a fact that makes it the case that you ought to eat to cabbage. As I have argued, evidence for the occurrence of a (potential) right-maker usually provides practical reasons. For example, it is natural to assume that the further facts that are supposed to make it the case that you ought to eat cabbage are facts that make it in some way or other good for you to eat cabbage. On this assumption, the book's assertion is evidence that you ought to eat cabbage at least partly because it is evidence that eating cabbage is good for you. If you have sufficient evidence that eating cabbage is

---

<sup>69</sup> See Broome (2008, 101–102), Kearns and Star (2008, 47–49).

<sup>70</sup> See esp. Broome (2004a). Broome goes so far as to suggest that reasons can be *defined* in terms of their explanatory role for 'ought'-claims; I am here only relying on the weaker idea that reasons play such a role.

good for you, then given your evidence, eating cabbage is likely to be good for you. The fact that eating cabbage is likely to be good for you is, however, a plausible reason to eat cabbage, and it might indeed make it the case that you ought to eat cabbage. So again, while Broome's example can be taken to provide an objection to the evidence analysis, because we might be reluctant to say that the book's assertion is itself a reason to eat cabbage, it is no objection the evidence principle. On the assumption that the book's assertion is sufficient evidence that you ought to eat cabbage, it remains plausible to think that there are evidence-based reasons to eat cabbage, which are, on balance with other evidence-based reasons, decisive.

This concludes my discussion of the evidence principle. I have first argued that the evidence principle is intuitively plausible, that it is supported by general considerations about reasons and that denying it has awkward implications. I have then shown that the evidence principle avoids a number of natural complaints that can be raised against the related, but importantly different, thesis that reasons can be analysed in terms of evidence for 'ought'-propositions. Finally, I have dismissed an objection that applies to both of these views by showing that it implicitly relies on the view that only good-making features can be practical reasons, which I have rejected above.

## 7.9 Practical rationality, and reasons for actions and intentions

My aim at this point of the chapter is still to defend some more specific response requirements of rationality that will later be used to explain the appearance of structural irrationality. With the evidence principle in hand, I can now finally turn to requirements on intentions. Above, I have argued for the following general practical requirement:

RRR *Intention*: If A has decisive evidence-based reason (not) to intend to  $\phi$ , then A is rationally required (not) to intend to  $\phi$ .

But what are our reasons for intention? On a natural view, our reasons for intentions are just the same as our reasons for action:

AIL\* *Action/intention link*: If A has decisive evidence-based reason (not) to  $\phi$ , then A has decisive evidence-based reason (not) to intend to  $\phi$ .

We could conclude from RRR and AIL\* that we are rationally required to intend all actions we have decisive reason to perform. But the view that reasons for intentions and actions are the same is vulnerable to objections. One kind of counterexample involves the assumption of state-given reasons for intention: you have no reason to drink the toxin, but since you have been promised a large reward for *intending* to drink it, you have strong reason to intend to drink it.<sup>71</sup> As I have indicated above, the right response here seems to me to deny that such reasons exist. That you have been promised a reward for intending to drink a toxin is not a reason to intend to drink it, but a reason to want that you intend it or a reason for bringing this intention about. The general rationale for this response is the idea that reasons must be capable of guiding agents, and thus of figuring in certain kinds of explanations. Something can be a reason for a response only if it is possible for agents to give this response *for* this reason, and this condition is not satisfied by state-given reasons.

Another kind of case involves reasons for non-intentional processes or actions in a very broad sense of 'action'. It might be argued that even though you have reason to wake up early tomorrow, it might not be the case that you have reason to *intend* to wake up early tomorrow. Since waking up is not something one can do intentionally, intending to do it might not make much sense to begin with; it might even have very bad effects, as when your intending to wake up early keeps you from falling asleep and thereby causes you to oversleep the next day. Again, it seem to me that the right response is to deny that there are, strictly speaking, any reasons for processes such as waking up. What might seem to be such a reason is actually a reason for an intentional action such as setting the alarm clock. The rationale for this response is just the same as above: waking up is not something that can be done for a reason and hence cannot be favoured by a reason.

---

<sup>71</sup> See Kavka (1983).

Even putting aside such questionable cases, there still seem to be plausible examples in which reasons for action and intentions come apart. Suppose that some action is likely to have good effects, which provides a reason to intend its performance. But suppose that unbeknownst to you, you cannot really perform this action. Since 'reason' implies 'can', you then have no reason to perform the action. But that you cannot perform the action does not mean that you cannot intend it. If your evidence does not indicate that you cannot perform the action, you may still have an evidence-based reason for the intention. This is a counterexample to the identification of reasons for action with reasons for intention, but it does not call into question AIL\*. Even if not all reasons for intentions are reasons for action, it might still be true that all reasons for action are reasons for intention, and that therefore decisive evidence-based reasons for action ground rational requirements on intentions.

However, there are also counterexamples that work in the other direction. In one kind of case, the favoured action is simply so far in the future that there is no need to intend it already – one might be permitted to form the intention at some later point even if one already has decisive reason for the action. Perhaps you have decisive reason to attend a conference that takes place in a year, but since there is no need to make the decision now, your reasons permit you not to form any intention now.<sup>72</sup> This involves a misunderstanding of AIL\*, however. If you have reason to attend a conference *in a year*, AIL\* entails only that you have reason to intend *in a year* to attend the conference; it does not entail that you have reason to intend *now* to attend the conference.<sup>73</sup>

A more serious kind of case is the one in which you are certain to perform the favoured action no matter whether you intend it or not – perhaps because you intend another action the performance of which ensures the performance of the favoured action. A doctor has decisive reason to save a patient's life by giving him a particular drug that is necessary for his survival. She also has decisive reason to relieve the patient of a mild pain. She intends to save his life, and she knows that the drug that the patient needs to survive will also relieve him of his mild pain. In such a case, it might seem that

---

<sup>72</sup> This point is suggested by some remarks of Raz (2005b, 8).

<sup>73</sup> More exactly, if you *now* have a reason to attend the conference *in a year*, AIL\* entails only that you *now* have a reason to intend *in a year* to attend the conference; it does not entail that you *now* have a reason to intend *now* to attend the conference. See the next chapter for more on reasons and time.

her reasons permit the doctor to refrain from intending to relieve the patient of his pain, even though she has decisive reason to relieve him of his pain.

There is a similar problem for claims about *structural* practical rationality. Recall:

(A)\* *Weakness of will*: Necessarily, if A believes that she ought to  $\phi$ , and A does not intend to  $\phi$ , then A is irrational.

(B)\* *Instrumental irrationality*: Necessarily, if A intends to  $\phi$ , and A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A does not intend to  $\psi$ , then A is irrational.

As was pointed out in Chapter 1 already, if you are certain that you will do what you believe you ought to do no matter whether you intend to do it, then it might not be irrational of you not to intend it, as (A)\* entails. And if you are certain that you will take the means to your ends no matter whether you intend to take these means, it might not be irrational of you not to intend them, as (B)\* entails.<sup>74</sup>

In the former case, some authors have responded to this problem by understanding weakness of will as a failure to respond to beliefs about reasons for *intention* rather than action.<sup>75</sup> The problem with this interpretation of (A)\* is that it restricts weakness of will to cases that are untypical and marginal at best. As T.M. Scanlon puts it, our normative judgements are usually “content-directed” rather than “attitude-directed”.<sup>76</sup> The most common instances of weakness of will therefore consist in a clash between an intention-state and a belief about the *content* of this state, i.e. a belief about reasons for *action* rather than *intention*.

Other authors have tried to specify the relevant conditions. For example, John Broome introduces into both (A)\* and (B)\* the further condition that A has a belief with the following content: ‘If I were not now to intend what I believe I ought to do (or what I believe to be the necessary means to my end), then, because of that, I would

<sup>74</sup> See esp. Kamm (2000, §4). See also Broome (2002, 91–92; 2013, 162–63). In the context of Chapter 1, I wanted to bracket these problems, and in order to avoid unnecessary complexity, I did not mark claims (A)\* and (B)\* with an asterisk. I apologise for the inconsistency.

<sup>75</sup> See Kolodny (2005, 521–527), Scanlon (1998, 25).

<sup>76</sup> Scanlon (2007, 90–91).

not do it.<sup>77</sup> A similar move could be made with respect to the action/intention link: we might plausibly claim that if A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that [if she were not now to intend to  $\phi$ , then, because of that, she would not  $\phi$ ], then A has decisive evidence-based reason to intend to  $\phi$ . And we could then conclude that in such a case, she is also rationally required to intend to  $\phi$ .

I think it is plausible that Broome's condition is indeed sufficient, but it seems also quite restrictive, and I want to leave open the possibility that a weaker condition can do the job. For my purposes, it suffices to say that there are certain explicable background conditions that connect reasons for actions to reasons for intentions. I shall work with the placeholder "condition C". The structural irrationality claims might thus be revised as follows:

- (A) *Weakness of will*: Necessarily, if A believes that she ought to  $\phi$ , and A believes that condition C obtains, and A does not intend to  $\phi$ , then A is irrational.
- (B) *Instrumental irrationality*: Necessarily, if A intends to  $\phi$ , and A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A believes that condition C obtains, and A does not intend to  $\psi$ , then A is irrational.

The reader may substitute for "condition C" the condition that A will not perform the relevant action unless she now intends to perform it. But the reader may also replace it with another, perhaps weaker, condition that she deems sufficient. Analogously, I propose the following revision of AIL\*:

AIL *Action/intention link*: If A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that condition C obtains, then A has decisive evidence-based reason to intend to  $\phi$ . If A has decisive evidence-based reason not to  $\phi$ , then A has decisive evidence-based reason not to intend to  $\phi$ .

---

<sup>77</sup> Cf. e.g. Broome (2010, 289–290).

Note that the qualification involving condition C occurs only in the first of these claims. Those objections to ALL\* that have not been rejected on independent grounds do not call into question that one always has decisive evidence-based reason against intending an action one has decisive evidence-based reason *not* to perform, and I also do not see any other reasons for questioning this.

Next, recall the evidence principle that I defended in the last section:

*Evidence principle (EP)*: If A has sufficient evidence that she ought to  $\phi$ , and A can  $\phi$ , then A has decisive evidence-based reason to  $\phi$ .

The evidence principle and the action/intention link together entail:

*Interim conclusion (IC-1)*: If A has sufficient evidence that she ought to  $\phi$ , and A can  $\phi$ , and A has sufficient evidence that condition C obtains, then A has decisive evidence-based reason to intend to  $\phi$ . If A has sufficient evidence that she ought not to  $\phi$ , and A can omit to  $\phi$ , then A has decisive evidence-based reason not to intend to  $\phi$  (*from ALL and EP*).<sup>78</sup>

We can now remove the ‘can’ condition from this claim. The ‘can’ condition was introduced into the evidence principle as an independent constraint on any reason. Sufficient evidence that one ought to  $\phi$  provides sufficient evidence-based reason to  $\phi$ , provided that one can  $\phi$ . However, since we are now concerned with reasons for *intention*, the restriction that one can perform the *action* is no longer necessary. As we have seen, one can have evidence-based reason to intend an action even though one cannot in fact perform the action.

It might be argued that we need to introduce the different condition that one can *intend* the action. But if one cannot form a particular intention on the basis of potential reasons, then this is not because of external obstacles, but because of psychological incapacity. The psychological capacity to form intentions is a precondition

---

<sup>78</sup> The second sentence follows because ‘ $\phi$ -ing’ in the evidence principle covers not only actions but also omissions.

of one's being subject to reasons or rational requirements on intention in general; so all the requirements on intentions mentioned are implicitly conditional on one's capacity to form intentions. As we need not explicitly mention that A has a rational capacity as a condition of a claim about reasons or rational requirements, we need not explicitly mention that A is capable of forming an intention. So I shall continue to formulate principles of reasons and rationality in such a way that they implicitly take for granted that one can form the attitudes in question.

If we remove the 'can' condition from *IC-1*, we get:

*Interim conclusion (IC-2)*: If A has sufficient evidence that she ought to  $\phi$ , and A has sufficient evidence that condition C obtains, then A has decisive evidence-based reason to intend to  $\phi$ . If A has sufficient evidence that she ought not to  $\phi$ , then A has decisive evidence-based reason not to intend to  $\phi$ .

Next, recall the response requirement of rationality proposed above:

RRR *Intention*: If A has decisive evidence-based reason (not) to intend to  $\phi$ , then A is rationally required (not) to intend to  $\phi$ .

RRR *Intention* and *IC-2* finally entail the following rational requirement:

PR *Practical rationality*: If A has sufficient evidence that she ought to  $\phi$ , and A has sufficient evidence that condition C obtains, then A is rationally required to intend to  $\phi$ . If A has sufficient evidence that she ought not to  $\phi$ , then A is rationally required not to intend to  $\phi$ .

By calling this claim "practical rationality", I do not want to suggest that it exhausts all practical response requirements of rationality. It is only one practical implication of RRR. But it is an important implication, and as I shall argue below, it helps us to understand the phenomenon of structural irrationality and explain it in a way that does

not presuppose the existence of structural requirements of rationality. This will have to wait until Chapter 9, however.

### 7.10 Summary and prospects

In this chapter I have defended the view that rationality requires us to respond to our available reasons, understood as reasons that are part of an agent's body of evidence, and discussed some important implications of it. The response requirements that follow from this account are in an important sense normative. The reasons that are available to a person form at least a significant subgroup of this person's reasons, and response requirements are guaranteed to be supported by what I have called "decisive evidence-based reasons", i.e. by the preponderance of available reasons.

In the next chapter, I will go further. I shall argue that what we ought, or have overall decisive reason, to do *just is* what we have decisive *available* reason to do. It follows that the response requirements mentioned are normative in the strongest possible sense; they are always supported by decisive reasons. Chapters 9 and 10 will finally explore how the appearance of structural irrationality can be explained in terms of these response requirements. Given what I have argued in this chapter, this would amount to an explanation of structural irrationality in terms of rational requirements that are normative in the sense of 'being supported by an important subgroup of reasons'. Given what I will argue in Chapter 8, it would amount to an explanation of structural irrationality in terms of rational requirements that are in the strongest sense normative. The reader who is content with the weaker conclusion may therefore skip the next chapter.

## An evidence-relative account of reasons

---

This chapter defends the *evidence-relative view*, according to which what we ought, or have decisive reason, to do depends only on our evidence-based reasons. I start by introducing the debate between objectivist and perspectivist views about ‘ought’ and clarifying my thesis as a version of perspectivism (8.1). I then put forward three objections to the objective view: that it has counterintuitive normative implications, that it cannot account for the normative significance of risks and chances, and that it does not capture the intuitive connections between ‘ought’ and legitimate criticism (8.2). Finally, I argue that the objective view not only fails to provide agents with guidance in circumstances of uncertainty, but seriously *misguides* them to making extremely irresponsible decisions (8.3). The remainder of the chapter is concerned with the main challenge for the evidence-relative view, namely to account for phenomena (most notably advice) in which the use of ‘ought’ is geared to evidence that is better than the evidence currently available to the agent (8.4; 8.6; 8.9). I argue that this challenge can be met by paying close attention to the role of *time* in the truth conditions of reason judgements (8.5; 8.7-8.9).

### 8.1 Objectivism and perspectivism about ‘ought’

Imagine a doctor who is faced with a patient’s disease that she knows will lead to death unless treated shortly.<sup>1</sup> Two possible treatments are available: A and B. After careful consideration of the available evidence, the doctor concludes that treatment A will cure the patient, and B will kill him. Unbeknownst to her, however, in fact treatment B is the cure, while A will lead to the patient’s death. What ought the doctor to do: give A or give B?

---

<sup>1</sup> This is a modification of an example by Jackson (1991, 462–463) that I will return to below. I assume that there is no reason for the doctor not to try to help the patient.

Let us call facts about a person's beliefs, knowledge or evidence facts about that person's *perspective*. We may then ask more generally: Does what an agent ought to do depend on the agent's perspective, or is it perspective-independent? Objectivists about 'ought', such as G. E. Moore and J. J. Thomson, claim that 'ought' is independent of the agent's perspective.<sup>2</sup> Hence, they would hold that the doctor ought to give treatment B – the one that in fact cures the patient. Perspectivists like H. A. Prichard and W. D. Ross, on the other hand, believe that 'ought' depends on the perspective of the agent – a view that is sometimes spelled out in terms of the agent's actual beliefs, and sometimes in terms of the evidence available to the agent.<sup>3</sup> Both of these versions of perspectivism hold that the doctor ought to give A, not B.

Others again try to solve the puzzle by distinguishing different senses of 'ought'.<sup>4</sup> According to them, all that we can say is that the doctor ought to give A, relative to her perspective, and that she ought to give B, relative to all the facts. I do not deny that it can be useful to speak of what an agent ought to do relative to certain considerations, and that different qualified notions of 'ought' might be important in their own right. Nevertheless, I believe that there is a substantial question at issue

<sup>2</sup> See Moore (1912, 80–82); Thomson (1983; 1990, 231–232). Note that Thomson slightly modifies her view in her (2008, 187–99), but remains opposed to the idea that 'ought' depends on the agent's perspective. See Bykvist (2011) and Graham (2010) for two recent defences of objectivism, though both limit their discussion to moral obligation and leave open how it relates to the overall 'ought' of deliberative conclusions at issue here.

<sup>3</sup> In phrasing the issue in terms of perspective-dependence, I follow Gibbons (2010b, 335). What I call "perspectivism" is often referred to as "subjectivism," but I think it is appropriate to reserve the latter term for a belief-relative view only. Belief-relative perspectivists include Howard-Snyder (2005); Jackson (1991, esp. 464–465); Prichard (1932); and W. D. Ross (1939, 146–167). Note that Ross, who advanced an objective view in his earlier *The Right and the Good* (1930) changed his mind completely after becoming convinced by Prichard's arguments. Chapter VII of his *Foundations of Ethics* (1939) is basically an affirmative summary – Dancy (2002, 244) calls it a "slavish account" – of Prichard's lecture. Different versions of evidence-relative perspectivism can be found in Andric (2013a); Dancy (2000a, ch.3); Gibbons (2010b); Mason (2013); Robertson (2011); Scanlon (2008, 47–52); and Zimmerman (2008). As above, some of these authors focus on moral duties rather than the deliberative 'ought'.

<sup>4</sup> Dorsey (2012); Ewing (1947, 112–144); F. Feldman (2012); R. Feldman (1988b); Gibbard (1990, 42–43; 2005, 340–341); Jackson (1991, 471–472); Jackson and Smith (2006, 268–271); Oddie and Menzies (1992, 512); Parfit (2011a, 150–164); J. Ross (2006, 167–176); Schroeder (2008; 2009); Smith (2006, 147; 2008, 252–61); and Wedgwood (2003) all introduce different senses of 'ought' though some of them accept the dominance of one or more of these senses. Ewing "generally prefers" the evidence-relative sense (1947, 128); Gibbard thinks that the objective 'ought' is "not genuinely normative" (2005, 344) and can be defined in terms of the belief- or evidence-relative ought (which he does not seem to distinguish); Parfit claims that for decision-making only the belief- and evidence-relative notions are relevant (2011a, 158–161). Alternatively, but in the same spirit, 'ought'-statements are sometimes taken to be elliptical, i.e., claimed to be meaningless unless they at least implicitly refer to a particular body of information.

between objectivists and perspectivists when it comes to the notion of ‘ought’ involved in the deliberative question “What ought I to do?” and deliberative conclusions of the form “I ought to  $\phi$ ”. Conclusions of this sort are directly response-guiding; a rational agent must be capable of guiding her responses on the basis of judgements about what she ought to do in this sense. But in order to make a rational decision guided by a belief that one ought to do something, one needs a univocal concept of ‘ought’ that figures in such beliefs. It is perfectly consistent to believe “I ought to  $\phi$ , relative to X” and “I ought not to  $\phi$ , relative to Y”, but one cannot rationally intend both to  $\phi$  and not to  $\phi$ . There must be *one* sense of ‘ought’, the belief in which is the relevant one for deliberative conclusions. We need to be able to judge “I ought to  $\phi$ , *full stop*”. At any rate, this is what I shall assume in the following discussion.

It is natural to consider the deliberative ‘ought’ as the central, unqualified sense of ‘ought’ and regard other possible senses of ‘ought’ as qualified senses.<sup>5</sup> This claim, however, will not function as a substantial assumption of the argument. This chapter is about the question of whether the deliberative ‘ought’ depends on perspective, and I will refer to this concept by using the word ‘ought’ without qualification only for the sake of convenience.

My aim is to defend a version of perspectivism, but I will be concerned with the evidence-relative view only. The belief-relative (or subjective) view is problematic for several reasons. For one, it seems to entail that one can bootstrap an ‘ought’ into existence by mere wishful thinking, which is implausible. For another, it is incompatible with the factivity of reasons, provided the natural assumption that the deliberative ‘ought’ is grounded in reasons. It also strikes me as phenomenologically inappropriate. Retrospectively, agents can deem their own normative judgements as false when they were based on false beliefs, even if they were based correctly on these beliefs. This would make little sense if the truth of such judgements would just depend on what they actually believed. A detailed discussion of the subjective view, however, is beyond the

---

<sup>5</sup> See Broome (2013, 22–25) and J. Ross (2012a, 163–164), who both claim that the central sense of ‘ought’ is the one figuring in the beliefs that are subject to the claim that it is irrational not to intend what one believes one ought to do. I also sympathise with Davidson’s view (1969, 38–39) that judgements concluding deliberation must be unqualified (or “unconditional” as he puts it) in order to be practical (although he takes the conclusion of a deliberation to be a betterness-judgement rather than an ‘ought’-judgement).

scope of this chapter. Here I will simply assume that a plausible perspectivist position will be spelled out in terms of evidence.

### *The evidence-relative view*

What exactly is the claim of the evidence-relative view? Many authors suggest that the objective view can be understood as the claim that one ought to  $\phi$  whenever  $\phi$ -ing is one's best option, while the perspectivist view can be understood as the claim that one ought to  $\phi$  whenever  $\phi$ -ing is the expected, or prospective, best option, relative to one's beliefs or evidence.<sup>6</sup> One problem with this setup is that it cannot plausibly be applied to 'oughts' that relate to beliefs or other attitudes.<sup>7</sup> A further problem is that it presupposes the substantial thesis that what we ought to do is always a function of the value of our actions. This is controversial: many philosophers hold that the right is not in this way determined by the good. Whatever account is correct here, it should be clear that the question of how 'ought' is sensitive to information is a different one and should not be conflated with the first by defining positions on perspective-dependence in terms of value-determination.

An alternative route is suggested by the assumption that the deliberative 'ought' is a function of a person's *reasons*. Given that an agent ought to  $\phi$  if and only if she has decisive reason to  $\phi$ , the evidence-relative view might be stated as follows:

*The evidence-relative view:* A has decisive reason to  $\phi$  iff A has decisive evidence-based reason to  $\phi$ .

---

<sup>6</sup> Authors who define objectivism and perspectivism in terms of value include Jackson (1991) and Moore (1912, 80–82), who discuss the question right from the start within a consequentialist framework, but also R. Feldman (1988b); Kolodny and MacFarlane (2010; unpublished); and Zimmerman (2008), who purport to give an account that is independent of consequentialism. Broome is more cautious in restricting his discussion to situations, where – as he puts it – “consequentialism [...] applies” (2013, 36). The thesis that 'ought' is a function of value is a version of the thesis that Broome calls “teleology”. For a discussion of teleology as a substantial thesis, see Broome (1991, ch.1) and (2004b, ch.3.1). While many authors seem to take for granted that it is unproblematic to account for the information-relativity of 'ought' in terms of value, Zimmerman (2008, 2–5) gives at least a rudimentary illustration of his opinion that this approach is compatible with different sorts of substantive moral theories.

<sup>7</sup> Here I agree with Raz (2011, 41–45) that epistemic reasons are not to be explained in terms of value, such as a putative value of having true beliefs.

As explained in the last chapter, what an agent has decisive evidence-based reason to do is determined by the preponderance of her evidence-based reasons, where a reason *R* is evidence-based if *R* is part of the agent's evidence.<sup>8</sup> The evidence-relative view thus claims that what an agent has *overall* decisive reason to do is determined by the preponderance only of her *evidence-based* reasons. It is natural to ask what must be true on the level of *pro tanto* reasons in order for this view to be true. A straightforward option is to say that all reasons are evidence-based; only facts that are part of an agent's evidence can be reasons:

*Evidence constraint\**: *R* is a reason for *A* to  $\phi$  only if *A*'s evidence includes *R*.

There are alternative theoretical options, but the version of the evidence-relative view I defend here is formulated in terms of the evidence constraint.<sup>9</sup> As the further discussion will make clear, however, I only accept a *synchronic* version of this constraint: if *R* is a reason for me to do something *now*, then *R* must be part of my evidence *now*; for *R* to be a reason to do something at a later time, it needs only to be the case that *R* would be part of my evidence at that later time if certain conditions (discussed below) were satisfied. Yet, according to this view, facts that do not meet either of these constraints

<sup>8</sup> Strictly speaking, this condition on *R* being an evidence-based reason only holds synchronically, i.e. at *t*, *R* is an evidence-based reason for *A* to  $\phi$  at *t* only if *A*'s evidence at *t* includes *R*. For diachronic cases, in which *R* is a reason to do something at a later time, see the discussion in later sections.

<sup>9</sup> One alternative option is to say that all kinds of facts can be *pro tanto* reasons, but only facts that are part of an agent's evidence can be "adequate" reasons, i.e. reasons that can be *sufficient* to justify a certain response in a particular situation. For example, Raz (2011, 113–120) distinguishes between different epistemic and availability constraints for something to be a reason and for something to be an adequate reason (cf. also Kolodny 2011b, 341–342). This view is less radical in allowing for the existence of *pro tanto* reasons that are not evidence-based, but it raises the question of how there could be reasons that do not have any bearing on the truth of deliberative conclusions. It is not clear to me that it makes sense to accept facts as reasons, but exclude them from making any contribution to conclusions about what one has sufficient or decisive reason to do. A further possibility is to distinguish between the reasons that an agent *has*, and the reasons that *there are* for this agent (see e.g. Lord 2010). According to such a conception of the evidence-relative view, what an agent ought to do is determined only by the reasons that she *has*, which are her evidence-based reasons. The relevant distinction is stipulative: neither ordinary nor philosophical discourse on reasons consistently distinguishes between reasons one has and reasons there are. Again, I am not fully convinced that it is useful. Nonetheless, the main arguments and claims of this chapter may also be expressed in either of these alternative views on *pro tanto* reasons.

are not reasons. Following Allan Gibbard, we might still call some of these facts “potential reasons”<sup>10</sup>:

*Potential reason:* R is a potential reason for A to  $\phi$  iff (i) R is the case, and (ii) R would be a reason for A to  $\phi$  if A’s evidence included R.

The objective view can be understood as denying the evidence constraint, and maintaining that all potential reasons are also actual reasons.<sup>11</sup> Intermediate views are possible, of course. Such views deny the evidence constraint but accept some weaker constraint on reasons, such as the condition that R must in principle be knowable by a human being, or by the agent herself.<sup>12</sup> I shall address such intermediate views in the course of discussing the objective view.

Sections 8.2-8.3 are concerned with the rejection of the objective view and its semi-objectivist relatives. Sections 8.4-8.8 are concerned with the defence of the evidence-relative view. I focus on a set of problems which many, including myself, take to be the most serious challenge to the evidence-relative view. First, if what an agent ought to do depends on her limited information, how can a better-informed adviser give good advice to this agent in terms of what the agent ought to do? Second, how can

---

<sup>10</sup> Cf. Gibbard (1990, 162).

<sup>11</sup> This is on the assumption that all actual reasons bear on the question of what the agent ought to do (in the sense relevant for deliberative conclusions). If this assumption is denied, e.g. by introducing the constraint that only evidence-based reasons bear on this question (because only they are “adequate reasons” or “possessed reasons”), then the objective view has to be understood as denying this constraint as well.

<sup>12</sup> For the former view, see Thomson (2008, 198): “we ought to do a thing only if a human being can know that we ought to”. For the latter view, see Raz (2011, 110): “if some people cannot know of a fact it does not constitute a reason for them, even though other people can know about it”. Raz’s view is extremely difficult to pin down. On the one hand, he embraces an epistemic filter that is individuated to particular agents and excludes facts that are unknowable for them as reasons. On the other hand, he suggests as a criterion for unknowability that “no one at the time would have found out about it regardless of their best efforts” (2011, 110, n. 6). What is crucial for the forthcoming discussion is that Raz, even though accepting certain epistemic constraints, ultimately wants to deny that “temporary epistemic limitations affect the force of reasons” (2011, 126). Hence, that a fact must be knowable by an agent in order to provide a reason for her is, according to Raz’s account, compatible with its being the case that a fact is a reason for an agent in a particular situation, even though in the particular situation the agent would not find out about it regardless of her best efforts. This is what distinguishes Raz’s view from the one I put forward in Section 8.6 (which could be understood in terms of ‘knowability’ as well), and it is also what makes Raz’s account vulnerable to the misguidance objection discussed in Section 8.3. For a helpful interpretation of Raz’s position on epistemic constraints, see Kolodny (2011, 341–343).

we understand an agent's seeking new evidence, or an adviser's sharing evidence not yet available to the agent, as ways of finding (or helping to find) an answer to the agent's deliberative question, "What should I do?" if the correct answer to that question depends on the agent's perspective? My aim is to develop an evidence-relative account of reasons that answers these questions by paying close attention to the role of *time* in the truth conditions of reason-judgements.

In response to the above-mentioned problems for the evidence-relative view, some philosophers have recently argued for relativist theories of 'ought', according to which what an agent ought to do does not depend on the evidence of the agent herself, but on the evidence of the *speaker*, or even the *assessor*, of the ought-statement.<sup>13</sup> Such accounts raise their own problems, for example by allowing for a multitude of practically incompatible truths about what an agent ought to do at a given time, but a discussion of these views lies beyond the scope of this chapter. However, a defence of the evidence-relative view – the view that the deliberative 'ought' invariantly relates to the evidence of the *agent* – will help to undermine the need for such relativist accounts, which are generally motivated by the supposed failure of non-relativist accounts to resolve the problems in question.

## 8.2 Three objections to the objective view

Let me start with an example of Frank Jackson's that illustrates some of the problems of the objective view:

Jill is a physician who has to decide on the correct treatment for her patient, John, who has a minor but not trivial skin complaint. She has three drugs to

---

<sup>13</sup> See Björnsson and Finlay (2010); Kolodny and MacFarlane (2010; unpublished). The former view is sometimes referred to as "contextualism", and the latter as "relativism", but both views have contextualist as well as relativist aspects. See also Henning (forthcoming) for a contextualist treatment of the issues discussed here. I should note that for the general argument in favour of the normativity of rationality, the difference between my agent-relative account, on the one hand, and speaker- or assessor-relative accounts, on the other hand, may not matter very much. All of these accounts agree that the 'ought' of deliberative conclusions, when assessed from the standpoint of the deliberating agent, is relative to the agent's evidence. If rationality can be understood in terms of responding to evidence-based reasons, as I argue in Chapter 7 and Chapters 9 and 10, it follows that it is true, from the standpoint of the deliberating agent, that she necessarily ought to do what rationality requires of her.

choose from: drug A, drug B, and drug C. Careful consideration of the literature has led her to the following opinions. Drug A is very likely to relieve the condition but will not completely cure it. One of drugs B and C will completely cure the skin condition; the other though will kill the patient, and there is no way that she can tell which of the two is the perfect cure and which is the killer drug. What should Jill do?<sup>14</sup>

As Jackson points out, the intuitively correct answer to that question is that Jill ought to prescribe drug A. This is also what the evidence-relative view entails. The relevant evidence-based reasons are the following facts: drug A is very likely to relieve the condition; there is a 50 per cent chance that drug B will completely cure the condition and a 50 per cent risk that drug B will kill the patient, and there is a 50 per cent chance that drug C will completely cure the condition and a 50 per cent risk that drug C will kill the patient. The balance of these reasons clearly weighs in favour of giving A and against giving B or giving C. So what Jill ought, intuitively, to do is provided by the balance of her evidence-based reasons.

*First objection: counterintuitive normative implications*

The objective view, in contrast, seems incapable of giving the intuitively correct answer in this case, for there is a fact of the matter which of the drugs provides the complete cure. Let us say that this is drug C (“C” as in “cure”). Since the objectivist denies the evidence constraint, he cannot exclude this fact from Jill’s reasons. Once we accept that the fact that drug C completely cures John is a reason for Jill to give drug C, it is difficult to see how this reason could be outweighed by any other consideration that is relevant in this case. What better reason could there be? Surely the fact that C provides a complete cure is not outweighed by the fact that A relieves the patient’s condition. Could it be outweighed by the fact that there is a 50 per cent risk that drug C will kill the patient? That does not make any sense. The risk in question is epistemic, it is relative to a set of information that does *not* include the fact that drug C will cure the

---

<sup>14</sup> Jackson (1991, 462–463). Regan (1980, 265) presents a structurally equivalent case, as do Parfit’s mine shafts example (2011a, 159) and J. Ross’s three envelope problem (2006, 174).

patient. Once we accept that the fact that C is the cure as a reason to give C, we cannot at the same time take the fact that there is a risk that C is not the cure as a reason not to give C. There is no coherent standpoint from which both of these facts could provide reasons. Could the fact that C is the cure be outweighed by the fact that Jill does not know whether C is the cure?<sup>15</sup> That also does not make sense. Perhaps the fact that Jill does not know whether C is the cure is a reason for her not to give C. But this could only be so if the fact that C is the cure is not at the same time a reason for her to give C. These facts could not both be reasons in the same situation; they could not be weighed against each other. This is not because weighing them is practically impossible from Jill's point of view, but because there is *no* single point of view from which these two facts can sensibly be weighed against each other.

Hence, once we allow the fact that C is the cure as a reason, we seem stuck with the conclusion that Jill has decisive reason and thus ought to give C.<sup>16</sup> Moreover, it seems that the fact cannot be blocked as a reason by introducing a constraint that is weaker than the evidence constraint (such as the constraint that a reason-giving fact must in principle be knowable). To see this, let us stipulate in the example that a colleague of Jill knows whether B or C is the cure, but that Jill cannot reach the colleague and must act immediately.<sup>17</sup> It remains intuitively plausible that the correct conclusion for her to draw is that she ought to give A. That it is in principle possible for

<sup>15</sup> Dancy (2000a, 56; 2002, 236) makes a suggestion along these lines in response to an example of Prichard's (1932, 93). Dancy's aim in this context is not to defend the objective view – as it is understood here – but to show that one can make sense of Prichard's point that ignorance is normatively significant while maintaining that reasons are facts, i.e. without following Prichard in becoming a subjectivist. His suggestion that facts about one's ignorance can be reasons is compatible, and, I think, makes most sense in combination with the evidence-relative view defended here. Indeed, Dancy suggests himself that there is an "agent-relative epistemic filter" for facts to be reasons, and he provides independent support for this claim (2000a, 56–59).

<sup>16</sup> That is, unless we make one of the moves mentioned above and say that only evidence-based reasons count for the truth of deliberative conclusions, which would be incompatible with the objective view.

<sup>17</sup> Compare the first variant of Jackson's example discussed by Raz (2011, 121). At first appearance, Raz seems to argue that his own preferred epistemic constraint, according to which a reason must be knowable by the agent, entails that the fact that C is the cure is not a reason even in this variant of the case (2011, 123). If his constraint is to be understood in this way, then I think there is no substantial difference between Raz's view and the one I defend in more detail below. However, as the progression of Raz's discussion suggests, he ultimately wants to deny that "temporary epistemic limitations affect the force of reasons" (2011, 126), and he withdraws his (apparent) conclusion that "Jill's best reason is to give John drug A" (2011, 123) in favour of the view that "the best reason supports one of the other drugs", but giving A "is the best approximation to what she has best reason to do" (2011, 124). I conclude that Raz understands his own view in such a way that (at least in the variant case at issue here) Jill has decisive reason, or ought, to give C, but not A. This is what makes it vulnerable to all the objections discussed in this section and the next.

her to know whether B or C provides the cure does not change the fact that it is in her situation far too risky to give any other treatment than A. We thus have arrived at a first argument for the evidence-relative view: In contrast to its competitors, the evidence-relative view provides the intuitively correct answer in Jill's case.

*Second objection: epistemic probabilities can be reasons*

A second argument is already suggested by these remarks. It is clear that risks, or more generally evidential probabilities, can be reasons. We constantly have to make decisions in circumstances of uncertainty, and such decisions can be guided by considerations about risks and chances – considerations that count for or against these decisions and are thus reasons.<sup>18</sup> Jill's case is only one case in point. It is unclear how the objective view can make sense of this. As we have seen, once one accepts the fact that one's  $\phi$ -ing will have a certain effect as a reason, one cannot at the same time accept the fact that one's  $\phi$ -ing will *likely* (or *unlikely*) have this effect as a reason. There is, as I have put it, no coherent standpoint from which both of these facts can count as reasons. But that means that the objectivist must deny that probability-related considerations are reasons whenever there is an objective fact of the matter about whether one's  $\phi$ -ing has a certain effect. By the same token, intermediate views must deny that probability-related considerations are reasons whenever there is a fact of the matter about whether one's  $\phi$ -ing has a certain effect that is in principle knowable. In my view, this amounts to nothing less than a denial of an important part of ethical reality. So while the evidence-relative view naturally accounts for the normative significance of epistemic uncertainty, its competitors seem committed to denying, at least in a wide array of cases, that risks and chances can provide reasons.

*Third objection: the connection between 'ought' and criticisability*

A third argument can be based on the intuitive connections between what we ought to do and what we are criticisable for. The objectivist maintains that Jill ought to give C.

---

<sup>18</sup> Raz (2011, 115) accepts that evidential probabilities can provide practical reasons, but he does not address the problem that this fact poses for the objective view.

If one fails to do what one ought to do, one seems open to legitimate criticism. But it would not be legitimate to criticise Jill for not giving C. It would, however, be legitimate to criticise Jill for not giving A. Hence, the evidence-relative view captures, while its competitors fail to capture, the intuitive connection between ‘ought’-statements and legitimate criticism.

In response, objectivists claim that there are cases of blameless wrongdoing; they maintain that one need not be criticisable for an action one ought not to have performed if one was inculpably ignorant about its wrongness.<sup>19</sup> The problem with this reply is that Jill knows that either B or C provides the complete cure and A does not. So if objectivism were true, then Jill could *know* that giving A is wrong and would still not be blameworthy for doing it. The objectivist thus not only needs to allow for blameless wrongdoing, he needs to allow that one can *knowingly* do wrong without being in *any way* criticisable. This strikes me as quite implausible. Moreover, even if there is blameless wrongdoing, that does not conversely show that agents can be blameworthy without doing any wrong.<sup>20</sup> Indeed, as I have argued in Chapter 2, criticisability entails the violation of a decisive reason or ‘ought’. But as was pointed out, Jill would be criticisable for not giving A, and yet according to objectivism she would not have violated a decisive reason in doing so.

#### *Intermediate conclusion*

Let me sum up these objections against objectivism and its semi-objectivist relatives. First, the objective view has counterintuitive normative implications in cases such as Jill’s, in which agents have to decide in circumstances of known uncertainty. Second, the objective view seems incapable of accounting for the fact that risks and chances

---

<sup>19</sup> See e.g. Graham (2010, 93–94); Moore (1912, 81–82); Thomson (2008, 191).

<sup>20</sup> Graham argues that “a person may be morally blameworthy even when she does what she knows she’s morally obliged to do. For example, X might be obliged to chop off Y’s leg [...] but if X does so not to help Y but, rather, to cause Y excruciating pain, X will be blameworthy even though she does what she knows she is obliged to do” (2010, 94). This does not entail that a person may be blameworthy without doing any wrong. It might be wrong to chop off someone’s leg in order to cause him pain even if it is right to chop off his leg. And intuitively, X is not blameworthy for chopping off the leg, but for his doing so with the purpose of causing pain.

provide reasons. Third, it cannot make sense of the intuitive connections between ‘ought’ and legitimate criticism.

The discussion focused on practical reasons, but similar points can be made with respect to reasons for belief as well. Suppose that I throw a fair coin and do not let you see whether it landed heads or tails.<sup>21</sup> What should you believe? Suppose the coin landed heads. If we do not accept the evidence constraint, then this fact is among your reasons. If it is, we seem forced to conclude that you have decisive reason, and thus ought to believe, that the coin landed heads. What better reason could there be? So the objective view is committed to saying that you ought to believe that the coin landed heads. As before, this conclusion is not prevented by a constraint that is weaker than the evidence constraint, for the fact that the coin landed heads is in principle knowable. Yet intuitively, you should not believe that the coin landed heads; you should suspend judgement on that question. This is what the evidence-relative view entails: since your evidence-based reasons together do not provide sufficient support for believing that the coin landed heads, you should not believe it. So the evidence-relative view captures our intuitive judgement in this case much better than the objective view. It is also unclear how the objective view about reasons for belief can make room for the significance of epistemic probabilities. Last but not least, you seem to be criticisable if you believe that the coin landed heads, while you do not seem criticisable if you suspend judgement. So the evidence-relative view gets the connection between criticisability and ‘ought’ right, while the objective view gets it wrong.

### 8.3 The misguidance argument

In this section, I shall provide a further argument against the objective view, which is concerned with the guiding role of reasons. I take it that the point of deliberation is, at least in part, to evaluate reasons, and rationally guide our responses in accordance with our normative judgements. So one, if not the most important, role that reasons and ‘oughts’ play in deliberation is that they provide guidance. In most cases, our situation is characterised by at least partial uncertainty. How can we make responsible decisions

---

<sup>21</sup> I borrow this example from Gibbard (2005, 340).

and rationally guide our behaviour in such situations if our reasons depend on facts that are beyond our epistemic reach?

Traditionally, objectivists have assumed that we could be guided by our justified (but possibly false) beliefs about what we ought to do, objectively understood.<sup>22</sup> But Jackson's example shows that this is not true, for Jill *knows* that giving A is definitely not what she ought to do, objectively understood. So for agents in situations of uncertainty, objectivists do not seem to have much more to offer than Karl Kraus' notorious piece of advice: "In case of doubt, decide in favour of what is correct".

In fact, things are even worse. As I shall argue now, the objective conception of the deliberative 'ought' not only fails to account for normative guidance in circumstances of uncertainty, it seriously *misguides* agents, forcing them, on pain of irrationality, to make extremely irresponsible decisions. The point can again be illustrated by Jill's case. We can start with the obvious: it is irresponsible to impose a 50 per cent risk of death on someone in order to gain a 50 per cent chance of curing a minor complaint; hence, the only responsible decision for Jill to make is to give drug A. To strengthen this point, let us assume that not giving any treatment would be likely to lead to a serious deterioration of John's disease that might ultimately lead to his death as well. Thus:

(1) The only responsible decision for Jill is to give drug A.

Proponents of the objective and the evidence-relative view agree on this premise; it cannot reasonably be denied that if Jill acts responsibly, she will give drug A.<sup>23</sup> Next, let us suppose that Jill is an objectivist and believes the implications of this view. Jill must then believe that she either ought to give B or ought to give C. Let us suppose that she can only give one drug (for example because the financial resources only suffice for one treatment). Then Jill must believe that giving A is incompatible with doing what she ought to do. If an action prevents one from doing what one ought to do, then one ought not to perform this action. This is a general application of the principle that

---

<sup>22</sup> Cf. Moore (1912, 82).

<sup>23</sup> See e.g. Graham (2010, 97).

reasons for action transmit to reasons for taking the necessary means and to reasons against incompatible actions. So if Jill believes what the objective view implies, she believes that she ought not to give drug A:

- (2) If Jill believes in the objective view, then she believes that she ought not to give drug A.

Next, recall the structural irrationality claim (A)\*, according to which a person cannot be rational if she does not intend what she believes she ought to do. A less demanding and independently plausible variant of this claim is that a person cannot be rational if she positively intends or decides in favour of an action she believes she ought *not* to perform. Hence:

- (3) It is irrational to decide in favour of an action that one believes one ought not to perform.

Now, it seems clear that believing the correct theory about 'ought' must be compatible with rationally making responsible decisions. A theory about the deliberative 'ought' which entails that an agent cannot rationally make the only responsible decision must be false:

- (4) Believing the correct theory about 'ought' is compatible with rationally making responsible decisions.

Assumptions (1)-(4) entail that the objective view is false:

- (5) If Jill believes in the objective view, then she cannot rationally decide to give drug A (*from 2 and 3*).
- (6) If Jill believes in the objective view, then she cannot rationally make the only responsible decision (*from 1 and 5*).

(7) Therefore, the objective view is not the correct theory about ‘ought’ (from 4 and 6).

I shall refer to this as the *misguidance argument*. According to it, the objective view not only fails to account for positive guidance in circumstances of uncertainty, it seriously *misguides* agents into making irresponsible decisions. If Jill believes in the objective view, she cannot rationally make the only responsible decision; she must, on pain of irrationality, risk the death of her patient by either refraining from any intentional action or by giving B or C. We would not want our doctors to be objectivists.

As before, the argument also applies to views that accept an epistemic constraint on reasons that is weaker than the evidence constraint. We may once again assume that a colleague of Jill’s knows whether B or C is the cure. We may also assume that Jill knows that her colleague knows. As a consequence, if she believes in some weaker form of objectivism in such a situation, she would still have to believe that she ought not to give drug A. But if she cannot reach her colleague, and she cannot wait any longer to begin the treatment, the only responsible decision for her to make is still to give drug A.

#### *Evaluating the misguidance argument*

Let us reconsider the premises of the misguidance argument. That the only responsible decision for Jill is to give drug A (1) is beyond dispute. That objectivism entails that Jill ought not to give A (2) is similarly uncontroversial: to my knowledge, it has not been denied and is often explicitly embraced by objectivists.<sup>24</sup> It could only be denied by rejecting the principle that one ought to omit actions that are, or are known to be, incompatible with actions that one ought to perform, which is intuitively plausible and can be supported by independent arguments.<sup>25</sup> (4), likewise, seems beyond reasonable dispute. If it does not constitute a *reductio* of a view about the deliberative ‘ought’ that it forces Jill, on pain of irrationality, to risk John’s death, then I do not know what

---

<sup>24</sup> For example by Bykvist (2011, 34–35) and Graham (2010, 97–98).

<sup>25</sup> In my “Instrumental Normativity” (unpublished), I argue that this principle is entailed by the best explanation for the fact that ‘oughts’ must be jointly satisfiable.

would. It thus seems that the only way for objectivists to avoid the conclusion is to deny (3): the claim that it is irrational to decide in favour of an action that one believes one ought not to perform.

This claim, however, seems essential for the deliberative 'ought'. It is the very point of practical deliberation to seek guidance in decision-making. It cannot be rational to maintain a deliberative conclusion to the effect that one ought not to perform an action and then decide to do it nonetheless (it might, of course, be rational to change one's mind about what one ought to do, but this is beside the point). So in abandoning (3), the objectivist seems to concede that the 'ought' he has in mind is not, after all, the 'ought' that we employ in deliberative conclusions.

The objectivist might reply that (3) has to be qualified, just as the claim (A)\* that it is irrational to believe one ought to do something without intending to do it needed to be qualified. Two remarks are in order. First, the reason for qualifying (A)\* was that it does not seem irrational not to intend an action one believes one ought to perform if one expects that one will perform the action regardless of whether or not one intends it. But this reason does not seem to apply to (3). It remains intuitively irrational to intend an action one believes one ought not to perform *even if* one believes one will not perform this action regardless of one's intention. Hence, the reason for qualifying (A)\* does not likewise speak for a qualification of (3). Second, even if (3) is qualified in this way, that does not help the objectivist to escape the misguidance argument, for we may safely assume, as part of the example, that Jill knows that she will give drug A when she decides to do this.

So the objectivist must have another qualification for (3) in mind. And this qualification would seem to be that it is irrational to decide in favour of an action that one believes one ought not to perform *only if one believes of a particular alternative option that it is permitted*.<sup>26</sup> If Jill believes in the objective view, then she does not believe of a particular option that this option is permitted. Hence, the qualification of (3) saves her from being irrational in deciding to give drug A. One point to note about this reply is that it seems *ad hoc*; it introduces a condition on an intuitively plausible claim in order

---

<sup>26</sup> Some remarks of Bykvist (2011, 39) suggest that he would recommend such a qualification.

to save a controversial theoretical assumption. The objectivist might argue that Jill's case itself suggests the relevant qualification of (3). But this would only be so if it were independently plausible that Jill ought not to give A, which is not the case. As pointed out above, it is at least *prima facie* plausible to say that Jill can correctly conclude that she ought to give A.<sup>27</sup> The relevant qualification of (3) therefore cannot be justified by plausibility assumptions about Jill's case. It is not well motivated.

Second, if the objectivist runs this manoeuvre, he still owes us an explanation of how Jill can rationally make the decision for A. As an objectivist, Jill so far only believes that she ought not to give A, and does not believe of any other particular option that it is permitted – how is she going to decide, then? There must be some kind of judgement that licenses her to rationally decide in favour of A despite her judgement that she *ought not* to give A.<sup>28</sup> The objectivist, in effect, has to say that even though Jill ought not to give A, she *shmought* to give A. And he will then have to agree that an agent cannot be rational if she decides in favour of an action that she believes she *shmought* not to do. So for example, Krister Bykvist says that even though Jill ought not to give A, it is “sensible” for Jill to give A, it is “rational to prefer” giving A; and Jill “should be willing” to give A.<sup>29</sup> Similarly, Peter Graham claims that even though Jill morally ought not to give A, Jill ought to give A in a “pragmatic” sense “associated with means and ends”.<sup>30</sup> And Joseph Raz suggests that even though Jill's “best reason supports one of the other drugs”, Jill could rationally decide for A on grounds of the judgement that doing so “is the best approximation to what she has best reason to do in the circumstances”.<sup>31</sup>

Any such suggestion faces the same dilemma. Either we have sufficient reason to do what we *shmought* to do or not. If we have sufficient reason to do what we *shmought* to do, then, since Jill *shmought* to give A, Jill has sufficient reason to give A. But then objectivism is false, for objectivism entails that Jill has decisive reason not to give A, and one cannot have sufficient reason for an action one has decisive reason not

---

<sup>27</sup> Note that Bykvist agrees with this; he explicitly says that giving A is “the intuitively right option” (2011, 34).

<sup>28</sup> Andric (2013a) considers various options for what this judgement might be and convincingly refutes all of them.

<sup>29</sup> Bykvist (2011, 38–39).

<sup>30</sup> Graham (2010, 103)

<sup>31</sup> Raz (2011, 124).

to perform. If, on the other hand, we do not have sufficient reason to do what we shmought to do, it is unclear how we could rationally make decisions on grounds of 'shmought'-judgements. To say that we may not have sufficient reason to do what we shmought to do is to admit that the normative question "Why do what I shmought to do?" may not have an answer even though it can reasonably be asked. But if this is so, then it seems that we can rationally ignore what we shmought to do. The point is that what has normative authority for us are reasons; judgements about what is rational or sensible, even judgements about what is the best approximation to what one has reason to do, do not have normative force and thus cannot guide our decision-making *unless we assume that we have sufficient reason to follow them, which is exactly what the objectivist needs to deny.*

In any case, when the objectivist says that Jill shmought to give A, he ultimately concedes that there is an important concept that figures in deliberative conclusions, guides our decision-making, and makes an unqualified version of (3) true, and he concedes that this concept is not to be understood in the objectivist fashion. I will continue to call this concept "ought". As far as this concept is concerned, the misguidance argument stands. The objective view is not the correct theory about the deliberative 'ought'.

#### 8.4 The problem of advice

I have argued that the objective view is vulnerable to serious problems that the evidence-relative view avoids (as are related views that accept some constraint on reasons that is weaker than the evidence constraint). Yet the evidence-relative view faces its own challenges. In particular, it has problems accounting for phenomena in which the use of 'ought' is geared to evidence that is better than the evidence currently available to the agent. In the remainder of this chapter, I set out to meet this challenge by developing an account of reasons that captures the phenomena in question. I will start by discussing the problem that cases of *advice* pose for perspectivism.

When people ask themselves what they should do, they sometimes ask others for advice. It is thus natural to assume that advice typically seeks to answer the deliber-

ative question *for* the agent; advice and deliberation share the same topic.<sup>32</sup> The evidence-relative view claims that what an agent ought to do in the sense relevant for deliberation depends on this agent's evidence. It claims, for example, that since Jill has no evidence whether B or C provides the cure, she ought to give A. But now let us suppose that a third person enters the scene, one who knows that C is the cure and is able to disclose this information to Jill. Naturally, the third person will give Jill the following advice: "You ought to give C!" Indeed, it seems that this would be the only good advice a person with such knowledge could give. Yet on the face of it, the evidence-relative view entails that this adviser's statement would be literally false, since at the time of the advice, Jill's evidence-based reasons favour giving A and not C. Apparently, then, the evidence-relative view entails that a better-informed adviser cannot truly give the only good advice in the situation described. Call this the problem of advice.<sup>33</sup>

The problem of advice poses a serious challenge for the evidence-relative view: When Jill asks the adviser, "What ought I to do?" it seems that she does not want a report about what her current evidence already tells her to do, but rather hopes for something beyond that. And it would be more than natural to suppose that the better-informed adviser will answer her question correctly – and in the very sense that Jill was after when asking for advice – when he makes his answer dependent on his better knowledge and *not* on the evidence of the agent. On the face of it, however, it seems that the evidence-relative view cannot account for this phenomenon.

A common reply on behalf of the evidence-relative view maintains that by giving advice in accordance with superior information, advisers anticipate the truth of their own statement in the nearby future. In the process of advising Jill to give C, the adviser thus *changes* Jill's evidence and thereby *makes* it true that she ought to give C.

---

<sup>32</sup> I take this characterisation from Kolodny and MacFarlane (unpublished), to whom the discussion of this section is very much indebted. See also Williams (1989, 40): "advice aims to offer something as a candidate for a deliberative conclusion." What is central is that advice seeks to answer the question that agents *would* ask in (and thus has the same object as) deliberation; it is not necessary that agents *actually* deliberate about it.

<sup>33</sup> For this line of objection to perspectivism see, e.g., Kolodny and MacFarlane (2010, 119–120); Thomson (2008, 187–191). I stipulate in all adviser-cases that the adviser *knows* (and not merely justifiably believes) that C is the cure, because the problem of advice is driven in part by the intuition about the *truth* of an 'ought'-statement made by a better-informed person.

According to this suggestion, saying “You ought to give C” is a bit like saying “You want to go left at the light”, when we are asked for directions – not literally true, but anticipating a truth that in part will be brought about by the utterance itself (or by accompanying elucidations).<sup>34</sup>

While I agree that the adviser’s interference in the situation and the transmittance of information plays a crucial role for the understanding of the problem of advice, I believe that this suggestion is a confession of failure rather than a solution to the problem. The reply admits that the adviser’s statement is strictly speaking false at the time it is made. It thus cannot account for the intuition that an adviser gives the correct answer to the question an agent is asking when he grounds his advice in his better knowledge, but instead reduces his answer to a merely rhetorical device. A non-literal interpretation of “You want to go left at the light” seems unproblematic, since advice clearly does not aim at telling the agent what he already wants to do. A non-literal interpretation of “You ought to give C”, on the other hand, comes at a high price, since it must dismiss the plausible idea that advice typically aims at providing an answer to the question agents ask in deliberation.

Similar worries apply to responses to the problem of advice that allude to different notions of ‘ought’.<sup>35</sup> I said in the beginning that the ‘ought’ figuring in deliberative conclusions cannot be ambiguous between different qualified notions of ‘ought’. For the time being, this leaves open the possibility of other senses of ‘ought’ that do not compete with the deliberative ‘ought’. Perhaps there is a genuine second- or third-personal ‘ought’ that is different from the notion employed in deliberation. This ‘ought’ might be understood in the objectivist fashion, but more plausibly it will relate to the evidence of the second or third person.<sup>36</sup> On this construal, an adviser can truly say, “You ought to give C”, meaning to state that Jill ought, relative to his – the adviser’s – evidence, give C, which is compatible with the claim that Jill ought to give A in the deliberative sense.

---

<sup>34</sup> Cf. Prichard (1932, 94); Ross (1939, 152–153). For a recent version of this reply, see Gibbons (2010b, 356), who is also the source of the analogy to sentences of the form “you want to go left at the light”.

<sup>35</sup> See e.g. Jackson (1991, 471–472).

<sup>36</sup> It is more plausibly relative to the evidence of the other person because that other person can find herself in the same informational state as Jill did before, and it would be irresponsible of her to give advice that could only be followed by imposing a 50 per cent risk of death on the patient.

But while this suggestion preserves the truth of the adviser's statement, it does so only by depriving the adviser and the agent of a common subject matter. As a result, we are forced to give up the natural idea that advisers seek to answer the agent's question "What ought I to do?" for them. For according to the suggestion, the adviser is concerned with what the agent ought to do *relative to the adviser's evidence*, while this is not the concern of the agent.

As Niko Kolodny and John MacFarlane have argued, the suggestion also stands in conflict with the phenomena of disagreement and retraction that we find in these cases. Suppose that Jill considers her options and concludes, "I ought to give A". Knowing that C provides the cure, an adviser could express disagreement with this conclusion, maintaining, "*No!* You ought to give C." But if he meant 'ought' in a sense different from the one employed by the agent, then there would not be any disagreement between them. Moreover, it seems that Jill could retract her earlier statement, saying "Oh, you're right – what I really ought to do is give C". Again, this would not make much sense if both original statements addressed different issues. If the dialogue we imagined is not an expression of complete confusion, it seems it must have *one* sense of 'ought' as a common subject – the very sense of 'ought' that an agent is after when he asks the deliberative question, "What ought I to do?"<sup>37</sup>

To sum up, proponents of the evidence-relative view seem forced to accept that an adviser who grounds an 'ought'-statement in his better knowledge either utters a sentence that is literally false or talks past the agent by employing a different notion of 'ought'. In both cases, they have to reject an attractive conception of advice and its engagement with deliberation that emerges naturally from linguistic observations. In what follows, I will put forward and defend a solution to the problem of advice that avoids these results.

---

<sup>37</sup> Cf. Kolodny and MacFarlane (unpublished). In defence of different notions of 'ought', Björnsson and Finlay (2010, 17–25) present an interesting solution to the disagreement and retraction cases. Whether or not their solution is convincing, however, they ultimately have to deny the plausible idea that advice, at least in part, seeks to answer the deliberator's question "What should I do?"

### 8.5 Reasons, time, and advice

In order to solve the problem of advice, we need to take account of time. Reasons are sensitive to time in two ways. The first of these concerns the favoured response. Suppose that you need mouth-to-mouth resuscitation in order to survive, and I am the only one who is in a position to help. Other things equal, I have a reason to perform mouth-to-mouth resuscitation. If I hesitate too long, and start doing it only when you are already dead, then I did not conform to this reason. It follows that the reason implicitly referred to a period in time; it was a reason to perform mouth-to-mouth resuscitation *at a particular time t*. The second way in which reasons are sensitive to time concerns the occurrence of the reason itself. For example, before you needed help, I did not have a reason to perform mouth-to-mouth resuscitation, and once you no longer need my help, I also no longer have this reason. Hence, a statement to the effect that A has a reason to  $\phi$  implicitly refers to two points in time; it states that *at some time t*, A has a reason to  $\phi$  *at some time t\**.

The relevant points in time might be the same, but they need not be. For example, when I promise you on Monday that I will help you move on Wednesday, then *on Monday* I come to have a reason to help you *on Wednesday*. This is important, since reasons create further instrumental reasons, e.g. reasons against incompatible actions. When I promised you on Monday to help you on Wednesday, then on Tuesday I have a reason against taking off for a one-week road trip. It is thus essential that the reason to help you is already present before the favoured action is supposed to take place. Yet it is present only from the moment on at which the promise was given: on Sunday, I might have had no reason against taking off for a one-week road-trip.

This time-sensitivity of reasons requires a clarification of the evidence-relative view. A synchronic version of the evidence constraint reads:

*Synchronic evidence constraint:* At  $t$ ,  $R$  is a reason for  $A$  to  $\phi$  at  $t$  only if at  $t$ ,  $A$ 's evidence includes  $R$ .

I take this claim to be supported by the arguments against the objective view given above. However, in many situations we are interested in the truth conditions of reason statements that are not in this way synchronic. For example, both deliberation and advice proceed on the assumption that the favoured response takes place *after* the judgement that a reason is present is made. For such diachronic cases, at least two different interpretations of the evidence constraint are available:

*Judgement-relative evidence constraint:* At  $t_1$ , R is a reason for A to  $\phi$  at  $t_2$  only if at  $t_1$ , A's evidence includes R.<sup>38</sup>

*Response-relative evidence constraint:* At  $t_1$ , R is a reason for A to  $\phi$  at  $t_2$  only if at  $t_2$ , A's evidence includes R.

The judgement-relative constraint requires that the reason-giving fact must be included in the agent's evidence at the time at which the reason judgement is supposed to hold. The response-relative constraint, in contrast, requires that the reason-giving fact must be included in the agent's evidence at the time at which he is supposed to give the response. Both of these constraints entail the synchronic evidence constraint (when  $t_1=t_2$ ), but they have different implications for diachronic cases.

An example: There is an earthquake on Sunday, which engulfs thousands of people into misery. Monday is a bank holiday, but on Tuesday you could transfer some money to an aid organisation. You do not know about the earthquake until Monday afternoon. Monday at noon, someone says that you have a reason to donate money on Tuesday. This is compatible with the response-relative constraint, but not with the judgement-relative constraint. According to the response-relative constraint, what counts is that by Tuesday you will know about the earthquake. If this condition is satisfied, then from the moment of the earthquake on, you might have a reason to donate money on Tuesday. According to the judgement-relative constraint, what

---

<sup>38</sup> This claim corresponds to the view that Kolodny and MacFarlane attack: "*S* ought (at *t*) to  $\phi$  iff  $\phi$ -ing is the best choice available to *S* in the light of what *S* knows at *t*" (2010, 118). They briefly consider the response-relative view I am defending here in their companion paper (unpublished); my adoption and development of it has greatly benefited from their discussion.

counts is that you know about the earthquake at the time of the reason judgement. Since the judgement is made before you know about the earthquake, it is false.

The difference is crucial for the problem of advice. The discussion of this problem has so far taken for granted an interpretation of the evidence-relative view along the lines of the judgement-relative constraint, according to which what counts is the agent's evidence at the time of the advice rather than the time of the action. As I shall now argue, the alternative interpretation along the lines of the response-relative constraint not only provides an attractive solution to the problem of advice, it is also the more plausible view for independent reasons.

According to the response-relative evidence constraint, a fact can be a reason for a person to do something at a certain point in time only if *at that time* the agent's evidence includes this fact. Hence, statements according to which this person has a reason can be true *before* the agent's evidence includes the relevant fact. Since the deliberative 'ought' is a function of an agent's reasons, the same is true for 'ought'-statements. This interpretation of the evidence-relative view suggests a simple solution to the problem of advice: A better-informed adviser can truly state, "You ought to give C", since what Jill ought to do depends only on her evidence at the time of her action, and by that time her evidence will include the fact that drug C is the cure – because the adviser will have told her.<sup>39</sup>

This solution bears some resemblance to the one I rejected earlier on the ground that it reduces advice to a merely rhetorical device. But in contrast to that solution, the response-relative constraint allows for the literal truth of the adviser's statement and thus ensures that advisers can take their advice to be a straightforward answer to the agent's deliberative question, "What ought I to do?" – for this question is now itself understood as reaching beyond the current state of evidence of the agent, aiming rather at the evidence that can be collected until the decision has to be made. The adviser may sensibly take herself to be disagreeing with Jill's earlier conclusion, and

---

<sup>39</sup> Normally, it will be this additional transmittance of evidence, which usually accompanies advice, that ensures the truth of the 'ought'-statement, but it is also possible that the adviser's 'ought'-statement *itself* provides the relevant evidence, e.g. because it makes it the case that relative to the agent's evidence, C is likely to be the cure.

Jill, after taking into account the adviser's information, may sensibly retract her earlier belief and acknowledge that she was wrong to believe that she ought to give A.

The picture of advice emerging from this view ascribes to it a dual character. Like other statements, 'ought'-claims in advice are subject to the standard of truth that functions as an ideal for sincere utterances. At the same time, however, advice aims at improving the agent's decision situation, e.g., by providing additional information. Therefore, advice does not just passively represent a certain state of affairs, but actively contributes to its obtaining; it thus might very well cause its own truth. To borrow from Bernard Williams, "advice...has to be understood, in part, in terms of its own intended effects."<sup>40</sup> This certainly is a peculiarity that might need some getting used to, but on reflection I think we have to accept it.

Once we acknowledge that proponents of the evidence constraint have to make a decision about the time at which the evidence counts as relevant, it falls into place that this point should be the time of the action, not that of the judgement. Compare the parallel question in rational choice theory. It seems obvious that the rationality of a choice at a certain point of time  $t$  depends on the evidence or beliefs that the agent has at  $t$ , not on the evidence or beliefs the agent has at the time at which that question is considered. *Prima facie*, I see no reason to suppose that things should be different with respect to what the agent ought to do. Or consider *ex post* judgements about what an agent ought to have done in the past. It would be against the spirit of the evidence-relative view to think that what the agent ought to have done can depend on information that she might have gained years after she had to make the decision, so it is very natural to say that what she ought to have done depends on the evidence she had when she was supposed to act. But if this is the correct account of *ex post* judgements, it would be astonishing if it were not the correct account of prospective judgements, too.

Moreover, the view I have suggested not only accommodates second- or third-personal phenomena, such as advice, but can plausibly be motivated from the perspective of the first person as well. For instance, suppose that Jill knows that other people in the room know which treatment – B or C – is the cure. In this case, it does not seem

---

<sup>40</sup> Williams (1989, 42).

appropriate for her to claim confidently, “I ought to give A”, but rather to say, “I don’t know what I ought to do.”<sup>41</sup> This is completely in line with the response-relative constraint, but it would not make much sense if the relevant evidence were the evidence at the time of the judgement. Or suppose that Jill needs to make the decision tomorrow in order to prevent a serious deterioration of the disease. She may say to herself today: “I believe that I ought to give treatment A, because that is what my evidence suggests now. However, let’s see whether I can gather more information until I actually have to make a decision.” This thought seems to presuppose that she can *correct* her belief that she ought to give treatment A on the basis of *new* information. And, indeed, when Jill learns that treatment C is the cure the next day, she may say to herself, “Yesterday, I thought I ought to give treatment A, but now I know that I was wrong. I ought to give treatment C.” If we want these first-personal thoughts to express truths, then we should take ‘ought’-statements to be dependent on the evidence available at the time of the act, not the evidence available at the time of the judgement.

It thus seems that the response-relative interpretation of the evidence constraint not only provides an attractive solution to the problem of advice, but also accounts for important first-personal phenomena, and is hence more plausible than its rival for independent reasons.

## 8.6 The problem of seeking and sharing new evidence

The case of advice, however, raises another question that poses difficulties for the evidence-relative view: Why should an adviser with better information bother to share his information with an advisee? According to the response-relative interpretation, by transmitting his evidence to Jill the adviser *makes* it true, as it were, that she ought to give C. If, on the contrary, the adviser withholds his information, then it is true that the doctor ought to give A, because that is what her evidence will suggest at the time of her action. But if deliberation seeks to answer the question “What should I do?”, and advice seeks to help answer that very question, then it seems hard to explain why sharing the information counts as *helping* the agent. After all, withholding the informa-

---

<sup>41</sup> See Björnsson and Finlay (2010, 13) for a similar point.

tion would make it just as easy – if not easier – for the agent to find a true (albeit different) answer to the question both are trying to answer.<sup>42</sup>

Let me be clear about what is, and what is not, the question that poses a problem for the evidence-relative view. The problem is not to answer the question: Why ought the adviser to share the information? According to the evidence-relative view, the adviser ought to share the information because, for example, his evidence suggests that doing so will likely lead to the cure of the patient. The question is: Why would the adviser's sharing his information count as *advice*, as helping to find a correct answer to the agent's deliberative question? This question provides a serious challenge for the evidence-relative view, for if the truth of an answer to the deliberative question is relative to whatever evidence the doctor happens to have at the time of her action, how can improving that evidence contribute to finding a correct answer?

Even worse for the evidence-relative view, a similar question arises even in the absence of an adviser, from the perspective of the first person alone. It seems that when deliberating about what to do, agents sometimes seek evidence that is not yet available to them. Again, the problem is not to explain why agents sometimes ought to seek new evidence. According to the evidence-relative view, one ought to seek evidence just when one's evidence-based reasons suggest seeking evidence. What remains unexplained is why an agent's looking for better evidence can contribute to finding an answer to her deliberative question. According to the evidence-relative view, the occurrence of new evidence *changes* what the agent ought to do. But if that is so, how can seeking new evidence contribute to finding out what the agent ought to do? If what the agent ought to do depends on the evidence that she happens to have at the time of the action, then it seems that in order to find out what she ought to do, she could just as well do nothing or even destroy available evidence – indeed this could be very helpful in coming to a clear-cut deliberative conclusion!

Let me start with the first-personal problem: How can seeking new evidence contribute to finding out what one ought to do? Objectivism has a straightforward

---

<sup>42</sup> Compare Björnsson and Finlay (2010, 14–16) for a discussion of a similar question that can be directed at their contextualist view. They argue that, despite appearances, Kolodny and MacFarlane's relativist account faces the same difficulty.

answer to that question: It will normally bring one closer to what one objectively ought to do. But we have rejected objectivism above, so we need an alternative answer to that question. On the face of it, the question seems perfectly legitimate, and the evidence-relative view cannot make sense of it. Consider Jill again, this time on the eve of her decision. She asks herself, “What ought I to do tomorrow? Give A? Or go for B or C?” The evidence available to her *now* suggests giving treatment A, and if she does not seek more evidence, that is also what her evidence will suggest tomorrow. But suppose that if she put a lot of effort into researching new evidence that night, then tomorrow she would have sufficient evidence to give C and really cure the patient of his disease. Why ought she to seek this evidence? Proponents of the evidence-relative view can answer this question: Presumably, her evidence now suggests that by seeking evidence she could make her patient’s life significantly better. So far, so good. It seems, though, that we want to say more. We want to say that, if Jill really spends all night seeking evidence, she could perfectly well understand herself not only as doing what she now ought to do (namely, seeking evidence), but also as making an effort to find out what treatment she ought to give tomorrow. And this is puzzling for a view that holds that what treatment she ought to give tomorrow depends on what her evidence will suggest tomorrow, regardless of whether she will have sought further evidence or not.

#### *A possibilist reply*

The problem is that the current interpretation of the response-relative constraint is *actualist* rather than *possibilist*: it ties reasons for future responses to the evidence that the agent will *happen to have*, rather than the evidence that the agent *could* have under favourable circumstances. If what treatment the doctor ought to give depended on the *best* evidence she *could* have tomorrow, then we would have an explanation for why seeking more evidence can be a way of finding out what one ought to do. However, it seems too strong – and against the spirit of the evidence-relative view – to say that what one ought to do depends on the *best* evidence that one could come to have at the time of the action. Suppose that Jill’s evidence on the eve of her decision strongly suggests seeking new evidence in the library by going through various medical journals.

In fact, she cannot find any new evidence in these journals, but she could find evidence if she were to call a former colleague who, unbeknownst to her, just discovered the relevant information days ago. So if what treatment Jill ought to give tomorrow depends on the *best* evidence she could come to have by tomorrow, then it is true that she ought to give C. But if she does what her evidence suggests she do all along, then she will never have the best evidence that she could have, and she will never be in the position to responsibly give treatment C. It is against the spirit of the evidence-relative view that these counterfactual truths could influence what Jill ought to do, independently of any evidence that was ever really available to her.

What I suggest instead is that the relevant evidence is the evidence that the agent *would* have if she conformed to her decisive reasons and thus sought evidence as she ought to. This gives us the following interpretation of the evidence constraint:

*Possibilist response-relative evidence constraint:* At  $t_1$ , R is a reason for A to  $\phi$  at  $t_2$  only if A's evidence would include R at  $t_2$  if A conformed to her decisive reasons at every  $t$  between  $t_1$  and  $t_2$ .

This is a *possibilist* interpretation of the response-relative evidence constraint, because it deems as relevant not the evidence that the agent *actually will have*, but the evidence that she *could have*, where the relevant sense of 'could' is given by those possible worlds in which the agent conforms to her decisive reasons.

The view I am suggesting can be spelled out in some more detail. I take as basic the idea that a fact R is, at a certain time  $t$ , a (*pro tanto*) reason for A to  $\phi$  at  $t$ . The synchronic evidence constraint here entails that this can only be so if A's evidence at  $t$  includes R. Whether at  $t$  A has *decisive reason* to  $\phi$  at  $t$  depends on her *pro tanto* reasons at  $t$  for and against  $\phi$ -ing at  $t$  (and possibly for and against other responses that she could give at  $t$ ). Against this background, we can now give the following characterisation of *pro tanto* and *decisive* reasons for future responses:

*The evidence-relative view:*

*Pro tanto reason:* At  $t_1$ ,  $R$  is a reason for  $A$  to  $\phi$  at  $t_2$  only if (i) at  $t_1$ ,  $R$  is the case, (ii) if  $A$  conformed to her decisive reasons at every  $t$  from  $t_1$  to  $t_2$ , then  $A$ 's evidence at  $t_2$  would include  $R$ , and (iii) if  $A$ 's evidence at  $t_2$  included  $R$ , then, at  $t_2$ ,  $R$  would be a reason for  $A$  to  $\phi$  at  $t_2$ .

*Decisive reason:* At  $t_1$ ,  $A$  has decisive reason to  $\phi$  at  $t_2$  only if (i) at  $t_1$ , there are some facts  $R_1, \dots$ , and  $R_n$ , such that (ii) if  $A$  conformed to her decisive reasons at every  $t$  from  $t_1$  to  $t_2$ , then  $A$ 's evidence at  $t_2$  would include  $R_1, \dots$ , and  $R_n$ , and (iii) if  $A$ 's evidence at  $t_2$  included  $R_1, \dots$ , and  $R_n$ , then, at  $t_2$ ,  $R_1, \dots$ , and  $R_n$  would provide decisive reason for  $A$  to  $\phi$  at  $t_2$ .<sup>43</sup>

Because of condition (ii) in the characterisation of *pro tanto* reasons, this view obviously entails the possibilist response-relative evidence constraint. Perhaps less obviously, it also entails the synchronic evidence constraint. For if  $t_1=t_2$ , then the antecedent of (ii), which requires that  $A$  conforms to her decisive reasons at every  $t$  between  $t_1$  and  $t_2$ , is trivially satisfied. Consequently, if  $t_1=t_2$ , then (ii) requires that  $A$ 's evidence at  $t_1=t_2$  includes  $R$ , as the synchronic evidence constraint claims.

This possibilist interpretation of the evidence-relative view provides a solution to the problem of seeking evidence. Jill, on the eve of her decision, deliberates about what treatment to give. What treatment she ought to give does not depend on the evidence that she will happen to have, regardless of what she does until tomorrow, but rather on the evidence that she would have if, from now on, she did what she ought to

---

<sup>43</sup> These formulas might also be understood as giving necessary *and* sufficient (rather than merely necessary) conditions. However, the claim about sufficient conditions is more vulnerable, and the claim about necessary conditions suffices for the present purposes (I am currently uncertain with respect to the claim about sufficient conditions). This statement of the evidence-relative view contains some revisions of the claim that I called "dynamic perspectivism" in earlier work, and which I stated as follows: "A ought (at  $t_1$ ) to  $\phi$  at  $t_2$  if, and only if, A ought (at  $t_1$ ) to  $\phi$  at  $t_2$  relative to the evidence that would be available to  $A$  at  $t_2$  if  $A$  were seeking evidence from  $t_1$  until  $t_2$  as  $A$  ought (at every  $t$  between  $t_1$  and  $t_2$ ) to" (Kiesewetter 2011b, 16). One difference is that the view suggested here takes as relevant the possible worlds in which one conforms to *all* one's decisive reasons, and not only to one's decisive reasons for seeking evidence. This step is suggested by a general possibilist conception of the deliberative 'ought' (see below). A further difference is that the view suggested here requires that statements about reasons or 'oughts' generally presuppose the presence of the relevant reason-giving fact. It should be clear that this is a requirement, yet it simply did not occur to me before I tried to express the evidence-relative view in terms of reasons rather than 'oughts'.

and thus also sought evidence as she ought to. This explains why we can understand her seeking evidence not only as doing something that she ought to do right now, but also as a way of finding out what she ought to do tomorrow.

### *Implications*

The view allows for the same utterance, e.g., “Jill ought to give A at t”, to be true when uttered at one point in time, but false at another. This is because it contains a contextualist element that relativises reason-statements to a certain time, which might be different from the time of the action. Suppose, for example, that Jill has to make a decision tomorrow afternoon (at  $t_3$ ). At the eve of her decision (at  $t_1$ ), her evidence suggests giving A tomorrow (at  $t_3$ ), but also seeking further evidence now, and seeking further evidence would in fact reveal to her, by  $t_3$ , that C is the cure. In such a case, the possibilist view implies that Jill ought (at  $t_1$ ) to give treatment C at  $t_3$ . But suppose that, contrary to what her evidence suggests, Jill does *not* seek further evidence. On the next morning (at  $t_2$ ), she wakes up and asks herself again what treatment she ought to give at  $t_3$ . Now assume that her evidence no longer suggests seeking further evidence. Suppose that if Jill started to seek evidence now (at  $t_2$ ), this would no longer lead her to improve her evidence until  $t_3$  – simply because in order to find better evidence, she would need more time than she has until  $t_3$ . The possibilist view then entails that Jill ought (at  $t_2$ ) to give treatment A at  $t_3$  – again, because this is what her evidence at  $t_3$  would suggest if she were doing what she ought to do at every t from  $t_2$  to  $t_3$ . Hence, the possibilist view entails some fluctuation of what Jill ought to do at  $t_3$ . At  $t_1$ , she ought to give C at  $t_3$ ; while at  $t_2$  and  $t_3$ , she ought to give A at  $t_3$ . Generally speaking, the truth of any claim that A ought to do something at some later point in time may vary from one earlier point in time to another, depending on the actual evidence the agent has at these earlier points in time.

The idea that ‘ought’-statements implicitly refer to two points in time – a time at which the action is supposed to take place *and* a time at which the requirement is supposed to hold – is not new; it can, and has been, defended on independent

grounds.<sup>44</sup> It might be worried that because of this contextualist element, the possibilist view entails a form of relativism similar to the one that it was supposed to avoid. As I see it, however, the view is not relativist in a problematic sense. “A ought to  $\phi$ ” works like “A has a new haircut”. At any particular time, these statements are true or false independently of the speaker or the assessor. But, of course, “A has a new haircut” may be true today and false in two months. There does not seem to be any worrying relativism involved here; similarly, I do not think that we should worry in the case of ‘ought’-statements.

Moreover, as I have argued above, the relevant contextualist element is a necessary element of any plausible view that identifies ‘oughts’ with decisive reasons. Everyone should accept that “you have decisive reason to help your friend move on Wednesday” might be false on Sunday (before you have promised to help), but true on Tuesday (since the promise you have given on Monday gives you derivative reasons not to perform actions on Tuesday that are incompatible with your helping to move on Wednesday). The evidence-relative view provides a particular account of the truth-conditions of ‘ought’-statements, but it is not a particular feature of this view that ‘ought’-statements are time-relative.

Nevertheless, the particular account of time-relativity involved in the evidence-relative view raises the question of how it avoids the disagreement-related objections that have been brought forward against views that relativise ‘ought’-statements to the bodies of information of different speakers (discussed above). On a simple application of the view, an agent’s utterance “I ought to  $\phi$ ”, and an adviser’s utterance “No! You ought not to  $\phi$ ”, made at a slightly later time, would indeed have to be relativised to different points in time, with the result that they could not be taken to express a genuine disagreement. But as in the case of “A has a new haircut”, it seems very natural to suppose that “A ought to  $\phi$ ” relativises to a *period* of time rather than a single *point*. And as people can of course disagree – within a certain period of time – about whether A has a new haircut, so they can disagree about whether A ought to  $\phi$ , as long as the ‘ought’ refers to the same period.

---

<sup>44</sup> See, e.g., Zimmerman (2008, 127-28, and esp. 151-68). Cf. also Prichard (2002b).

Finally, it is important to see that introducing the possibilist interpretation of the evidence-relative view is not simply an *ad hoc* move to solve the problems that this view would otherwise face. Actualist accounts of the normative relevance of the future – i.e. accounts that, broadly speaking, take as normatively relevant only what *would* actually happen rather than what *could* happen as a result of responsible agency – are quite generally plagued with problems.<sup>45</sup> It is thus natural to think that the evidence-relative view should take a possibilist form. A minimal constraint that possibilists accept (and actualists have to deny) is that A ought to  $\phi$  only if  $\phi$ -ing is part of every maximal strategy such that if A were to carry out this strategy, A would always be doing what she ought to do.<sup>46</sup> A plausible time-sensitive version of this constraint (applied to responses in general) gives us:

*General possibilist constraint:* At  $t_1$ , A ought to  $\phi$  at  $t_2$  only if  $\phi$ -ing at  $t_2$  is involved in every maximal course of responses from  $t_1$  forward, such that if A's responses followed that course, then at every time  $t$  from  $t_1$  forward, A would be responding as she ought to at  $t$ .

In other words, continuous conformity with one's synchronic decisive reasons guarantees conformity with one's diachronic decisive reasons. This is also independently plausible. In order to conform with one's decisive reasons, one should not be required to do anything more than to give the responses that one has, at every time  $t$ , decisive reason to give at that time.

The general possibilist constraint and the synchronic evidence constraint together make the possibilist response-relative evidence constraint almost inevitable. The general possibilist constraint requires that A ought, at  $t_1$ , to  $\phi$  at  $t_2$  only if A's  $\phi$ -ing

---

<sup>45</sup> As has been forcefully argued by several authors, among them J. Ross (2012b, 75–80) and Zimmerman (1996, ch. 6; 2008, ch. 3). Actualism also has some plausible features, as esp. Jackson and Pargetter (1986) have argued. I try to explain these features within a framework that accepts the possibilist constraint (introduced below) in my “Instrumental Normativity” (unpublished).

<sup>46</sup> A strategy is maximal just when no other strategy includes it. Note that the possibilist constraint entails no particular conception about how the goodness of outcomes determines what we ought to do (which possibilists normally aim to offer); it is merely a minimal constraint that rules out actualism. Consequently, it is not vulnerable to objections against particular possibilist accounts, as discussed e.g. in J. Ross (2012b, 80–89).

at  $t_2$  is a necessary consequence of her conformity with all synchronic 'oughts' up until  $t_2$ . Plausibly, this condition is satisfied only if the following conditional holds: If A conforms to all synchronic 'oughts' up until  $t_2$ , then she ought, at  $t_2$ , to  $\phi$  at  $t_2$ . The synchronic evidence constraint tells us that the consequent of this claim holds only if the relevant reason-giving facts are part of the agent's evidence at  $t_2$ . So both these constraints together require that at  $t_1$ , A ought (and thus has a reason) to  $\phi$  at  $t_2$  only if the following conditional holds: If A conforms to all synchronic 'oughts' up until  $t_2$ , then the relevant reason-giving facts are part of her evidence at  $t_2$ . And this is exactly what the possibilist response-relative evidence constraint tells us. Hence, the possibilist evidence constraint is not simply an *ad hoc* move; it is strongly supported by the conjunction of the synchronic evidence constraint and the general possibilist constraint, which are both well motivated.<sup>47</sup>

### 8.7 Sharing evidence and the good

While the possibilist interpretation of the evidence-relative view presents an attractive solution to the problem of *seeking* evidence, it does not help with the problem of *sharing* evidence as it is posed from the perspective of an adviser. The possibilist view accounts for the idea that Jill can understand her seeking new evidence as a way of finding out what treatment she ought to give, because what treatment she ought to give does not depend on the evidence she will happen to have, but on the evidence she will have if she does what (and thus seeks for evidence as) she ought to. But this cannot explain why an adviser can understand his sharing of information as a way of helping Jill to find out what she ought to do, since what Jill's evidence would suggest if she sought for evidence as she ought to may still be different from what the adviser's better evidence suggests. We still need an explanation of how sharing this evidence can count as helping to answer a question that could just as well be answered correctly (albeit differently) in the absence of any advice.

---

<sup>47</sup> I owe thanks to an anonymous referee for the *Journal of Ethics and Social Philosophy* for calling my attention to the connections between my time-sensitive version of the evidence-relative view and possibilism in general.

The problem, I believe, cannot be solved by further revising of the evidence-relative view, but only by reconsidering the picture of advice and deliberation that it presupposes. The key is to remind ourselves that we need not deny that evidence-independent normative or evaluative considerations play *any* role in deliberation. The argument against objectivism establishes that the ‘ought’ of deliberative conclusions depends on evidence, but this does not imply that reaching such conclusions is the one and only concern of deliberative agency. On the contrary, it is quite plausible that agents in their deliberation are also concerned with the pursuit of the good, i.e. with promoting, protecting or respecting certain values. As Gunnar Björnsson and Stephen Finlay point out, better information puts agents in a better position to pursue these values.<sup>48</sup> In sharing evidence, then, advisers do not literally help agents to find a true answer to their question “What should I do?”, but they nevertheless respond to a fundamental concern of deliberative agency, namely the pursuit of the good.

What this means is that the objective good, though not directly determining the ‘ought’ of deliberative conclusions, still functions as an evaluative ideal that we strive for in practical deliberation.<sup>49</sup> The air of paradox is only apparent, as the case of belief shows. Objective truth functions as an ideal in theoretical deliberation, but it does not settle directly what we ought to believe. As a conscientious believer, I aspire to believe the truth, even though this is not what I ought to believe without qualification. Recall the coin example from above. When you know that a fair coin has landed, but you do not know how, it seems that you ought not to believe it landed heads, and you ought not to believe it landed tails. Believing either proposition, without having evidence supporting one proposition rather than the other, seems to involve not taking truth seriously. What this shows is that maintaining the ideal of objective

---

<sup>48</sup> See Björnsson and Finlay (2010, 16), in defence of their contextualist view. I make use of their idea for perspectivism, here, but adopt neither their speaker-relativism, nor the more or less consequentialist assumption that ‘ought’ has to be understood in terms of greatest expected value, nor the claim that the agent’s interest in reaching true ought-judgements is merely instrumental. It is difficult to say something about the relation between ‘good’ and ‘ought’ while remaining neutral with respect to certain substantial issues in moral theory, but it is probably safe to say that one ought to take the option with the greatest expected value in case one has no stronger reason not to take it.

<sup>49</sup> See Oddie and Menzies (1992, 513) for the thesis that “the correct regulative ideal for the moral agent is that of maximizing objective value.” As I see it, this view (which Oddie and Menzies call objectivism) is compatible with the rejection of the objective view about the deliberative ‘ought’.

truth is compatible with following an evidence-relative ‘ought’ rather than an objective one – indeed, it even seems that taking truth seriously, here, *requires* one to form one’s beliefs in accordance with the evidence rather than with the facts. Truth functions as an evaluative ideal in believing, but it does not directly settle what we ought to believe. Similarly, Jill will give the treatment that is suggested by her evidence, partly because she is concerned with the objective well-being of her patient. Again, what is objectively good functions as an ideal in deliberation, even though it does not directly settle what we ultimately ought to do. Advisers can thus understand their sharing of evidence as a response to the deliberative concerns of the agent without assuming that they are helping to answer the more specific question, “What ought I to do?”.

## 8.8 Objections

Let me address some objections against the view that I have proposed in this chapter. The first objection comes from Niko Kolodny and John MacFarlane.<sup>50</sup> They maintain that a view along the lines of the one suggested here is committed to an implausibly indirect picture of advice. Intuitively, an adviser who knows for certain that C is the cure will conclude directly that Jill ought to give C. But according to the evidence-relative view, this inference is only justified if certain conditions about the transmission of evidence are met. Thoughts about the effects of the advice, however, will appear to the adviser as being beside the point. It thus seems that the evidence-relative view presupposes an account of advice that is phenomenologically inappropriate.

I agree that in a normal case, it would be rather peculiar for an adviser to go through considerations about the effectiveness of information transmittance before issuing advice. But I disagree that the evidence-relative view requires any such thing. The view makes a claim about a necessary condition on statements about reasons and ‘ought’; it does not entail anything about what role this condition should play in one’s actual considerations. For an analogy, consider the plausible claim that “A ought to  $\phi$ ” is true only if A is alive. In most situations it would be strangely indirect and beside the point to *think* about the question of whether A is a zombie, a robot or a living being

---

<sup>50</sup> Cf. Kolodny and MacFarlane (unpublished).

before giving him advice. But that of course does not mean that A's being alive is *not* a necessary condition for the truth of "A ought to  $\phi$ ". Rather, that A is alive will normally play a role in the *background* of one's consideration, perhaps as a dispositional belief. Similarly, thoughts about how one's advice may alter the informational state of the advisee typically will not occur in the foreground of an adviser's considerations. But advice is also typically carried out on the *background* assumption that it will improve the informational state of the advisee (note that otherwise advice seems pretty much pointless). The evidence-relative view thus allows for an adviser to conclude that Jill ought to give C directly from the fact that C is the cure – the supposition that the advice will alter the agent's information will normally play its role as a background assumption and not in the foreground of the adviser's considerations.

In order to evaluate the evidence-relative view, we should therefore ask whether it provides correct conditions for 'ought'-judgements, not whether these conditions figure in the explicit considerations of advisers (or deliberators). At this point, a second objection by Kolodny and MacFarlane comes into play. They imagine a situation in which the adviser justifiably expects that he will not succeed in transmitting his evidence to the agent, maybe because the agent mistrusts him or is out of earshot. While the evidence-relative view appears to entail that Jill ought to give A in this case, we may think it natural for the adviser to insist that she actually ought to give C. The objection is that the evidence-relative view incorrectly entails that this belief of the adviser would be false.

But I think it is far from obvious that the adviser's belief would be true in such a case. Consider, first, the possibility that the adviser is out of earshot: He simply cannot reach the doctor before she has to make a decision. I think that in such a case Jill ought to give treatment A, since every other action would constitute an unacceptable risk in her situation. The adviser will of course regret that he is out of earshot, but on reflection he should agree that Jill ought to give treatment A in her situation, even though he knows that C is the cure.<sup>51</sup>

---

<sup>51</sup> Kolodny and MacFarlane (unpublished) seem to acknowledge this point later, when stating that the agent's evidence can be the relevant parameter in contexts in which advice is impossible. Note that for all that has

In the second case, the adviser is within earshot, but he anticipates that the agent does not trust him. This case is more interesting. As I see it, we should distinguish the case in which the agent's mistrust is justified from the case in which it is not. If Jill is not justified in mistrusting the adviser, then it is true, on the account I propose, that she ought to give C. This is because if she did what she ought to, which includes trusting the adviser, she would come to know that C is the cure by the time of the action. So in this case, the evidence-relative view actually entails that the adviser can truly say "You ought to give C".

Things look different if the agent's suspicion *is* justified. Under such circumstances, Jill's evidence will not support giving C even if she does what she ought to until the time of the action. The evidence-relative view therefore entails that the adviser's statement "You ought to give C" is indeed false. Again, on reflection it seems to me that this implication is correct. If Jill justifiably mistrusts the adviser, then deciding on his advice to give the treatment that may well kill the patient seems clearly too risky; hence, Jill ought to give A in these circumstances. The adviser, in turn, will find it unfortunate that the situation is such that his advice cannot justifiably be trusted, but on reflection he should acknowledge that Jill ought to give A in such a situation.

Let me turn to the third objection. Suppose that a fully informed bystander voluntarily refuses to share his information with Jill. The evidence-relative view here implies that the bystander could truly say, "You ought to give A", even though he knows that C is the cure – and this seems counterintuitive.<sup>52</sup> I agree that something is wrong with that statement, but what is wrong, I maintain, is not that it is literally false. Again, I think it is plausible that Jill really ought to give A if the bystander withholds his information, for it would clearly be too risky for her to do anything else as long as she lacks better information. So what is wrong with the bystander's statement, then? Note that the bystander could *also* truly say, "You ought to give C", if he gave up his refusal to share the evidence. Note also that what he ought to do, relative to his evidence, is make the second statement rather than the first, so here we already have a clear sense of

---

been said so far, the adviser could still truly believe that the doctor ought to give C in some qualified sense that is not at issue here.

<sup>52</sup> Thanks to an anonymous referee for the *Journal of Ethics and Social Philosophy* for both pressing this objection and suggesting a reply to it.

what is wrong with his statement. But there is also a pragmatic explanation of why his statement seems wrong. Given that deliberation is concerned not only with what one ought or has reason to do, but also with the pursuit of values, and given that better information puts agents in a better position to pursue these values, as suggested in the last section, it is plausible to assume that ‘ought’-statements generally carry the implicature that they are based on the best information that can be made available to the agent. And this implicature is, of course, false when the bystander states, “You ought to give A”, even though he knows that C is the cure. Hence, what is wrong with the bystander’s statement is that it carries a false implicature, not that it is literally false.

To sum up, neither objection poses an actual problem for the evidence-relative view. *Pace* Kolodny and MacFarlane, it does not require an implausibly indirect account of an adviser’s considerations. Moreover, the possibilist response-relative interpretation of the evidence-relative view subtly distinguishes between different kinds of cases in which a better-informed bystander anticipates that he cannot transmit the evidence to the agent, or refuses to do so voluntarily.

Let me finally turn to the worry that there are legitimate uses of “ought” that the evidence-relative view is not able to capture. If what I have argued in this chapter is correct, then an agent’s reasons, and consequently the truth of her deliberative conclusions about what she ought to do, depend on her evidence. Moreover, my account explains how the ‘ought’ of deliberation and the ‘ought’ of advice can still be the same concept. Though I have not explicitly argued for this, there seems to be no reason to think that the ‘ought’ used in criticism is any different. Hence, the evidence-relative view seems to provide the correct truth conditions for ‘ought’ as it is used in the three central domains of normativity: deliberation, advice, and criticism.

But perhaps this is not enough. Some authors think that there is an important use of “ought”, which is independent of the agent’s evidence, and need not have a function in either deliberation, advice, or criticism.<sup>53</sup> This would be a sense in which someone who knows that C is the cure, but also knows that there is no way for Jill to find this out, could still truly say “Jill ought to give C”. I did not deny that there is such

---

<sup>53</sup> See e.g. Henning (forthcoming, 11).

a sense of ‘ought’; my claim was merely that this is not the sense relevant in deliberation, advice, or criticism. I admit that it is natural to think that in some contexts, e.g. in moral theorising, we can reasonably abstract from epistemic limitations of individuals and still use the word “ought” in a legitimate way. For the account offered here, this poses the question of how we should conceive of such an evidence-independent notion of ‘ought’, and how it is related to the deliberative ‘ought’. I shall conclude with a suggestion about how this question might be answered.

One possibility is that the notion in question is the non-agential evaluative ‘ought’, which signifies that it would be desirable, good, or best if some state of affairs were to obtain (as in “there ought to be world peace”).<sup>54</sup> It seems, however, that the proponent of a non-teleological view in ethics at least need not have this sense of ‘ought’ in mind when she uses ‘ought’ in abstraction from the agent’s epistemic limitations. My alternative suggestion is that the notion in question could be elucidated by the concept of a *potential reason*: a fact that would be a reason if it were part of the agent’s evidence. Accordingly, we might say that certain contexts license a use of “ought” that is provided by potential reasons:

*Potential decisive reason*: A has potential decisive reason to  $\phi$  iff there are some facts  $R_1, \dots,$  and  $R_n,$  such that if A’s evidence included  $R_1, \dots,$  and  $R_n,$  then  $R_1, \dots,$  and  $R_n$  would provide decisive reason for A to  $\phi$ .<sup>55</sup>

It is important to see that a sense of ‘ought’ that is provided by what an agent has *potential* decisive reason to do is only derivatively normative, and is not directly guiding: Jill could believe that she ought, in this sense, not to give A, while rationally deciding to give A on grounds of her judgement that she has in fact decisive reason to give A. Yet the relation to the guiding, directly normative ‘ought’ is obvious: all that it takes for a potential reason to become actual is that the agent comes to know the relevant reason-

<sup>54</sup> See e.g. Andric (2013b, 5–6) for this suggestion. For helpful discussion of the deliberative and the evaluative ‘ought’, see Schroeder (2011b).

<sup>55</sup> A full definition that takes account of time is this: At  $t_1,$  A has potential decisive reason to  $\phi$  at  $t_2$  iff at  $t_1,$  there are some facts  $R_1, \dots,$  and  $R_n,$  such that if A’s evidence at  $t_2$  included  $R_1, \dots,$  and  $R_n,$  then, at  $t_2,$   $R_1, \dots,$  and  $R_n$  would provide decisive reason for A to  $\phi$  at  $t_2.$

giving fact. Insofar as there are contexts in which normative questions can be discussed by using the word “ought” in abstraction from the epistemic situation of particular agents, as seems natural to think, my conjecture is that the topic of such discussions are our potential decisive reasons.

## 8.9 Conclusion

Let me conclude by way of a summary. After discussing some preliminary objections to the objective view in Section 8.2, Section 8.3 presented an argument that I think is fatal for it. We had to reject objectivism in order to preserve the truth of:

- (a) An agent believing the correct view about ‘ought’ can rationally avoid taking a 50 per cent risk of killing someone in order to have a 50 per cent chance of avoiding a minor complaint.

I turned to difficulties for the evidence-relative view in Section 8.4, considering cases of advice. We wanted to say:

- (b) An adviser can correctly advise an agent to do what the adviser’s better evidence suggests, thereby answering the agent’s deliberative question, “What ought I to do?”

In Section 8.5, I have presented, defended against objections, and motivated independently the idea that reasons and ‘oughts’ are relative to the agent’s evidence at the time of the response rather than the time of the judgement, which allows for the truth of (b). However, two further problems with this suggestion arose. Intuitively, we wanted to say:

- (c) Seeking evidence not yet available to oneself might be a way for an agent to find an answer to her deliberative question, “What ought I to do?”

- (d) Sharing evidence otherwise unavailable to an agent might be a way for an adviser to help the agent find an answer to her deliberative question, “What ought I to do?”

As we have seen in Section 8.6, both (c) and (d) cannot be explained by an *actualist* interpretation of the evidence-relative view, according to which ‘ought’ depends on the evidence that an agent *happens to have* at the time of the response. Yet I defended a *possibilist* version of the evidence-relative view that instead takes ‘ought’ to depend on the evidence that the agent will have at the time of her action, if until that time she does what she ought to do. This can explain why (c) is true, but it cannot explain why (d) is true. As far as I can see, we cannot preserve the spirit of the evidence-relative view while holding that (d) is *literally* true. But as I have argued in Section 8.7, we can hold something similar to (d), namely:

- (d)\* Sharing evidence not yet available to an agent might be a way of helping the agent to respond to her deliberative concerns.

The truth of (d)\* can explain the most significant part of the problem of sharing evidence. It can explain why sharing evidence can count as helping the agent in her deliberation. But ultimately, we must deny that it can count as helping the agent to find an answer to the more specific question of what she ought to do. This is a bullet not too hard to bite, I think. Sharing information in advice helps the agent insofar as it brings her into a better position to pursue the good. In this respect, advice responds to the evaluative ideal of deliberation rather than to its deontic focus.

If the arguments given in this chapter are correct, then the notion of ‘ought’ or ‘decisive reason’ relevant for deliberative conclusions is a function only of an agent’s evidence-based reasons, i.e. of reasons that either are included in an agent’s body of evidence, or would be included at the time of the favoured response if the agent conformed to the reasons that already are included in the evidence. In the last chapter I have argued that rationality requires conformity with one’s decisive evidence-based reasons. So if what I have argued in both of these chapters is correct, then rationality

requires conformity with decisive reasons, period. In other words, response requirements of rationality are in the strongest possible sense normative; they are necessarily supported by decisive reasons. We can now return to the phenomenon of structural irrationality, and the question of how it might be explained in terms of response requirements of rationality.



## Explaining structural irrationality

---

The aim of this chapter is to provide the outline of a general explanation of structural rationality in terms of non-structural requirements of rationality, i.e. rational requirements to respond to reasons. The general idea is that internal incoherence is not by itself forbidden by rationality, but only indicates that at least one of the attitudes involved is insufficiently supported by available reasons. I argue that a successful explanation of this kind would amount to a vindication of the normativity of rationality (9.1) and avoid the problems that befall the assumption of structural requirements of rationality (9.2). After discussing an important objection to the general idea (9.3), I provide a more detailed account of the explanation for different types of structural irrationality: weakness of will (9.4), weakness of belief (9.5), and *modus ponens* irrationality (9.6).

### 9.1 Explaining structural rationality: the general idea

Let me recall where we are. Chapters 2 to 6 provided arguments against both the structuralist and the mixed view, which motivated the development of a non-structuralist view that embraces the normativity of rationality and denies the existence of structural requirements of rationality. Chapter 7 defended a response conception of rationality that involves non-structural requirements of rationality to respond to available reasons. The arguments given in Chapter 8 entail that such requirements are necessarily normative in the strongest possible sense, by being supported by decisive reasons. The aim of the present chapter is to provide the outline of an explanation of the phenomenon of structural irrationality in terms of these (necessarily normative) response requirements. I start in this section by presenting the general idea of the explanation. This presentation will be sketchy and partly based on simplified claims; later sections provide a more detailed account.

To begin with, recall the non-structural requirements of rationality defended in Chapter 7 (the asterisk indicates that these are simplified versions):

TR\* *Theoretical rationality*: If A has sufficient evidence that p, then A is rationally required to believe p. If A lacks sufficient evidence that p, then A is rationally required not to believe p.

PR\* *Practical rationality*: If A has sufficient evidence that she ought to  $\phi$ , then A is rationally required to intend to  $\phi$ . If A has sufficient evidence that she ought not to  $\phi$ , then A is rationally required not to intend to  $\phi$ .

As argued in Chapter 7, these requirements are entailed by the more general requirement of rationality that requires conformity with decisive evidence-based reasons:

*Rationality as responding to reasons (RRR)*:

*Intention*: If A has decisive evidence-based reason (not) to intend to  $\phi$ , then A is rationally required (not) to intend to  $\phi$ .

*Belief*: If A has decisive evidence-based reason (not) to believe that p, then A is rationally required (not) to believe that p.

It follows that all the response requirements mentioned are necessarily supported by decisive evidence-based reason. As argued in Chapter 8, what we have decisive reason to do is what we have decisive evidence-based reason to do. It follows that all the response requirements mentioned are necessarily supported by decisive reasons, and thus normative in the strongest possible sense.

We have arrived at a set of rational requirements that are necessarily normative. That is a significant result, but it obviously falls short of explaining the normative force associated with *structural* irrationality. The third and last step of the account I am proposing aims to do exactly this, by providing a theory that explains structural irrationality in terms of non-structural requirements such as TR\* and PR\*. If successful, such a theory would amount to an explanation of structural irrationality in terms of a

set of rational requirements that are necessarily normative. It would thereby show that the phenomenon of structural irrationality poses no threat to the normativity of rationality.

The guiding idea of the kind of explanation I have in mind is well-expressed by Peter Geach, when he discusses the mistake involved in logical inconsistency:

A man who falls into inconsistency does not incur the further evil of a special sort of wrongness, logical wrongness; it is only that logic suffices to show that somewhere or other (logic does not say where) he is wrong in a non-logical way.<sup>1</sup>

Extending this idea to structural irrationality more generally, we might say that such irrationality does not involve a special kind of failure to be structurally rational, but only shows that somewhere or other one has failed to be rational in a non-structural way. What accounts for this irrationality, then, is not a structural requirement in favour of one attitude that is conditional on another attitude (as the narrow-scope account claims), nor a structural requirement in favour of a disjunctive response (as the wide-scope account claims), but a *disjunction* of non-structural requirements.<sup>2</sup> The relevant patterns of incoherent attitudes are rendered irrational by the fact that they guarantee the violation of at least one non-structural requirement.

Let me briefly illustrate this kind of explanation, using the example of weakness of will. The *akrates* believes that he has decisive reasons, or *ought* to do something, but does not intend to do it. This strikes us as irrational, as violating a standard of rationality. It is very tempting to say that the standard just *is* that rationality requires one to intend to act in accordance with one's normative beliefs. But drawing on Geach, we might just as well say that there is no such "special sort of wrongness" in having these incoherent attitudes. Rather, the incoherence only *indicates* that something *else* has gone wrong. The *akrates* cannot be fully rational, but the explanation may not be that he has

---

<sup>1</sup> Geach (1977, 166).

<sup>2</sup> By a disjunction of requirements, I mean a claim of the form "for all agents A, A is required to  $\phi$  or A is required to  $\psi$ " (not a claim of the form "all agents are required to  $\phi$  or all agents are required to  $\psi$ ").

violated the self-governance requirement, but that someone who is akratic has necessarily violated some other rational requirement.

What is this other requirement? Well, the *akrates* believes that he ought to  $\phi$ , but he does not intend to  $\phi$ . Now, either he has sufficient evidence for his belief or he does not. If he does not, he violates the non-structural requirement TR\*, for he believes something for which he lacks sufficient evidence. But if he *does* have sufficient evidence for his belief, then he violates the non-structural requirement PR\*, for although he has sufficient evidence that he ought to  $\phi$ , he does not intend to  $\phi$ . So in either case, the *akrates* violates a non-structural requirement of rationality.

This argument can be summarised as follows:

- (1) A has sufficient evidence that she ought to  $\phi$ , or A lacks sufficient evidence that she ought to  $\phi$  (*logical truth*).
- (2) If A lacks sufficient evidence that she ought to  $\phi$ , then A is rationally required not to believe that A ought to  $\phi$  (TR\*).
- (3) If A has sufficient evidence that she ought to  $\phi$ , then A is rationally required to intend to  $\phi$  (PR\*).
- (4) Therefore, A is rationally required not to believe that she ought to  $\phi$ , or A is rationally required to intend to  $\phi$  (*from 1, 2, and 3*).

(4), in turn, entails the structural irrationality claim (A)\*:

(A)\* *Weakness of will*: If A believes that she ought to  $\phi$ , and A does not intend to  $\phi$ , then A is irrational.

This is how the irrationality of weakness of will can be explained by non-structural requirements alone, without positing any structural requirement of rationality. Recall that (A)\* is not a claim about a rational requirement; as it stands, it merely states

sufficient conditions for irrationality.<sup>3</sup> It is natural to think that the conditions mentioned in (A)\* are sufficient for irrationality because there is a *structural* principle of rationality that requires us not to be in this condition. But as the argument shows, there is an alternative explanation available: the conditions mentioned are sufficient for irrationality not because of a structural requirement but because of a disjunction of non-structural requirements, as it occurs in (4). The key is Geach's insight that the mistake involved in incoherence is not the incoherence itself – the incoherence only shows that a mistake has been made. And this mistake, I suggest, lies in the violation of a non-structural requirement of rationality, which requires us to respond correctly to our evidence.

The general idea of this explanation has a great deal in common with Niko Kolodny's account of coherence requirements in "How Does Coherence Matter?" (2007a) and "The Myth of Practical Consistency" (2008a). The most important difference is that Kolodny's target is only what he calls "requirements of formal coherence as such".<sup>4</sup> Like T.M. Scanlon, Kolodny thinks that rationality consists in responding to one's beliefs about reasons; both therefore need an explanation for the fact that certain forms of incoherence among attitudes that do not seem to have normative content (such as ordinary beliefs and intentions) appear to be structurally irrational, as in the case of instrumental irrationality. Scanlon's explanation is that beliefs and intentions *have* normative content; simply put, he explains first-order structural requirements (such as the instrumental requirement) in terms of second-order structural requirements (such as the self-governance requirement).<sup>5</sup> Kolodny, in contrast, explains first-order structural requirements of rationality directly in terms of requirements of reasons. He argues, for example, that there is no rational requirement not to [believe p and

---

<sup>3</sup> As was argued in Chapter 1, claims that state sufficient (structural) conditions for irrationality do not themselves entail (structural) requirements of rationality, at least not in the sense of a code or standard, which is relevant here. Similarly, a disjunction of non-structural requirements, such as (4), which entails the same sufficient conditions for irrationality as some structural requirement, does not therefore entail this structural requirement.

<sup>4</sup> Kolodny (2007a, 229; 2008a, 366).

<sup>5</sup> See Scanlon (2007).

believe  $\neg p$ ], but only the general fact that one either lacks sufficient reason to believe  $p$ , or lacks sufficient reason to believe  $\neg p$ .<sup>6</sup>

What is crucial in our context is that Kolodny and Scanlon both *adhere* to the idea of structural requirements of rationality; they only object to structural requirements of a special kind. In particular, they are both committed to some version of the self-governance and the belief-guidance requirement.<sup>7</sup> Consequently, they have all the problems that such requirements pose, which I have discussed in detail above. In contrast, the account I am envisaging does away with structural requirements in general; it aims to explain all instances of structural irrationality in terms of non-structural requirements, not just instances of first-order structural irrationality.

A further important difference is that the account I am proposing explains structural irrationality in terms of non-structural requirements of *rationality*, which are shown to be necessarily normative, whereas Kolodny explains (first-order) structural requirements directly in terms of *reasons*. Thus, while my account can capture the widely shared intuition that instances of first-order structural irrationality are instances of irrationality, Kolodny's cannot do this. According to his account, a person who has beliefs with contradictory (non-normative) content is not irrational at all. To be sure, Kolodny maintains that such a person lacks sufficient reason for one of her beliefs, but as long as the person does not respond in any way that she actually *believes* to be defeated by reasons, Kolodny must deny that there is irrationality involved.

In sum, Kolodny explains first-order structural irrationality in terms of reasons, while I want to explain all kinds of structural irrationality in terms of non-structural requirements of rationality, which in turn are necessarily accompanied by reasons. These differences also explain why Kolodny's view is an *undermining* account of the normativity of rationality, while my view is a *vindicating* account. On Kolodny's view, there are on the one hand requirements of reasons that we can violate without committing any kind of irrationality, and on the other hand structural requirements of rationality that have no normative force. In my view, there are no structural requirements of

---

<sup>6</sup> See Kolodny (2007a).

<sup>7</sup> Note that Kolodny relies on the existence of such requirements not only in his argument against the normativity of rationality (2005), but also in his error theory about first-order requirements (2007a, 243; 2008a, 388).

rationality and all rational requirements are necessarily accompanied by reasons. I believe that my view has considerable advantages. Kolodny's adherence to non-normative structural requirements of rationality poses the need for the transparency account, which I have argued to be unsatisfactory in many respects. Kolodny ultimately has to deny that irrationality is criticisable, and that the normative pressure to revise one's attitudes in the light of recognised irrationality, or advise others to do so, is illusory. For the reasons given, my account is in a position to make sense of the commitments of these practices. Let us see whether it can be defended.

## 9.2 Structural and non-structural requirements, violation and satisfaction

Before turning to the more detailed explanation, I would like to take a closer look at how the *kind* of explanation of structural irrationality presented here compares with an explanation in terms of structural rationality, bracketing for a moment the normative question about rationality. This section considers some independent advantages of the non-structural explanation, while the next discusses a possible disadvantage. Consider the following schematic representation of the competing accounts of structural irrationality:

*Basic situation of structural irrationality:* If A is in the antecedent state (a), and A is not in the consequent state (c), then A is irrational.

*Narrow-scope account:* If A is in (a), then A is required to be in (c).

*Wide-scope account:* A is required [not to be in (a) or to be in (c)].

*Non-structural account:* A is required not to be in (a), or A is required to be in (c).

I have discussed a variety of problems for structural requirements of rationality in Chapter 6. An important upshot of this discussion was that both the narrow- and the wide-scope account of such requirements involve an overgeneralisation. Narrow-scope requirement claims state that in every possible world in which an agent is in the ante-

cedent state, he is rationally required to be in the consequent state. This means that in each of these worlds, an agent's refusal to adopt the consequent state constitutes the violation of a rational requirement.<sup>8</sup> As I have argued, however, this violation claim is too strong; it seems that there are possible worlds in which an agent's rationality is preserved despite his refusal to adopt the consequent state, provided that he gives up the antecedent state instead. For example, a weak-willed person might rationally refuse to intend to  $\phi$  if she gives up her belief that she ought to  $\phi$  on the basis of an assessment of her evidence. This is what the wide-scope account gets right. Yet the wide-scope account avoids this implausible violation claim only at the cost of maintaining an implausible satisfaction claim. The wide-scope account entails that in all possible worlds in which agents drop the antecedent state, these agents thereby come to satisfy the requirement, even though in many of these worlds, there are no rational processes available that could lead them to drop this attitude. For example, the wide-scope account entails that a weak-willed person always comes to satisfy a rational requirement by dropping her belief that she ought to  $\phi$ , no matter whether she has any rational basis for doing this.

Hence both narrow- and wide-scope requirements overgeneralise. Not in every world in which an agent is in the antecedent state and continues to refuse to adopt the consequent state is he violating a rational requirement (as the narrow-scope account entails); but nor does he satisfy a rational requirement in every world in which he gives up the antecedent attitude (which is entailed by the wide-scope account). Rather, it seems that in some worlds, agents are rationally required to adopt the consequent attitude, and in others they are required to drop the antecedent attitude. Note that this is exactly what the non-structural explanation entails. According to it, the irrationality that occurs if one is in an antecedent state but not in a consequent state is to be explained neither by a non-disjunctive requirement to be in the consequent state, which is conditional on one's being in the antecedent state, nor by a disjunctive requirement to [be in the consequent state *or* not-be in the antecedent state]. What explains the appearance of structural irrationality is instead a disjunction *of* require-

---

<sup>8</sup> As the discussion of Chapter 6.3 showed, this claim rests on the assumption defended in Chapter 3.4, namely that the requirement in question must ultimately be understood diachronically.

ments, such that necessarily, agents are rationally required to be in the consequent state *or* rationally required not to be in the antecedent state.

The non-structural explanation therefore avoids the problems that trouble both the narrow- and the wide-scope account quite independently of the normative question about rationality. According to the account I propose, which requirement the agent stands under depends on this agent's evidence. If an agent lacks evidence for her belief that she ought to  $\phi$ , there is no requirement to intend to  $\phi$ , and consequently no violation of a requirement if this agent drops her belief instead of forming an intention to  $\phi$ . In this way, the account avoids the implausible violation claims of the narrow-scope account. If an agent has sufficient evidence for her belief that she ought to  $\phi$ , then she is required to intend to  $\phi$ ; she cannot satisfy the demands of rationality by giving up this belief on no basis at all, as the wide-scope account implausibly entails. In this way, the non-structural account avoids the implausible satisfaction claims that wide-scope advocates are committed to. The account also avoids all the other problems of structural requirements of rationality discussed above. Since the requirements are conditional on evidence, they are accompanied by a rational basis on grounds of which agents can rationally satisfy them. In contrast to wide-scope requirements, they thus meet all versions of what I have called the "rational process constraint". Since the requirements are conditional on evidence *rather than* beliefs and intentions, the account neither licenses implausible bootstrapping, nor yields conflicting or undermining requirements, as was the case with the narrow-scope requirements. Since the requirements are not disjunctive, they avoid the peculiarities of wide-scope requirements discussed above; the requirements at issue are capable of guiding agents.

### 9.3 An objection: satisfaction in the wrong direction

Hence the non-structuralist account avoids the implausible implications of both the narrow- and the wide-scope account, in particular with respect to what responses count as a violation or a satisfaction of a rational requirement. However, as especially Niko Kolodny has emphasised, it might also be argued that non-structural requirements in turn fail to explain some natural intuitions about satisfaction that both

narrow- and wide-scope requirements account for.<sup>9</sup> Suppose that Tim believes that he ought to stay on a certain diet, but he does not intend to do so. Tim actually lacks sufficient evidence for this belief; in fact, his experience indicates that the diet in question is counterproductive. Yet he sincerely believes that he ought to continue with it, and he experiences his lack of intention as a blemish. Reflecting on this, he finally decides to stay on the diet, basing his intention on his sincere (but unjustified) belief that he ought to do so. Even though it is clear that Tim is making a mistake in keeping to his belief, it is also plausible to say that there is something that Tim gets right. Structuralist accounts can explain this intuition, for both the narrow- and the wide-scope view entail that in deciding as he does, Tim satisfies a rational requirement. According to the anti-structuralist view, however, the only rational requirement that applies to Tim is the one that requires him to give up his belief. The view therefore lacks the resources to explain that Tim gets at least something right when he forms his intention.

Tim's case suggests that one can at least sometimes escape structural irrationality in the "wrong direction" – from the perspective of non-structural requirements – and still count as satisfying a rational requirement. How could this be so if there were no structural requirements of rationality? I shall propound three alternative explanations that are compatible with the anti-structuralist view, and then argue that they offer a better account of the phenomena in question than structural requirements of rationality.

The first point to note is that an agent who escapes a structurally irrational state in the wrong direction might still count as manifesting a generally valuable disposition. Rational capacities are not exhausted by the ability to reflectively guide our responses, but consist to a considerable part in dispositions for certain responses. The disposition to believe and intend in accordance with one's assessment of reasons is quite plausibly part of the bundle of dispositions that constitute the rational capacity (as is, plausibly, the disposition to intend the means to one's intended ends). These dispositions are part of the general capacity to respond to reasons, and they ensure the

---

<sup>9</sup> See Kolodny (2007a, 236–237; 2008a, 387–388).

effectiveness of our agency in general.<sup>10</sup> They are therefore highly valuable. All of this is compatible with denying that we are rationally required (or have normative reasons) to respond, in each particular case, in the ways in which beings endowed with the rational capacity are generally disposed to respond.<sup>11</sup> But it provides a good explanation for the intuition that escaping a structurally irrational state in the wrong direction can seem rational or even praiseworthy. The situation is similar to the one in which a morally disposed person does the wrong thing, but acts on good motives. The responses in question appear in a positive light not because they satisfy a requirement, but because they manifest a valuable disposition.

A second explanation is suggested by Kolodny, who faces basically the same challenge with respect to his explanation of first-order structural irrationality in terms of reasons.<sup>12</sup> A person who is in an incoherent state typically has “second-order” evidence that one of his attitude-states is not sufficiently supported by reasons (or forbidden by a non-structural requirement of rationality). *Ceteris paribus*, he is therefore required to review his evidence and reassess the reasons that he takes himself to have. Such a process is not infallible and might lead to the wrong result. But someone who is led astray by his own reassessment, and therefore escapes incoherence in the wrong direction, might nevertheless count as satisfying a second-order requirement to review and reassess a particular group of attitudes.<sup>13</sup> The situation might be compared to one in which a giver notices that the presentee is unhappy with a present, and responds to this fact by replacing the present with another one which is even less suitable.<sup>14</sup> Such a person gets something right because she correctly responds to a second-order reason to reassess her choice, even though by doing so she continues to fail to respond to her first-order reasons for buying a suitable present.

Last but not least, the intuition that escaping incoherence in the wrong direction can amount to satisfying a requirement might also be grounded in the fact

---

<sup>10</sup> See e.g. Raz (2005a, 17–18) and Bratman (2009a, 417–418).

<sup>11</sup> Compare Raz’s (2005b, 10–12) response to Schroeder (2005b, 4–6).

<sup>12</sup> Cf. Kolodny (2007a, 242–244; 2008a, 387–390).

<sup>13</sup> A person could count as satisfying such a requirement independently of whether she would conclude her reassessment and come to have coherent attitude-states; the point here is just that a person’s coming to have coherent attitude-states is often a sign that she responds to a requirement to reassess.

<sup>14</sup> Thanks to Jan Gertken for suggesting this analogy to me.

that it is *correct* to *base* certain attitudes on one another in particular ways. It is correct, for example, to base a belief in  $q$  on a belief in  $p$  and a belief in  $p \rightarrow q$ . And it is correct to base an intention to  $\phi$  on a belief that one ought to  $\phi$ . In a broad sense of ‘reasoning’, basing attitudes on one another in these ways amounts to correct reasoning. And an agent who escapes an incoherent state in the wrong direction might nevertheless be judged to get something right because he accomplishes certain mental procedures that can count as correct reasoning.

It might be wondered whether this last response does not reinstate the problems of structural requirements of rationality. If basing attitude (b) on attitude (a) is correct, does it not follow that it is also rationally permitted? And if basing attitude (b) on attitude (a) is rationally permitted, does it not follow that *having* attitude (b) is rationally permitted when it is based on (a)?<sup>15</sup> Does this not mean that one can *make* an attitude (b) rationally permissible simply by adopting an attitude on which one can rationally base (b)? Further, if rationality is normative, do rational permissions not entail sufficient reasons? And would we not have to conclude that if it is correct to base (b) on (a), then we can bootstrap sufficient reasons for (b) into existence by merely adopting (a)?

This argument should be resisted. If basing an attitude (b) on (a) is correct or rationally permitted, it does not follow that having attitude (b) is also correct or rationally permitted, not even when it is based on (a).<sup>16</sup> If having attitude (a) is not rationally permitted, then having attitude (b) may also not be rationally permitted, even if the basing itself is permitted. The mistake of someone who bases an attitude on a rationally impermissible attitude does not lie in the basing itself; this basing might be rationally permissible or correct even if both attitudes are rationally impermissible. As far as I can see, the acceptance of rational basing permissions does not pose any of the problems that befall structural requirements of rationality and is compatible with the normativity of rationality.<sup>17</sup> By the same token, the anti-structuralist view I am

---

<sup>15</sup> This seems to be presupposed by an argument of Shpall (2013, 726).

<sup>16</sup> See also Broome (2013, 190).

<sup>17</sup> I take basing permissions to be claims of the form “for all agents  $A$  and attitude-state types (a) and (b),  $A$  is permitted to base  $b$  on  $a$ ”. Since such claims are unconditional, ambiguity between wide- and narrow scope readings does not occur.

defending here can refer to such basing permissions in order to explain intuitions that seem to support the existence of structural requirements of rationality.

The suggestion, then, is this. Our tendency to think that Tim satisfies a rational requirement even though he escapes his incoherent state in the wrong direction from the standpoint of non-structural requirements stems from three sources. First, by forming the intention to  $\phi$  against his evidence he might nevertheless manifest a generally valuable disposition to intend in accordance with his assessments of reasons. Second, his intention indicates that he satisfies the requirement to reflect on and reassess his reasons in case his evidence suggests that at least one member of a particular group of attitudes is insufficiently supported by reasons. Third, even though his intention is rationally impermissible, it might still be based in a rationally permissible way that can qualify as correct reasoning.

We can test this hypothesis by considering a case in which an agent – call him Tom – escapes an incoherent state in the wrong direction while at the same time failing to meet any of the aforementioned conditions. If the hypothesis is correct, our inclination to think of Tom as getting something right should be significantly lower than our inclination to think of Tim as getting something right, and eventually diminish after reflection. If, in contrast, the inclination to think of Tim as getting something right is to be explained by structural requirements of rationality, then we should have no hesitation to think of Tom as getting something right as well. So let us have a closer look at Tom's case. Just like Tim, Tom believes that he ought to stay on a certain diet, but he does not intend to do so. And just like Tim, Tom comes to intend to stay on the diet, even though his experience suggests that the diet is counterproductive. Yet while Tim was concerned about the inner tension in his attitudes, Tom does not care about it at all. His forming the intention to stay on the diet therefore cannot be interpreted as an expression of his concern about his second-order evidence that one of his attitudes cannot be rational or sufficiently supported by reasons. Moreover, Tom's intention is not based on his belief that he ought to stay on the diet. Perhaps he comes to have the intention by mere accident, or he comes to have it because of a weird disposition to form an intention to diet whenever he sees delicate food. As a consequence, there is no reason to think that the formation of his intention is an instantiation of a generally

valuable disposition; from what we know about Tom, his disposition to intend in accordance with his assessment of reasons is poorly developed anyway.

Is there any pretheoretical reason for thinking that Tom satisfies a requirement, that he gets something right or is in any sense praiseworthy for forming his intention? Speaking for myself, I do not share this intuition. In any case, the judgement that Tom is getting something right is disputable; it cannot be regarded as an independent criterion for a theory of rationality that it renders this claim true.

#### 9.4 Explaining weakness of will

In the last three sections, I have presented and defended the general idea of an explanation of structural irrationality in terms of non-structural requirements of rationality. I have argued that the explanation would amount to a vindication of the normativity of rationality, and that it has independent advantages over structuralist accounts. I have also argued that our tendency to think of agents that escape structural irrationality in the wrong direction as getting something right can be explained without adherence to structural requirements of rationality, and is indeed more comprehensible if it is explained in the ways suggested. The remainder of this chapter is concerned with a more accurate and detailed account and defence of this explanation. This section starts with the explanation of weakness of will.

First, as I have argued in Chapter 7, TR\* and PR\* need to be qualified as follows:

TR *Theoretical rationality*: If A has sufficient evidence for p, **and A attends to p**, then A is rationally required to believe p. If A lacks sufficient evidence for p, then A is rationally required not to believe p.

PR *Practical rationality*: If A has sufficient evidence that she ought to  $\phi$ , **and A has sufficient evidence that condition C obtains**, then A is rationally required to intend to  $\phi$ . If A has sufficient evidence that she ought not to  $\phi$ , then A is rationally required not to intend to  $\phi$ .

To recall, we can substitute for “condition C” the condition that A will not perform the relevant action unless she now intends to perform it, or some other sufficient condition that connects reasons for action with reasons for intention. We have also seen that (A)\* and (B)\* need to be similarly qualified:

- (A) *Weakness of will*: Necessarily, if A believes that she ought to  $\phi$ , **and A believes that condition C obtains**, and A does not intend to  $\phi$ , then A is irrational.
- (B) *Instrumental irrationality*: Necessarily, if A intends to  $\phi$ , and A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and **A believes that condition C obtains**, and A does not intend to  $\psi$ , then A is irrational.

Given these qualifications, the simplified argument given above, which purported to show how weakness of will can be explained in terms of non-structural requirements, can be revised as follows:

- (1) Either [A has sufficient evidence that she ought to  $\phi$  and A has sufficient evidence that condition C obtains], or it is not the case that [A has sufficient evidence that she ought to  $\phi$  and A has sufficient evidence that condition C obtains] (*logical truth*).
- (2.1) If it is not the case that [A has sufficient evidence that she ought to  $\phi$  and A has sufficient evidence that condition C obtains], then it is not the case that A has sufficient evidence that she ought to  $\phi$ , or it is not the case that A has sufficient evidence that condition C obtains (*logical truth*).
- (2.2) If A lacks sufficient evidence that she ought to  $\phi$ , then A is rationally required not to believe that she ought to  $\phi$  (*TR*).
- (2.3) If A lacks sufficient evidence that condition C obtains, then A is rationally required not to believe that condition C obtains (*TR*).
- (2) If it is not the case that [A has sufficient evidence that she ought to  $\phi$  and A has sufficient evidence that condition C obtains], then either A is rationally required

- not to believe that she ought to  $\phi$ , or A is rationally required not to believe that condition C obtains (*from 2.1-2.3*).
- (3) If [A has sufficient evidence that she ought to  $\phi$  and A has sufficient evidence that condition C obtains], then A is rationally required to intend to  $\phi$  (*PR*).
- (4) Therefore, either A is rationally required not to believe that she ought to  $\phi$ , or A is rationally required not to believe that condition C obtains, or A is rationally required to intend to  $\phi$  (*from 1, 2, and 3*).<sup>18</sup>

(4), in turn, entails the revised structural irrationality claim (A):

- (A) *Weakness of will*: If A believes that she ought to  $\phi$ , and A believes that condition C obtains, and A does not intend to  $\phi$ , then A is irrational.

## 9.5 Explaining weakness of belief

Recall weakness of belief, the theoretical analogue to weakness of will:

- (C) *Weakness of belief*: Necessarily, if A believes that she has sufficient evidence that p, and A does not believe that p, then A is irrational.

This claim might be explained in terms of non-structural requirements of rationality by way of the following argument:

---

<sup>18</sup> Since (1) and (2.1) are logical truths, they need not be stated as premises for this argument to be valid. For reasons of transparency, I state logical truths as premises nevertheless (here as well as in the arguments that follow in the next sections). The formal structure of the argument is this:

- (1)  $(P \& Q) \vee \neg(P \& Q)$ , *logical truth*  
 (2.1)  $\neg(P \& Q) \rightarrow (\neg P \vee \neg Q)$ , *logical truth*  
 (2.2)  $\neg P \rightarrow R$ , *TR*  
 (2.3)  $\neg Q \rightarrow S$ , *TR*  
 (2)  $\neg(P \& Q) \rightarrow (R \vee S)$ , *from 2.1-2.3*  
 (3)  $(P \& Q) \rightarrow T$ , *PR*  
 (4)  $R \vee S \vee T$ , *from 1, 2, and 3*

- (1) Either A has sufficient evidence that she has sufficient evidence that p, or A lacks sufficient evidence that she has sufficient evidence that p (*logical truth*).
  - (2) If A lacks sufficient evidence that she has sufficient evidence that p, then A is rationally required not to believe that she has sufficient evidence that p (*TR*).
  - (3.1) If A has sufficient evidence that she has sufficient evidence that p, then A has sufficient evidence that p (*premise*).
  - (3.2) If A has sufficient evidence that p, and A attends to p, then A is rationally required to believe that p (*TR*).
  - (3.3) Therefore, if A has sufficient evidence that she has sufficient evidence that p, and A attends to p, then A is rationally required to believe that p (*from 3.1 and 3.2*).
  - (3.4) If A believes that she has sufficient evidence that p, then A attends to p in the relevant sense (*premise*).
  - (3.5) Therefore, if A has sufficient evidence that she has sufficient evidence that p, and A believes that she has sufficient evidence that p, then A is rationally required to believe that p (*from 3.3 and 3.4*).
  - (3) If A has sufficient evidence that she has sufficient evidence that p, then either A is rationally required to believe that p or A does not believe that she has sufficient evidence that p (*equivalent to 3.5*).
  - (4) Therefore, either A is rationally required not to believe that she has sufficient evidence that p, or A is rationally required to believe that p or A does not believe that she has sufficient evidence that p (*from 1, 2, and 3*).
  - (5) Therefore, if A believes that she has sufficient evidence that p, then either A is rationally required not to believe that she has sufficient evidence that p, or A is rationally required to believe that p (*equivalent to 4*).
- (5), in turn, entails the structural irrationality claim (C).<sup>19</sup> The argument shows that (C) might be explained in terms of non-structural requirements of rationality alone.

---

<sup>19</sup> As above, logical truths are stated as premises for reasons of transparency. The formal structure of the argument is this:

Besides (TR), it is based on two additional premises, (3.1) and (3.4), which I shall briefly discuss in reverse order.

(3.4) claims that the attention condition in (TR) is satisfied when A believes that she has sufficient evidence for p. Recall that the attention condition was introduced in response to the problem of justificational fecundity: since our evidence supports an infinite number of propositions, we are required to believe only those supported propositions that we attend to. (3.4) claims that this condition is satisfied if we believe of a particular proposition that we have sufficient evidence for it. This is intuitively plausible. If you have sufficient evidence for p, and you believe that you have sufficient evidence for p, you cannot plausibly excuse yourself for not believing p by pointing out that you did not attend to p.

Suppose, to the contrary, that you could believe that you have sufficient evidence for p without attending to p in the sense relevant for TR to apply. If this were true, we would have good reason to think that claim (C) has to be qualified by way of an attention condition. If one could believe that one has sufficient evidence for p without paying the relevant kind of attention to p, then it would seem that it can be rationally permitted to refrain from believing p while believing that one has sufficient evidence for p, namely if one does not attend to p. However, a qualified version of (C), which includes the condition that A attends to p, is entailed by an analogous argument that does not require premise (3.4):

(3)\* If A has sufficient evidence that she has sufficient evidence that p, then A is rationally required to believe that p or A does not attend to p (*equivalent to 3.3*).

- 
- (1)  $P \vee \neg P$ , *logical truth*
  - (2)  $\neg P \rightarrow Q$ , *TR*
  - (3.1)  $P \rightarrow R$ , *premise*
  - (3.2)  $(R \& S) \rightarrow T$ , *TR*
  - (3.3)  $(P \& S) \rightarrow T$ , *from 3.1-3.2*
  - (3.4)  $U \rightarrow S$ , *premise*
  - (3.5)  $(P \& U) \rightarrow T$ , *from 3.3-3.4*
  - (3)  $P \rightarrow (T \vee \neg U)$ , *equivalent to 3.5*
  - (4)  $Q \vee T \vee \neg U$ , *from 1, 2, and 3*
  - (5)  $U \rightarrow (Q \vee T)$ , *equivalent to 4*

- (4)\* Therefore, A is rationally required not to believe that she has sufficient evidence that p, or A is rationally required to believe that p, or A does not attend to p (from 1, 2, and 3\*).
- (5)\* Therefore, if A believes that she has sufficient evidence that p, and A does not believe that p, and A attends to p, then A is irrational (from 4\*).

The lesson is that (3.4) is either true, or not needed in order to explain the correct version of (C).

Premise (3.1) states that one has sufficient evidence for any proposition p if one has sufficient evidence for the proposition that one has sufficient evidence for p. This claim, too, is intuitively plausible. If one's evidence makes it rational to believe that one has sufficient evidence for p, then this very evidence should also make it rational to believe p itself. The negation of (3.1) also has awkward consequences. Suppose you could have sufficient evidence that you have sufficient evidence for p, while having insufficient evidence for p itself. Then, if you attend to the question whether you have sufficient evidence for p, you would be rationally required to believe that you have sufficient evidence for p. But you would also be rationally required not to believe p. Both of these claims follow from (TR), which is widely accepted and has been defended in Chapter 7. So if (3.1) were false, you could be rationally required to believe that you have sufficient evidence for p, and at the same time rationally required not to believe p. But (C) tells us that you would be irrational in having this combination of belief-states. It follows that your evidence situation might be such that no matter what you do, your belief-states are irrational. This is an unacceptable conclusion; you cannot be forced into irrationality through no fault of your own.<sup>20</sup> Hence, either (3.1) is true, or (TR) or (C) must be false. In fact, all three claims are plausible and should be maintained as long as there is no good reason to give them up.<sup>21</sup> Since (3.1) is intuitively plausible,

<sup>20</sup> Note that even Kolodny, who thinks that rationality can require incompatible responses, accepts that this can be so only because "you have painted yourself into a corner", and "if you have painted yourself into a corner, then there was a way not to have done so" (Kolodny 2007b, 384).

<sup>21</sup> There is a worry with (3.1) that is worth mentioning, however. Suppose that there is a particular number x between 0.5 and 1 such that A has sufficient evidence for p if, and only if, the probability of p, conditional on A's evidence E, is higher than x. On this assumption, (3.1) seems to suggest that the probability of a proposition p must be greater than x if the probability of the proposition that the probability of p is greater

and is entailed by the independently plausible principles (TR) and (C), we have good reasons to take it to be true. Consequently, the argument stands. (C) can be explained without adherence to structural requirements of rationality.

## 9.6 Explaining *modus ponens* irrationality

After discussing second-order cases of structural irrationality (weakness of will and belief), I shall now continue with the explanation of first-order cases of structural irrationality, starting with:

(D)\* *Modus ponens irrationality*: Necessarily, if A believes that  $p$ , and A believes that  $p \rightarrow q$ , and A does not believe that  $q$ , then A is irrational.

As I already pointed out when I first introduced this claim in Chapter 1, (D)\* is an overgeneralisation. We are often rationally permitted not to clutter our minds with believing all the logical consequences of what we believe, and it would be surprising if this could never be the case with consequences that follow via *modus ponens*. Some authors suggest that we need to introduce into (D)\* the further condition that A *cares* whether  $q$  is the case.<sup>22</sup> Taking into account parts of the discussion of the uniqueness and the fecundity theses, however, I would suggest that the correct condition is rather that A *attends* to  $q$ . First, if an agent attends to a proposition, it seems to me that

---

than  $x$ , is greater than  $x$ , i.e.:  $P(P[p|E] > x | E) > x \rightarrow P(p|E) > x$ . However, this principle is not generally true if  $x < 1$ . So if we do not generally limit the propositions that we are justified in believing to those that have evidential probability 1 (which would be very restrictive), we have to ask how (3.1) can be true if the analogous probability principle is not. One general possibility is that the connection between sufficient evidence and probability is not as close as this argument presumes. Another possibility is that the probability standards for justified belief depend on the content of the proposition. Perhaps, in order to be justified to believe the proposition that one has sufficient evidence for  $p$ , this proposition needs to be more likely than the proposition  $p$  would need to be in order for it to be the case that one is justified to believe it (this might be due to the higher-order nature or generally the higher complexity of the former proposition). A further possibility is that due to conceptual connections between evidence for  $p$  and evidence for evidence for  $p$ , one cannot fail to have a very high degree of evidential probability for a truth about what one has sufficient evidence to believe. I do not have a decided opinion on that matter, but given these possibilities, it seems to me that the falsity of this probability principle should not be taken to call (3.1) into question. I owe thanks to Franz Huber and Thomas Kroedel for helpful discussion of the issue raised in this note.

<sup>22</sup> See e.g. Broome (2013, 157–58). In earlier work, Broome (2009, 64) worked with the alternative condition that it matters to A whether  $q$ . See also Harman's interest condition (1986, 55) and Nozick's expected utility rule (1993, 88) quoted in Chapter 7.7.

whether or not she cares about its truth is irrelevant for the question of whether belief in that proposition is rational.<sup>23</sup> So caring is not a necessary condition for *modus ponens* irrationality. Together with the conditions mentioned in (D)\*, it might be taken to be a sufficient condition, but the explanation for this is likely to be that the relevant sense of ‘caring’ entails the relevant sense of ‘attending’. If we understand caring such that one can care whether *q* in a subconscious way that does not involve any kind of attention, its relevance again seems to me questionable. Hence I would suggest the following revision of (D)\*:

(D)\*\* *Modus ponens irrationality*: Necessarily, if A believes that *p*, and A believes that  $p \rightarrow q$ , **and A attends to *q***, and A does not believe that *q*, then A is irrational.

This claim can be explained in terms of non-structural requirements of rationality on the assumption of the following principle:

*Evidence transmission*\*: If A has sufficient evidence that *p*, and A has sufficient evidence that  $p \rightarrow q$ , then A has sufficient evidence that *q*.

The following argument shows how:

- (1) Either [A has sufficient evidence that *p* and A has sufficient evidence that  $p \rightarrow q$ ], or it is not the case that [A has sufficient evidence that *p* and A has sufficient evidence that  $p \rightarrow q$ ] (*logical truth*).
- (2.1) If it is not the case that [A has sufficient evidence that *p* and A has sufficient evidence that  $p \rightarrow q$ ], then it is not the case that A has sufficient evidence that *p* or it is not the case that A has sufficient evidence that  $p \rightarrow q$  (*logical truth*).

---

<sup>23</sup> See also Harman (1986, 15): “Once one is explicitly considering whether or not to accept a conclusion, one cannot decide not to on ... grounds [of clutter avoidance].”

- (2.2) If A lacks sufficient evidence that p, then A is rationally required not to believe that p (*TR*).
- (2.3) If A lacks sufficient evidence that  $p \rightarrow q$ , then A is rationally required not to believe that  $p \rightarrow q$  (*TR*).
- (2) If it is not the case that [A has sufficient evidence that p and A has sufficient evidence that  $p \rightarrow q$ ], then either A is rationally required not to believe that p or A is rationally required not to believe that  $p \rightarrow q$  (*from 2.1-2.3*).
- (3.1) If A has sufficient evidence that p, and A has sufficient evidence that  $p \rightarrow q$ , then A has sufficient evidence that q (*evidence transmission\**).
- (3.2) If A has sufficient evidence that q, and A attends to q, then A is rationally required to believe that q (*TR*).
- (3.3) If A has sufficient evidence that q, then either A is rationally required to believe that q or A does not attend to q (*equivalent to 3.2*).
- (3) If A has sufficient evidence that p, and A has sufficient evidence that  $p \rightarrow q$ , then either A is rationally required to believe that q or A does not attend to q (*from 3.1 and 3.3*).
- (4) Therefore, either A is rationally required not to believe that p, or A is rationally required not to believe that  $p \rightarrow q$ , or A is rationally required to believe that q, or A does not attend to q (*from 1, 2, and 3*).<sup>24</sup>

As before, (4) entails the structural irrationality claim (D)\*\*.

---

<sup>24</sup> The formal structure of the argument is this:

- (1)  $(P \& Q) \vee \neg(P \& Q)$ , *logical truth*  
 (2.1)  $\neg(P \& Q) \rightarrow (\neg P \vee \neg Q)$ , *logical truth*  
 (2.2)  $\neg P \rightarrow R$ , *TR*  
 (2.3)  $\neg Q \rightarrow S$ , *TR*  
 (2)  $\neg(P \& Q) \rightarrow (R \vee S)$ , *from 2.1-2.3*  
 (3.1)  $(P \& Q) \rightarrow T$ , *evidence transmission\**  
 (3.2)  $(T \& U) \rightarrow V$ , *TR*  
 (3.3)  $T \rightarrow (V \vee \neg U)$ , *equivalent to 3.2*  
 (3)  $(P \& Q) \rightarrow (V \vee \neg U)$ , *from 3.1 and 3.3*  
 (4)  $R \vee S \vee V \vee \neg U$ , *from 1, 2, and 3*

*Evidence transmission and probability*

The problem with this explanation is that there are good reasons for doubting *evidence transmission\**. While it is plausible that evidence for  $p$  and evidence for  $p \rightarrow q$  generally transmits to evidence for  $q$ , it is questionable that the *strength* of this evidence is necessarily preserved. Some natural assumptions about the relation between evidence and probability suggest that this is not the case. First, it seems clear that we can have evidence for  $p$  that is sufficient for rational or justified belief, even if the evidential probability for  $p$  is less than 1. Otherwise, it would be difficult to see how we could ever be justified in believing a contingent proposition. Next, it is also clear that the evidence for  $p$  can be insufficient for rational or justified belief, even though the evidential probability is higher than 0. Now suppose that  $A$  has barely sufficient evidence for  $p$  and barely sufficient evidence for  $p \rightarrow q$ , and no independent evidence for  $q$ . In such a case, all the evidential probability for  $q$  comes from the evidential probability for  $p$  and  $p \rightarrow q$ . Yet the probability of a conjunction may be lower than the probability of each of the conjuncts. The evidential probability for  $q$  may therefore be lower than the evidential probability for  $p$  and the evidential probability for  $p \rightarrow q$ . Given that this is possible, it is difficult to see why there could not be cases in which the evidence probabilises each  $p$  and  $p \rightarrow q$  to a sufficient degree, while it does not probabilise  $q$  to a sufficient degree, which would falsify *evidence transmission\**.

Similar conclusions have been drawn from two paradoxes famously discussed in the philosophical literature: the lottery paradox and the preface paradox.<sup>25</sup> Both of these paradoxes can be described as cases in which we have sufficient evidence for each of a number of propositions that are jointly inconsistent. In the lottery case, we have sufficient evidence for each claim that a particular lottery ticket will not win, but we also have sufficient evidence that one lottery ticket will win. In the preface case, an author has sufficient evidence for each of the claims that he maintains in his book, but he also has sufficient evidence that not all of them are true. Taken at face value, these examples seem to show that we can have sufficient evidence for propositions not only

---

<sup>25</sup> See Kyburg (1961, 197) for the former, and Makinson (1965) for the latter paradox.

while *lacking* sufficient evidence for their logical implications, but also while *having* sufficient evidence that the logical implications are false.

There are attempts to solve these paradoxes without giving up the claim that evidence is closed under logical implication.<sup>26</sup> If successful, such theories might amount to a defence of *evidence transmission\**, and thus rescue the non-structural explanation of modus ponens irrationality that I have given above. Still, there are reasons to be sceptical about *evidence transmission\**, and it is worth exploring how an explanation for modus ponens irrationality could do without it.

The first thing I would like to note is that the considerations about sufficient evidence and probability brought forward not only cast doubt on *evidence transmission\**, they likewise call into question whether the conditions specified in (D)\*\* really are sufficient for irrationality. If it is possible to have sufficient evidence for each  $p$  and  $p \rightarrow q$ , while lacking sufficient evidence for  $q$ , why should it be necessarily irrational to refrain from believing  $q$  when one both believes  $p$  and believes  $p \rightarrow q$ ? Insofar as *evidence transmission\** is an overgeneralisation, (D)\*\* seems to be an overgeneralisation as well. The original paradoxes themselves suggest this conclusion. On the face of it, an author who believes each of the claims of his books for good reasons does not seem irrational when he also believes that not all of these claims are true.

Now, there still seem to be many possible cases of *modus ponens* irrationality. First, if *evidence transmission\** fails due to probability-related considerations, then this is because it is a *multiple* premise closure principle. Contrast:

*Single premise evidence transmission:* If  $A$  has sufficient evidence that  $p$ , and  $q$  follows logically from  $p$ , then  $A$  has sufficient evidence that  $q$ .

Given that logical truths have probability 1, the strength of the evidential probability for  $p$  completely transmits to the evidential probability for  $q$ . It follows that sufficient evidence for the conjunction ' $p$  and  $p \rightarrow q$ ' fully transmits to sufficient evidence for  $q$ . By

---

<sup>26</sup> See e.g. Kroedel (2012), who argues that high evidential probability gives rise only to epistemic permissions, and Littlejohn (2012), who argues that high evidential probability does not justify belief at all. I raise some objections against Kroedel's solution to the lottery paradox in my paper "Can the lottery paradox be solved by identifying epistemic justification with epistemic permissibility?" (unpublished).

the same token, probability considerations provide no objection to the claim that a person who believes ‘ $p$  and  $p \rightarrow q$ ’, attends to  $q$ , but does not believe that  $q$ , is necessarily irrational.<sup>27</sup> An argument analogous to the one above can explain this claim in terms of non-structural requirements of rationality, on the assumption of *single premise evidence transmission*.

Second, there will also be cases in which  $A$  is irrational in having *modus ponens* inconsistent beliefs even if she does not believe the conjunction ‘ $p$  and  $p \rightarrow q$ ’. These will be cases in which  $A$ , in addition to believing  $p$  and believing  $p \rightarrow q$ , has a sufficiently high *credence* in each of these propositions. Consider the following theorem of probability logics:

*Theorem:* If  $P(p) > 1 - \epsilon$ , and  $P(p \rightarrow q) > 1 - \epsilon$ , then  $P(q) > 1 - 2\epsilon$ .<sup>28</sup>

As Branden Fitelson points out, while this theorem allows that *some* degree of probability can be lost in *modus ponens* inferences starting from uncertain premises, “it also tells us that, when the premises are *highly* probable, the *amount* of probability that can be lost ... is rather small”.<sup>29</sup> Now let us suppose, for some appropriately small  $\epsilon$ , that if  $A$ ’s evidential probability is higher than  $1 - 2\epsilon$ , then  $A$  has sufficient evidence that  $p$ . On this assumption, the following multi-premise evidence transmission principle escapes the probability-related objections raised above:

*Evidence transmission:* If  $A$  has sufficient evidence that  $p$ , and  $A$  has sufficient evidence that  $p \rightarrow q$ , and  $A$ ’s evidential probability  $P(p) > 1 - \epsilon$ , and  $A$ ’s evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ , then  $A$  has sufficient evidence that  $q$ .

By the same token, probability-related considerations provide no objection to:

<sup>27</sup> Cf. Nozick (1993, 92): “I suggest that for belief legitimately to be transferred from premisses to conclusion in a deductive inference not only must each premiss be believed but also the conjunction of the premisses must be believed.”

<sup>28</sup> See Hailperin (1996, 205). ‘ $\epsilon$ ’ stands for an arbitrary constant, such that  $1 > \epsilon > 0$ .

<sup>29</sup> Fitelson (2010, 460).

- (D) *Modus ponens irrationality*: Necessarily, if A believes that  $p$ , and A believes that  $p \rightarrow q$ , and A's credence  $C(p) > 1 - \epsilon$ , and A's credence  $C(p \rightarrow q) > 1 - \epsilon$ , and A attends to  $q$ , and A does not believe that  $q$ , then A is irrational.

This claim can also be explained in terms of non-structural requirements of rationality, but for this we need a requirement that connects evidential probability with credence. Against the background of an evidentialist picture, the following principle should not appear controversial:

- CR *Credence rationality*: For every proposition  $p$ , every degree of credence  $C$ , and every degree of evidential probability  $P$ , A is rationally required not to have  $C(p) > P(p)$ .<sup>30</sup>

On the assumption of CR and the revised transmission principle, (D) can be explained as follows:

- (1) Either [A has sufficient evidence that  $p$ , and A has sufficient evidence that  $p \rightarrow q$ , and A's evidential probability  $P(p) > 1 - \epsilon$ , and A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ ], or it is not the case that [A has sufficient evidence that  $p$ , and A has sufficient evidence that  $p \rightarrow q$ , and A's evidential probability  $P(p) > 1 - \epsilon$ , and A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ ] (*logical truth*).
- (2.1) If it is not the case that [A has sufficient evidence that  $p$ , and A has sufficient evidence that  $p \rightarrow q$ , and A's evidential probability  $P(p) > 1 - \epsilon$ , and A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ ], then it is not the case that A has sufficient evidence that  $p$ , or it is not the case that A has sufficient evidence that  $p \rightarrow q$ , or it is not the case that A's evidential probability  $P(p) > 1 - \epsilon$ , or it is not the case that A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$  (*logical truth*).

---

<sup>30</sup> Note that since I am interested in a vindication of the normativity of rationality, I am also committed to the claim that CR is necessarily normative, i.e.: For every proposition  $p$ , every degree of credence  $C$ , and every degree of evidential probability  $P$ , A has *decisive reason* not to have a credence  $C(p) > P(p)$ .

- (2.2) If A lacks sufficient evidence that  $p$ , then A is rationally required not to believe that  $p$  (*TR*).
- (2.3) If A lacks sufficient evidence that  $p \rightarrow q$ , then A is rationally required not to believe that  $p \rightarrow q$  (*TR*).
- (2.4) If it is not the case that A's evidential probability  $P(p) > 1 - \epsilon$ , then A is rationally required not to have a credence  $C(p) > 1 - \epsilon$  (*CR*).
- (2.5) If it is not the case that A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ , then A is rationally required not to have a credence  $C(p \rightarrow q) > 1 - \epsilon$  (*CR*).
- (2) If it is not the case that [A has sufficient evidence that  $p$ , and A has sufficient evidence that  $p \rightarrow q$ , and A's evidential probability  $P(p) > 1 - \epsilon$ , and A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ ], then either A is rationally required not to believe that  $p$ , or A is rationally required not to believe that  $p \rightarrow q$ , or A is rationally required not to have a credence  $C(p) > 1 - \epsilon$ , or A is rationally required not to have a credence  $C(p \rightarrow q) > 1 - \epsilon$  (*from 2.1-2.5*).
- (3.1) If A has sufficient evidence that  $p$ , and A has sufficient evidence that  $p \rightarrow q$ , and A's evidential probability  $P(p) > 1 - \epsilon$ , and A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ , then A has sufficient evidence that  $q$  (*evidence transmission*).
- (3.2) If A has sufficient evidence that  $q$ , and A attends to  $q$ , then A is rationally required to believe that  $q$  (*TR*).
- (3.3) If A has sufficient evidence that  $q$ , then either A is rationally required to believe that  $q$  or A is not attending to  $q$  (*from 3.2*).
- (3) If A has sufficient evidence that  $p$ , and A has sufficient evidence that  $p \rightarrow q$ , and A's evidential probability  $P(p) > 1 - \epsilon$ , and A's evidential probability  $P(p \rightarrow q) > 1 - \epsilon$ , then either A is rationally required to believe that  $q$  or A is not attending to  $q$  (*from 3.1 and 3.3*).
- (4) Therefore, either A is rationally required not to believe that  $p$ , or A is rationally required not to believe that  $p \rightarrow q$ , or A is rationally required not to have a credence  $C(p) > 1 - \epsilon$ , or A is rationally required not to have a credence  $C(p \rightarrow q) > 1 - \epsilon$ .

$\epsilon$ , or A is rationally required to believe that q, or A is not attending to q (from 1, 2, and 3).<sup>31</sup>

And (4) is all we need in order to explain (D).

*Evidence transmission, knowledge closure, and cleverly disguised mules*

At long last, I shall consider a further reason to question *evidence transmission*, which is independent of the worries about probability raised above and therefore raises doubts about this argument as well. Fred Dretske and Robert Nozick have both argued that knowledge is not closed under known logical implication – an assumption that both see as operative in a typical form of sceptical argument. Since knowledge and evidence are closely connected, their arguments might be applied to evidence as well. Dretske's famous example is that while visiting the zoo, you might know that this animal is a zebra without knowing that it is not a cleverly disguised mule, even though you also know that something's being a zebra implies that it is not a cleverly disguised mule.<sup>32</sup> It might similarly be claimed that one can have sufficient evidence that this is a zebra, sufficient evidence that this implies that it is not a mule, and yet lack sufficient evidence that it is not a mule. If this is possible, this casts doubt even on *single premise evidence transmission* and the revised multiple premise *evidence transmission*. We can stipulate, for the purpose of this example, that the evidential probability for the analytic truth that something's being a zebra implies its not being a mule is 1, and that the evidential

<sup>31</sup> The formal structure of the argument is this:

- (1)  $(P \& Q \& R \& S) \vee \neg(P \& Q \& R \& S)$ , *logical truth*
- (2.1)  $\neg(P \& Q \& R \& S) \rightarrow (\neg P \vee \neg Q \vee \neg R \vee \neg S)$ , *logical truth*
- (2.2)  $\neg P \rightarrow T$ , *TR*
- (2.3)  $\neg Q \rightarrow U$ , *TR*
- (2.4)  $\neg R \rightarrow V$ , *CR*
- (2.5)  $\neg S \rightarrow W$ , *CR*
- (2)  $\neg(P \& Q \& R \& S) \rightarrow (T \vee U \vee V \vee W)$ , *from 2.1-2.5*
- (3.1)  $(P \& Q \& R \& S) \rightarrow X$ , *evidence transmission*
- (3.2)  $(X \& Y) \rightarrow Z$ , *TR*
- (3.3)  $X \rightarrow (Z \vee \neg Y)$ , *equivalent to 3.2*
- (3)  $(P \& Q \& R \& S) \rightarrow (Z \vee \neg Y)$ , *from 3.1 and 3.3*
- (4)  $T \vee U \vee V \vee W \vee Z \vee \neg Y$ , *from 1, 2 and 3*

<sup>32</sup> Dretske (1970, 1015–1016).

probability for the proposition that the animal in question is a zebra is sufficiently high for *evidence transmission* to apply.

So should we follow Dretske and Nozick on knowledge, and, in particular, should we conclude that the same is true for evidence? Both of these steps might reasonably be denied. First, siding with Dretske and Nozick on knowledge has considerable costs; it is unclear whether it is compatible with the natural assumption that we can gain knowledge via deduction at all. In any case, we would have to give up the plausible idea that a person who knows that *p* and who competently deduces from this that *q* by an argument that she knows to be logically valid thereby comes to know that *q*.<sup>33</sup> Second, arguments against the relevant closure principle about knowledge do not necessarily support the rejection of analogous principles about evidence. For example, Nozick rejects knowledge closure on grounds of his counterfactual analysis of knowledge. According to this account, 'A knows that *p*' implies that A would not believe that *p* if *p* were false.<sup>34</sup> Assuming this *sensitivity condition*, it does indeed seem to follow that knowledge is not closed under known implication. For example, it may be true that in the closest possible worlds in which the animals in the zebra compound are not zebras, you would also not believe that they are zebras. Yet at the same time it may be true that in the closest possible worlds in which the animals in the zebra compound are cleverly disguised mules, you would still believe that they are not cleverly disguised mules.<sup>35</sup> Besides the fact that the sensitivity condition is quite controversial,<sup>36</sup> it is clear that this argument cannot be used against *evidence transmission*, since 'A has sufficient evidence that *p*' does not imply anything that can play the role the sensitivity condition plays in it. Nozick's explanation for why you do not know that the animals are not disguised mules is fully compatible with its being the case that you have sufficient evidence to believe that proposition. So even if we follow Dretske and Nozick in

---

<sup>33</sup> Cf. Williamson (2000, 117–118).

<sup>34</sup> See Nozick (1981, 172–178). A major motivation for this condition is that it avoids counting so-called Gettier cases, first introduced by Gettier (1963), as cases of knowledge.

<sup>35</sup> Cf. Nozick (1981, 204–211).

<sup>36</sup> For a critical discussion of the sensitivity condition, see e.g. Williamson (2000, ch. 7).

denying the closure principle about knowledge, we might still maintain that both of the revised evidence transmission principles hold true.<sup>37</sup>

I am inclined to believe that they are. Dretske seems to question this. He asks:

Do you know that these animals are not mules cleverly disguised by the zoo authorities to look like zebras? If you are tempted to say ‘Yes’ to this question, think a moment about what reasons you have, what evidence you can produce for this claim. [...] Have you checked with the zoo authorities? Did you examine the animals closely enough to detect such a fraud?<sup>38</sup>

In response, we might say that given our experience with people in general and zoos in particular, it is most likely that these animals are not mules in disguise. If this were not so, this would cast doubt on the assumption that we have sufficient evidence to believe that the animals are zebras. Perhaps if we believed that the animals are not disguised mules, this belief would not constitute knowledge, but it seems to me plausible that we would at least be justified to have such a belief.

So there are reasons to resist the suggestion that considerations along the lines of Dretske and Nozick show the revised transmission principles about evidence to be false. Suppose, however, that they do show this. In this case, I think they would likewise show that the conditions specified in (D) are not sufficient for irrationality. If you can have sufficient evidence that this animal is a zebra without having sufficient evidence that this animal is not a cleverly disguised mule, then why should it be irrational to have the former belief without having the latter? So if the explanation I gave for the structural irrationality claim (D) fails for the reasons discussed here, it fails only because there is nothing to explain in the first place. In any case, non-structural requirements of

---

<sup>37</sup> It might be wondered whether this result would be incompatible with Williamson’s equation of evidence and knowledge (cf. 2000, ch. 9). But this would be a misunderstanding. If  $p$  is evidence for  $A$  that  $q$ , Williamson’s thesis implies that  $A$  knows that  $p$ , but not that  $A$  knows that  $q$ . *Evidence transmission*, in contrast, is a principle about the propositions that are *supported* by our evidence, not about the propositions that *constitute* our evidence. If the propositions that constitute our evidence do not obey a principle of closure or transmission, it does not follow that the propositions supported by our evidence likewise do not obey such a principle.

<sup>38</sup> Dretske (1970, 1016).

rationality explain everything there is to explain about *modus ponens* irrationality.<sup>39</sup> Let me thus turn to instrumental irrationality, which will be discussed in a separate chapter.

---

<sup>39</sup> I would like to thank the following people for their helpful contributions to a little survey I conducted about the questions discussed in this section: Max Barkhausen, Branden Fitelson, Josh Gillon, Alex Gregory, Tim Henning, Arnulf Köhncke, Thomas Kroedel, Errol Lord, Jack Woods.



## Instrumental rationality as a system of categorical imperatives

---

The foregoing chapter provided the general outline for an explanation of structural irrationality in terms of rational requirements to respond to available reasons, and defended a detailed account of the explanation of three out of the four paradigmatic instances of structural irrationality discussed in this thesis. This chapter completes the task by providing a non-structural account of the fourth remaining instance of structural irrationality: instrumental irrationality. The explanation faces a particular challenge. While an argument of the same general structure as the arguments considered in the last chapter can show that means/end-incoherence is irrational in *determined* cases, i.e. in cases in which the intended action is either required or forbidden by the available reasons, it fails to account for *underdetermined* cases, in which the action is merely permitted (10.1). I discuss and reject various ways of meeting this challenge, according to which the actual adoption of an intention provides an independent reason, or a tie-breaker, or excludes everything but intending the means as too costly (10.2). I also reject views that explain instrumental irrationality as instances of weakness of will (normativism), or of *modus ponens* irrationality (cognitivism) (10.3). My own favoured view is that means/end-incoherence is irrational because it increases the risk of engaging in pointless activity and thereby wasting valuable resources, which gives one a positive reason to make a decision between taking the means and giving up the end. This account takes on board some attractions of the wide-scope account, while avoiding its flaws (10.4). It concedes, however, that the irrationality of means/end-incoherence is ultimately to be explained by the interaction of different factors, and is not without exceptions (10.5). I conclude by offering a brief summary of the results of this thesis (10.6).

### 10.1 Explaining instrumental irrationality (in determined cases)

The most difficult task for my account is to explain instrumental irrationality:

- (B) *Instrumental irrationality*: Necessarily, if A intends to  $\phi$ , and A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A believes that condition C obtains, and A does not intend to  $\psi$ , then A is irrational.

The particular difficulty that arises when we try to explain this claim in terms of response requirements of rationality is that actions and intentions are often *underdetermined* by our reasons: we often have neither decisive reason for nor decisive reason against some action or intention. Before turning to this problem, I shall first present an argument of the same general structure as the arguments considered in the last chapter, showing how response requirements entail that means/end-incoherence is irrational in *determined* cases, i.e. in cases in which the agent has decisive (evidence-based) reason either for or against the intended action.<sup>1</sup> Besides the response requirements that I used in the last chapter, the argument is based on one additional premise:

*Evidence-based transmission*: If A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A has sufficient evidence that condition C obtains, then A has decisive evidence-based reason to intend to  $\psi$ .

This is an evidence-based version of the transmission principle (which states that we ought to take the necessary means to actions we ought to perform). It connects decisive evidence-based reasons for an action to decisive evidence-based reasons for an *intention* to take what is *likely* to be a necessary means to the action. The general rationale for

---

<sup>1</sup> The argument is in terms of decisive *evidence-based* reasons, but given what I have argued in Chapter 8, it could just as well be presented in terms of decisive reasons, sans phrase. Yet I aim to present the arguments in this chapter in such a way that the thesis of Chapter 8 is left open. Those who are reluctant to accept that decisive reasons are just decisive evidence-based reasons may still accept the conclusion that rationality is normative in the sense that it is supported by the balance of *evidence-based* reasons.

this principle is as follows. If you have decisive evidence-based reason to  $\phi$ , this is because of some reason-giving facts that are part of your evidence. If you add to your evidence facts that make it sufficiently likely that  $\psi$ -ing is a necessary means to  $\phi$ -ing, then all these facts together plausibly provide decisive evidence-based reason to  $\psi$ . If you further add to your evidence facts that make it sufficiently likely that you will not  $\psi$  unless you intend to  $\psi$  (or whatever condition C consists in), then all these facts together plausibly provide decisive evidence-based reason to *intend* to  $\psi$ .

An example: You know that Radiohead is a terrific live band, which gives you decisive evidence-based reason to attend their concert. The ticket seller told you that there are only a handful of tickets left, which gives you sufficient evidence that buying a ticket today is a necessary means to attending the concert. In this case, it seems plausible to say that you have decisive evidence-based reason to buy tickets today (the reason being e.g. that doing so is likely to be the only way to enable you to attend the concert). You also know that you are very unlikely to buy tickets today unless you intend to buy them. In this case, it seems plausible to say that you have decisive evidence-based reason to intend to buy the tickets today. This is all that the evidence-based transmission principle says.

Strictly speaking, the derivative reasons are in place only if the 'can'-condition is satisfied, which is an independent constraint on any reason. If the tickets are sold out, then your reason to attend the concert and your evidence that buying tickets is a necessary means to attending do not provide a reason to buy a ticket, for you have such a reason only if you can buy the ticket.<sup>2</sup> However, the evidence-based transmission principle is concerned only with the reason to *intend* taking the means, and as before, I am implicitly taking for granted as a condition of all claims about reasons and rational requirements that the agent is psychologically capable of forming the relevant attitude.

---

<sup>2</sup> It is no reply that in such a case, there is also no reason to attend the concert. There would be no such reason if you could not buy tickets and buying the ticket were *in fact* a necessary means to attending the concert, because in this case, you could not attend the concert. But the example is based only on the weaker assumption that you have *sufficient evidence* that buying a ticket is a necessary means to attending, which is compatible with its not being the case that buying a ticket is a necessary means to attending. In this case, you could attend (and might have reason to attend) even though you could not buy a ticket.

Given that we are assuming that A is capable of forming the intention to  $\psi$ , the evidence-based transmission principle is surely plausible.

On basis of this principle, we can run the following argument:

- (1) [A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A has sufficient evidence that condition C obtains], or not (*logical truth*).
  - (2.1) If it is not the case that [A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A has sufficient evidence that condition C obtains], then A lacks decisive evidence-based reason to  $\phi$ , or A lacks sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A lacks sufficient evidence that condition C obtains (*logical truth*).
  - (2.2) If A lacks sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, then A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing (*TR*).
  - (2.3) If A lacks sufficient evidence that condition C obtains, then A is rationally required not to believe that condition C obtains (*TR*).
- (2) If it is not the case that [A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A has sufficient evidence that condition C obtains], then A lacks decisive evidence-based reason to  $\phi$ , or A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A is rationally required not to believe that condition C obtains (*from 2.1-2.3*).
- (3.1) If A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A has sufficient evidence that condition C obtains, then A has decisive evidence-based reason to intend to  $\psi$  (*Evidence-based transmission*).

- (3.2) If A has decisive evidence-based reason to intend to  $\psi$ , then A is rationally required to intend to  $\psi$  (*RRR*).
- (3) If [A has decisive evidence-based reason to  $\phi$ , and A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, and A has sufficient evidence that condition C obtains], then A is rationally required to intend to  $\psi$  (*from 3.1 and 3.2*).
- (4) Therefore, A lacks decisive evidence-based reason to  $\phi$ , or A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A is rationally required not to believe that condition C obtains, or A is rationally required to intend to  $\psi$  (*from 1, 2, and 3*).

In other words:

- (5) *If A has decisive evidence-based reason to  $\phi$ , then A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A is rationally required not to believe that condition C obtains, or A is rationally required to intend to  $\psi$  (equivalent to 4).*<sup>3</sup>

As is obvious, (5) falls short of explaining (B). (5) deems means/end-incoherence irrational only if one has decisive evidence-based reason to pursue one's end. It cannot explain why not intending the means believed necessary to an intended action is irrational in cases in which one does not have such a reason. It might be argued:

---

<sup>3</sup>The formal structure of the argument is this:

- (1)  $(P \& Q \& R) \vee \neg(P \& Q \& R)$ , *logical truth*  
 (2.1)  $\neg(P \& Q \& R) \rightarrow (\neg P \vee \neg Q \vee \neg R)$ , *logical truth*  
 (2.2)  $\neg Q \rightarrow S$ , *TR*  
 (2.3)  $\neg R \rightarrow T$ , *TR*  
 (2)  $\neg(P \& Q \& R) \rightarrow (\neg P \vee S \vee T)$ , *from 2.1-2.3*  
 (3.1)  $(P \& Q \& R) \rightarrow U$ , *evidence-based transmission*  
 (3.2)  $U \rightarrow V$ , *RRR*  
 (3)  $(P \& Q \& R) \rightarrow V$ , *from 3.1-3.2*  
 (4)  $\neg P \vee S \vee T \vee V$ , *from 1, 2, and 3*  
 (5)  $P \rightarrow (S \vee T \vee V)$ , *equivalent to 4*

(6)\* If A lacks decisive evidence-based reason to  $\phi$ , then A is rationally required not to intend to  $\phi$ .

(7)\* Therefore, A is rationally required not to intend to  $\phi$ , or A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A is rationally required not to believe that condition C obtains, or A is rationally required to intend to  $\psi$  (from 4 and 6\*).

(7)\* does entail (B), but it rests on an implausible assumption, namely (6)\*. One can be rationally permitted to intend an action that one lacks decisive evidence-based reason to perform, at least unless one has decisive evidence-based reason *not* to perform it. In many cases, one will not have decisive reason either for or against an action, and in such cases intending the action will neither be required nor forbidden by rationality.<sup>4</sup> We might argue instead:

(6.1) If A lacks decisive evidence-based reason to  $\phi$ , then A also lacks decisive evidence-based reason *not* to  $\phi$ , or A has decisive evidence-based reason not to  $\phi$  (*logical truth*).

(6.2) If A has decisive evidence-based reason not to  $\phi$ , then A is rationally required not to intend to  $\phi$  (*RRR*).

(6) If A lacks decisive evidence-based reason to  $\phi$ , then A also lacks decisive evidence-based reason *not* to  $\phi$ , or A is rationally required not to intend to  $\phi$  (from 6.1 and 6.2).

(7) Therefore, A both lacks decisive evidence-based reason to  $\phi$  and decisive evidence-based reason not to  $\phi$ , or A is rationally required not to intend to  $\phi$ , or A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A is rationally required not to believe that condition C obtains, or A is rationally required to intend to  $\psi$  (from 4 and 6).

---

<sup>4</sup> See e.g. Chang (2009, 249); Raz (1998). I assume this claim here simply on the basis of its intuitive plausibility. My general account does not presuppose it, however; I merely concede this claim as a challenge for my account.

In other words:

- (8) *Unless A both lacks decisive evidence-based reason to  $\phi$  and decisive evidence-based reason not to  $\phi$ , A is rationally required not to intend to  $\phi$ , or A is rationally required not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or A is rationally required not to believe that condition C obtains, or A is rationally required to intend to  $\psi$  (from 7).*

(8) is closer to an explanation of (B) than (5) was, but it still falls short of an explanation of instrumental irrationality. (8) deems means/end-incoherence irrational only in such cases in which one either has decisive reason to pursue one's end or decisive reason not to pursue one's end. It cannot explain that it is irrational not to intend the means believed necessary to an intended action in *underdetermined* cases, in which pursuing the end is neither required nor forbidden by one's reasons. The argument thus provides an explanation of why means/end-incoherence is irrational in determined cases, but not in underdetermined cases.

## 10.2 The problem of underdetermination

This result highlights a more general puzzle about the relation between reasons and instrumental rationality. On the present view, rationality is concerned with responding to reasons. Practical rationality is partly concerned with forming intentions in accordance with available reasons and carrying them out, by coordinating them, forming further intentions for acknowledged means, etc. But reasons underdetermine choice. In many cases, we have neither decisive reason for nor decisive reason against an action. It is rationally permitted to intend such actions as well as not to intend them. Once we actually form such an intention, however, this seems to change what we can rationally do in a way that cannot – at least not obviously – be traced back to the reasons that we had before making the choice. It is an important feature of practical rationality that it coordinates our intentions in such underdetermined cases as well. How can this be reconciled with the idea that rationality is concerned with responding to reasons?

The problem of underdetermination does not just arise for the account I am proposing, according to which rational requirements can ultimately be understood as requirements of reasons themselves. Every view that maintains an important necessary connection between rationality and reasons, such as the transparency account or the apparent reasons account, faces a similar question. Just as reasons themselves underdetermine choice, *beliefs* about reasons or *apparent* reasons underdetermine choice as well. If rationality consists in responding to beliefs about reasons or apparent reasons, we still have to understand how rationality coordinates intentions in cases that are underdetermined by one's believed or apparent reasons.<sup>5</sup>

### *Intentions as reasons*

Let us consider an example of an underdetermined case. You are pondering what to do with the rest of your evening. You have sufficient reason to watch some movie, and, as you know, you have to go to the video store in order to do that. But you also have sufficient reason to listen to a record instead of watching the movie. As far as the story goes, you may rationally intend to go to the store or not. Why does this change once you form the intention to watch the movie? One possible answer is, of course, that intentions are themselves reasons. According to this *intention-based-reasons view*, by forming an intention in an underdetermined case, we turn this case into a determined case, because our intention provides an additional reason for the intended action and thus for taking the means to it.<sup>6</sup> Once you intend to watch the movie, your sufficient reasons for watching it become decisive, and thus require you to go to the store. The intention-based-reasons view is, however, a version of the narrow-scope account of instrumental rationality and inherits all the problems that we have discussed above. Most significantly, it entails implausible bootstrapping.<sup>7</sup> If intentions were reasons, it seems that we could turn impermissible actions into permissible or even required actions simply by adopting an intention we have no reason to adopt. This is unaccept-

---

<sup>5</sup> The puzzle may also be described independently of rationality, as a puzzle about the normative significance of intentions alone (cf. Kolodny 2011a, 51–52).

<sup>6</sup> Cheng-Guajardo (forthcoming) adopts this view in response to the underdetermination problem.

<sup>7</sup> See esp. Bratman (1987, 23–27) and the discussion in Chapter 4.

able. Even if this result could somehow be avoided, it is implausible to assume that we can make it the case that we have reason for an action simply by adopting the intention to perform the action.

An alternative option is a *tie-breaker view*, according to which intentions provide additional reasons only if the action in question is supported by sufficient reasons independently of the intention.<sup>8</sup> This view is tailored to account for the problem of underdetermination while at the same time avoiding the most implausible cases of bootstrapping. Besides being *ad hoc*, the tie-breaker view also strikes me as unstable. It is difficult to see why the actual adoption of an intention should add normative weight in a case in which the intended action is permissible but not in a case in which the intended action is impermissible. Moreover, the tie-breaker view is vulnerable to what John Broome calls the “objection from changing one’s mind”.<sup>9</sup> Broome deploys this objection against the intention-based-reasons view, but it applies equally to the tie-breaker view. Both of these views entail that once one has adopted an intention in an underdetermined case, one thereby makes it the case that one ought to perform the intended action. But this is implausible. If one has formed an intention in an underdetermined case, there is not necessarily something wrong with changing one’s mind instead of forming an intention for the means. Unless one has made further investments, giving up one’s intention is a permissible alternative to carrying it out. For example, there does not seem to be anything wrong with not intending to go to the video store if you change your mind and decide to listen to your record instead of watching the movie.

---

<sup>8</sup> Compare Chang’s “hierarchical voluntarism”, according to which we can create reasons by will in (and only in) situations that are underdetermined by non-voluntarist reasons. Note, however, that according to Chang the reason is provided not by any simple intention, but by a specific voluntary attitude of “taking’ [...] a consideration to be a reason” (Chang 2009, 247). The claim that it is *possible* for us to create voluntarist reasons in cases of underdetermination does not explain why means/end-incoherence is irrational in such cases, as long as we do not presuppose that intentions *necessarily* involve the specific voluntary attitude that creates such a reason (which is not, as far as I can see, part of Chang’s view). Somewhat surprisingly, Raz seems to subscribe to a tie-breaker view with respect to *wants* rather than intentions: “There remains the simple point that if of two acceptable options one wants one thing and does the other, one is acting irrationally. If when offered a pear or banana, I have reason to take one and it does not matter which one, then if I want the banana but take the pear, I have acted irrationally. [...] In these circumstances, wants are reasons, though in being limited to this case they are very peculiar reasons” (Raz 1998, 62).

<sup>9</sup> Broome (2001a, 99–100).

Finally, it can also be questioned whether the intention-based-reasons view or the tie-breaker view really do provide an answer to the puzzle we are considering. Our question is why it is irrational to intend to  $\phi$  but not to intend the acknowledged necessary means  $\psi$ -ing in an underdetermined case, in which one has no decisive reason either for or against  $\phi$ -ing. The answer given by both of the discussed views is that intentions provide an additional reason in underdetermined cases, a reason which transmits to the necessary means and in this way makes intending the means mandatory rather than permissible. But what if one not only forms an intention to  $\phi$ , but at the same time forms a further intention that (as one knows) can only be realised if one omits  $\psi$ -ing? So for example, what if you not only intend to watch the movie, but also intend to put your feet up immediately and not leave home anymore? The case would then again be underdetermined by reasons, even if intentions were reasons. We could not claim that intending to go to the store was mandatory in such a case, since the reason provided by the intention to watch the movie would be counterbalanced by the reason provided by the intention not to leave home. Yet it would be irrational in such a case to intend to watch the movie but not to intend to go to the store. Neither the intention-based reasons view nor the tie-breaker view seem capable of providing an answer in such a case. Call this the “counterbalancing objection”.

#### *Intentions as investments*

Both the intention-based-reasons view and the tie-breaker view seek to explain instrumental irrationality by postulating that intentions themselves provide reasons. Other authors have claimed that while intentions do not provide reasons themselves, they often change what reasons we have in more indirect ways, and these authors have sought to explain the puzzle in terms of such reasons.<sup>10</sup> A natural starting point is that in many cases in which one has adopted an underdetermined intention, one has made further investments that add to one’s reasons.<sup>11</sup> If one has bought a ticket for a concert

---

<sup>10</sup> See esp. Kolodny (2011a) and Scanlon (2004). Note that both see the puzzle as being concerned primarily with the normative significance of intentions as such – an issue that, due to their view about reasons and rationality, they emphatically distinguish from questions about instrumental rationality.

<sup>11</sup> See e.g. Broome (2001a, 99).

because one has formed a merely permitted intention to attend it, then the expected loss of the value of the ticket might make the case for going to the concert (and taking further necessary means) conclusive. More generally speaking, if one has already taken some means to the intended action, one can realise this action with less cost, which normally strengthens on balance one's reasons to perform it and thereby can turn an underdetermined case into a determined one. This is an important factor that should be kept in mind. Yet the explanation is not available if one has not made any investments or taken any means to the intended action.

A related thought, which promises to provide a more general explanation, is that deliberation, or more generally the formation of intentions and decisions, is typically itself a costly activity.<sup>12</sup> The idea would be that if you have formed an intention in an underdetermined case, you have conclusive reason not to spend time and energy on reconsidering this decision as long as the circumstances do not change. Accordingly, T.M. Scanlon argues that in such situations you have "a purely pragmatic second-order reason not to reconsider that decision".<sup>13</sup> We might thus argue that once you have already formed the intention to watch the movie, and since you have sufficient reason to do that, you have a pragmatic reason not to reconsider this decision. One problem with this suggestion is that you might conform to this second-order reason not to reconsider your decision and still end up not intending to go to the video store. That this is a failure of rationality therefore cannot be explained by the failure to conform to the second-order reason alone. Drawing on Scanlon, Niko Kolodny suggests that the cost of deliberation not only provides a second-order reason not to reconsider one's decision, but affects one's first-order reasons as well. His reasoning is that intending an action is itself some kind of means to the action which involves at least a small investment. Once this investment is made, performing the action is less costly and the balance of reasons must now count stronger in favour of performing it than prior to forming the intention. Since prior to forming the intention the reasons for and against

---

<sup>12</sup> See already Bratman (1987, 26).

<sup>13</sup> Scanlon (2004, 246). Compare also Raz's claim that decisions are "exclusionary reasons", i.e. "reasons to disregard further reasons and arguments" (Raz 1975b, 490), which Scanlon discusses critically.

the action balanced each other, the agent must now have conclusive reason to perform the intended action and thus to take the necessary means to it.<sup>14</sup>

What about Broome's point that there is nothing in itself wrong with changing one's mind? As Kolodny himself concedes, changing one's mind is not always costly.<sup>15</sup> If you wait for the bus and have nothing better to do, it seems perfectly permissible to revise your intention. Moreover, the cost of revising an intention is often negligibly low, and it is disputable whether low cost changes of this sort generally turn the balance in cases of conflicting reasons, as Kolodny's argument presumes. To give an example, suppose that you have sufficient reasons for visiting each of two funerals that will take place simultaneously. Next, the railroad company issues a special offer, which means that you would spare five Euros if you went to one of the funerals rather than the other. Going to this funeral is now less costly than it was before. Yet it does not seem plausible to say that you therefore have decisive reason to go to that funeral; it seems that you still have sufficient reason to visit the other one.<sup>16</sup> Similarly, it is questionable whether one has decisive reason to carry out an intention one has adopted in an underdetermined case simply because revising the intention would involve some minimal cost. Hence the objection from changing one's mind seems to apply to Kolodny's suggestion as well. If revising one's intention has no costs, or these costs are negligible with regard to the primary reasons at issue, one does not have decisive reason to pursue one's intention – one may just as well change one's mind.

The cost explanation is also vulnerable to the counterbalancing objection. If in an underdetermined case, you both intend the end and intend not to take the means, then the cost of revising the former intention is counterbalanced by the cost of revising the latter. In such a case, the irrationality of intending the end without intending the means cannot be explained by the fact that one has already paid the costs of forming the intention for the end and has now stronger reason to pursue it than not to pursue it, for the same argument would apply to the intention not to take the means. Even

---

<sup>14</sup> Cf. Kolodny (2011a, 53–54).

<sup>15</sup> Cf. Kolodny (2011a, 54).

<sup>16</sup> Perhaps there are pragmatic explanations available that explain why we would not *say* that you lack sufficient reason to go the slightly more expensive funeral even if this is in fact true. Still, someone who denies that a case of this structure is possible, ultimately denies the possibility of incommensurability, which is supported by a variety of phenomena. See Chang (1997) for literature on this question.

though some cases of instrumental irrationality in underdetermined cases might be explained by deliberative costs, some other cases cannot be explained in this way.

*The predictive significance of intentions*

Kolodny offers a further explanation for how intentions affect one's reasons, which starts off from what Scanlon calls the "the predictive significance of intent"<sup>17</sup>, i.e. the fact that intentions often matter because they change what is likely to hold in the future. He argues as follows. If one intends to  $\phi$ , one generally makes it more likely that one will take certain insufficient means to  $\phi$ -ing. If one is more likely to take such means, then it is also more likely that taking *another* insufficient means will bring about one's  $\phi$ -ing. Since this other means thereby becomes *more effective* than before, one has *more reason* to take this means than one had before.<sup>18</sup> So for example, we might assume that once you intend to watch the movie, it is more likely that you will install the DVD player, which is an insufficient means to watching the movie. Since you are now more likely to install the DVD player, going to the video store, which is another insufficient means, becomes a more effective means to watching the movie; that is, it is now more likely that going to the store will bring it about that you actually watch the movie. For our purposes, let us assume that the second means (going to the store) is not only insufficient but also necessary. Kolodny's point about the predictive significance of intention can then be used to argue as follows: Prior to the adoption of an intention in an underdetermined case, the reasons for and against taking the necessary means balance each other. Since through the adoption of an intention, every insufficient means becomes a more effective means, after the adoption one must have decisive reason to take the (insufficient) necessary means.

It is worth noting that if this argument were successful, it could also only provide a partial answer to the problem of underdetermination. First, the effectiveness of *sufficient* means is not increased by the adoption of an intention (sufficient means are always fully effective). So if the necessary means in question is also sufficient, the inten-

---

<sup>17</sup> Scanlon (2008, 13).

<sup>18</sup> See Kolodny (2011a, 56).

tion makes no difference for its effectiveness. Second, as Kolodny mentions himself, there are cases in which intending an action does not even make insufficient means more effective, namely if one is independently very likely or unlikely to take other insufficient means, no matter whether one intends the action.<sup>19</sup>

In addition, I am sceptical that the argument is indeed successful. It rests on the assumption that the more effective a means to an action one has reason to perform is, the more reason one has to take this means.<sup>20</sup> This might seem plausible in the case of non-necessary means: *ceteris paribus*, one has more reason to take a more effective means than a less effective means.<sup>21</sup> Yet in the case of necessary means we have no such choice, and here the assumption strikes me as unwarranted. Intuitively, one's reasons for  $\phi$ -ing provide equally strong reasons for taking the necessary means to  $\phi$ -ing.<sup>22</sup> Therefore, how likely it is that one's taking the means will in fact bring about one's  $\phi$ -ing is neither here nor there, as long as we assume that one *can*  $\phi$  (which I take to be entailed by the assumption that one has reason to  $\phi$ ), and thus *can* take the necessary means to  $\phi$ -ing. If this is so, then the argument fails to establish its conclusion.

To sum up, there are reasons to doubt that either the intention-based-reasons view, or the tie-breaker view, or considerations about the predictive significance of intentions can provide a solution to the problem of underdetermination. Considerations about further investments or the cost of deliberation itself help to explain why many initially underdetermined cases turn into determined cases after one has adopted an intention. But some cases will remain underdetermined after the formation of an intention. If in such cases there is nothing wrong with not intending the means but instead *revising* one's intention for the end, why is there anything wrong with not intending the means and *keeping* the intention for the end?

<sup>19</sup> Cf. Kolodny (2011a, 57–58).

<sup>20</sup> See Kolodny (2011a, 56): "If some condition makes my taking some step, M [...] a more effective means to some goal, E, that there is reason for me to pursue, then, because of this, there is more reason for me to M."

<sup>21</sup> Note, however, that it is not uncontroversial that reasons transmit to non-necessary means in the way Kolodny suggests. For criticism of such a liberal transmission principle, see e.g. Rippon (2011, esp. §8).

<sup>22</sup> See e.g. Schroeder (2009, 245). I give an argument for this claim in my "Instrumental Normativity" (unpublished, §4).

### 10.3 Alternative explanations: normativism and cognitivism

Some authors have suggested explaining instrumental rationality by appealing to a version of the theory that intentional action is action under the guise of the good. T.M. Scanlon, for example, aims to account for the normative appearance of instrumental rationality by appealing to the claim that “insofar as a subject has [...] intentions, it must see these as responsive to its assessment of the reasons for these states”.<sup>23</sup> And Mark Schroeder’s *apparent reasons account* of instrumental rationality is based on the hypothesis that “intending entails believing that you ought”.<sup>24</sup> In both cases, the general idea is that since intentions involve attitudes with normative content, means/end-incoherence can be understood as a variant of *akrasia* or weakness of will, i.e. as a case in which an agent’s attitudes are not in line with his own beliefs about what he has reason to do.<sup>25</sup> Let us call this the *normativist approach*.<sup>26</sup> Since I have already explained the irrationality of weakness of will above, the normativist approach could be employed on behalf of the account I am suggesting.

The approach faces a dilemma, however. On the first horn, we do indeed assume that intentions entail ‘ought’-beliefs. In this case, instrumental irrationality could be explained as a variant of *akrasia*. On the present account, we could argue that one either has decisive reason to revise the end-intention (namely if one lacks sufficient evidence for the ‘ought’-belief involved), or one has decisive reason to revise the

<sup>23</sup> Scanlon (2007, 87–88). One problem with Scanlon’s approach is that the “must” in this quotation (as well as in many other passages of this article) oscillates between a constitutive and a rational sense of necessity. Since Scanlon’s general aim is to explain the normative appearance of rational necessity in terms of the normative content that attitudes have (as a matter of constitutive necessity), it seems that the relevant sense of “must” has to be one of constitutive necessity. Other passages suggest otherwise, however. I am here assuming an interpretation according to which the “must” is constitutive, but different interpretations of Scanlon remain possible.

<sup>24</sup> Schroeder (2009, 237). In the course of discussing objections to this claim, Schroeder seems willing to concede that intentions only entail believing that one has “adequate reason” for the intended action. As I shall presently argue, this claim is too weak to render means/end-incoherence irrational.

<sup>25</sup> For reasons of brevity, I focus on normative *beliefs*. Insofar as normative truth-taking attitudes such as ‘seeing R as a reason’ or ‘taking R to be a reason’ are different from belief, analogous points could be made in that terminology.

<sup>26</sup> Korsgaard subscribes to a particular version of the normativist approach, according to which willing an end involves not only believing that one has reason to pursue it, but believing that the willing itself provides the reason: “for the instrumental principle to provide you with a reason, you must think that the fact that you will an end *is a reason* for the end. It’s not exactly that there has to be a *further* reason; it’s just that you must take the act of your own will to be normative for you” (Korsgaard 1997, 245–246). For a helpful critical discussion of Korsgaard’s view, see Wallace (2001, §1).

means/end-belief (namely if one lacks sufficient evidence for this belief), or one has decisive reason to intend the means (namely if one has sufficient evidence for both of the former propositions). The problem is that the assumption that intentions involve 'ought'-beliefs is rather obviously false. It is not only possible, but in many cases entirely rational to intend an action that one does not believe one ought to perform. It is not only possible that our choices are underdetermined by reasons, it is also possible that we acknowledge this fact and still make decisions.<sup>27</sup>

On the second horn of the dilemma, we assume only that intentions involve the belief that one has *sufficient reason* for the intended action. Indeed, this seems to be the strongest assumption that the normativist approach can hope to establish (and it is still quite a strong assumption).<sup>28</sup> Yet this claim does not suffice for our purposes. There is nothing wrong with not intending what one believes one has merely *sufficient* reason to do, and consequently nothing wrong with not intending the means believed necessary to what one believes one has sufficient reason to do.

A different approach seeks an explanation of instrumental rationality in terms of theoretical rationality.<sup>29</sup> The general idea is that intentions are inextricably linked with certain beliefs about one's own future behaviour, that incoherent means/end-attitudes thus involve inconsistent beliefs, and that instrumental irrationality can therefore be traced back to theoretical kinds of irrationality such as *modus ponens* irrationality. Following Michael Bratman, I shall call this view *cognitivism about instrumental rationality*.<sup>30</sup> According to one version of the cognitivist view – put forward by Gilbert Harman, Kieran Setiya, and David Velleman – intending to  $\phi$  involves believing (or

<sup>27</sup> Schroeder replies that intentions do not entail an *antecedent* 'ought'-belief, but maintains that once one intends, one has this belief. The assumption in the background is that intentions themselves provide additional reasons in antecedently underdetermined cases (cf. Schroeder 2009, 237–238). I have rejected this assumption above. But note that even if it were true, it would remain unclear why it should be *impossible* for an agent to intend an action without believing that this intention provides an additional reason.

<sup>28</sup> It is a strong assumption particularly because of the possibility of *akrasia* in the sense of intending what one believes one ought not to do (cf. Wallace 2001, §1). I assume that such cases would have to be understood as involving contradictory beliefs (cf. also Schroeder 2009, 238–239). It is questionable, however, that all cases of *akrasia* involve contradictory normative beliefs. The objection is a general problem for the normativist view, it applies independently of whether intentions are supposed to involve beliefs in 'oughts' or sufficient reasons.

<sup>29</sup> See Harman (1976); Setiya (2007a); Velleman (1989, esp. ch. 3); Wallace (2001, §4).

<sup>30</sup> See Bratman (1991; 2009c; 2009d).

even amounts to a particular kind of belief) that one is going to  $\phi$ .<sup>31</sup> Let us assume that the relevant beliefs in the means/end-relation and condition C include the proposition that one is not going to  $\phi$  unless one now intends to  $\psi$ . Further, let us assume that the means/end-incoherent agent does not falsely believe that he intends to  $\psi$ . It then follows that the means/end-incoherent agent is *modus ponens* irrational: he believes that he is going to  $\phi$ , he believes that he is not going to  $\phi$  unless he now intends to  $\psi$ , but he does not believe that he now intends to  $\psi$ . I have explained *modus ponens* irrationality above in terms of non-structural requirements of rationality. The cognitivist view could therefore be employed on behalf of my account.

One problem for this explanation is that agents can be means/end-incoherent while *falsely* believing that they intend to take the means, in which case their incoherence does not involve an inconsistency in beliefs.<sup>32</sup> In response to this worry, R. Jay Wallace argues that in the relevant deliberative contexts, one is independently rationally required to form accurate beliefs about one's intentions.<sup>33</sup> This seems to presuppose that the condition of having or not having an intention is, in Timothy Williamson's terminology, luminous, i.e. that we are always in a position to know, or at least have sufficient evidence, whether we have or lack an intention.<sup>34</sup> This assumption is controversial. It seems, however, that a reasonable response on behalf of the cognitivist is available even if intention-states are not necessarily luminous. First, cases in which we lack evidence for our intention-states will be exceptional. We do at least usually have sufficient evidence that we have or lack an intention. Second, as Wallace argues convincingly, it is independently questionable that means/end-incoherence is irrational in the exceptional cases in which we lack evidence for our intention-states.<sup>35</sup> Irrationality is something that we can avoid by employing our rational capacities; conditions of irrationality must therefore be rationally accessible. Insofar as intention-states

<sup>31</sup> See Harman (1976, §1) – though see also Harman (1986, ch. 8) for a less committed view; Setiya (2007b, pt. 1); Velleman (1989, ch. 4). The view is often motivated by reference to Anscombe's (1957, 1) remark that the characteristic expression of an intention is the assertoric utterance of the sentence "I am going to  $\phi$ ".

<sup>32</sup> See e.g. Bratman (2009c, §4; 2009d, §4); Brunero (2005, 6; 2009a, §2).

<sup>33</sup> Cf. Wallace (2001, 22; 2006, 117–120).

<sup>34</sup> See Williamson (2000, esp. ch. 4) for general scepticism about the luminosity of mental states.

<sup>35</sup> Wallace (2001, 23): "... it seems to me doubtful that intentions that are cut off in this way from conscious belief really do introduce rational constraints on our further attitudes, of the kind represented by the instrumental principle." See also Kolodny (2007a, 253).

are not necessarily rationally accessible, structural irrationality claims like (B) must themselves be restricted to rationally accessible intention-states.

A more significant problem for the cognitivist view is that the assumption that intending to  $\phi$  necessarily involves believing that one is going to  $\phi$  seems too strong. One might intend to perform actions without being certain enough to believe that one will succeed in them, such as producing ten carbon copies with a typewriter (to cite a famous – though arguably outdated – example by Donald Davidson).<sup>36</sup> Wallace suggests that the cognitivist view can be defended on the much weaker (and much more plausible) assumption that intending to  $\phi$  involves believing that it is possible for one to  $\phi$ . According to his view, the means/end-incoherent agent has the following beliefs:

- (1) It is possible that I  $\phi$ .
- (2) It is possible that I  $\phi$  only if I intend to  $\psi$ .
- (3) I do not intend to  $\psi$ .<sup>37</sup>

And having these beliefs is *modus ponens* irrational. While I am sympathetic to the thesis that intentions involve a belief in (1), I have doubts that (2) correctly represents the relevant belief about necessary means. If A believes that an intention to  $\psi$  is necessary for  $\phi$ -ing, then A believes something like (4): ‘In all accessible possible worlds, I will  $\phi$  only if I intend to  $\psi$ ’. It follows that (5) ‘It is possible that I  $\phi$  only if it is possible that I intend to  $\psi$ ’. It also follows that (6) ‘I will actually  $\phi$  only if I actually intend to  $\psi$ ’. But it does not follow that ‘It is possible that I  $\phi$  only if I *actually* intend to  $\psi$ ’, as Wallace’s premise (2) states. Another way to make this point is as follows. If the agent’s belief about necessary means is true, and he remains in a means/end-incoherent state, then it will turn out that he does not achieve his end. But it will not turn out that achieving his end was not possible. By the same token, *belief* in the necessary means relation and the absence of the means-intention should not be inconsistent with *belief* in the possibility of achieving the end, but only with belief that the end will

<sup>36</sup> Davidson (1971, 50; 1978, 91–96). See also Bratman (1987, 37–38).

<sup>37</sup> See Wallace (2001, 21; 2006, 112).

actually be achieved. For these reasons, I am sceptical that an appeal to the weak link between intention and belief (1) suffices to convict the means/end-irrational person of logical inconsistency.<sup>38</sup>

Finally, there is the natural worry that the cognitivist view simply fails to account for the practical dimension of means/end-coherence.<sup>39</sup> According to cognitivism, the concern for means/end-coherence is ultimately exhausted by a concern for representing the world as it is, but it seems plausible that means/end-coherence is related to a dimension of rationality that is distinctively concerned with practical matters such as achieving our ends or being effective agents. Some cognitivists seem to suggest that the idea of a genuinely practical rational concern with means/end-coherence beyond the concern with practical reasons that we have independently of our aims is an illusion.<sup>40</sup> Cognitivism can thus be understood as a kind of error theory that aims to explain the impression that means/end-incoherence is irrational independently of whether our ends accord with our reasons. As we have seen, however, antecedent reasons often underdetermine whether we should pursue or refrain from pursuing a particular end. So even if there is nothing practically rational about intending the means to ends that are not sufficiently supported by practical reasons, it remains plausible that there is a distinctive practical dimension to establishing means/end-coherence in underdetermined cases.

---

<sup>38</sup> In the postscript to his original article, in which Wallace responds to related worries about his conception of possibility put forward by Brunero (2005, 4–6), Wallace concedes that “so long as it is possible for the agent to change their intentions regarding the means, it is possible for them to achieve the end” (2006, 115). He argues, however, that the sense of ‘possibility’ relevant for his argument is constrained by one’s present intentions. Accordingly, something is possible in the relevant sense only if it is compatible with the execution of one’s present intentions. The problem is that on this reading, (1) is not, as Wallace originally claimed, “much weaker, and hence more plausible” (2001, 20) than the assumption of other cognitivists that intending to  $\phi$  involves believing that one is going to  $\phi$ . That intending to  $\phi$  involves believing that  $\phi$ -ing is compatible with the execution of all one’s other intentions is a logically independent assumption, and it is quite a strong one. We should be wary of accepting it without a strong argument for it. *Prima facie*, it seems *irrational* rather than *conceptually impossible* to intend an action that one does not believe to be compatible with the execution of one’s other intentions. Furthermore, it is not clear that a belief in (2) must be involved in means/end-incoherence even in this constrained sense of ‘possibility’, for one’s *forming* an intention to take the means is normally compatible with the execution of one’s present intentions. At least typically, means/end-incoherent agents must regard it as possible for them to *form* the intention to take the means; as long as they believe that, there is no incompatibility with their belief that it is possible for them to achieve the end even if they believe that they do not presently have the intention to take the means.

<sup>39</sup> See e.g. Brunero (2005, 4).

<sup>40</sup> Cf. Setiya (2007a, 673), Wallace (2006, 119–120).

#### 10.4 A reason to decide

For reasons that will become apparent shortly, I believe it is helpful at this point to take a look back at the wide-scope account of instrumental rationality. The wide-scope view states that we are rationally required [not to intend to  $\phi$ , or not to believe that  $\psi$ -ing is a necessary means to  $\phi$ -ing, or to intend to  $\psi$ ]. One of the various problems for this account discussed in earlier chapters was the difficulty of finding a reason in favour of satisfying this disjunctive requirement that necessarily holds (Chapters 5.1–5.3). Worse, I have argued that not even a possible reason can be found for which one can make this disjunction true. This is partly due to the belief-element in the disjunction: there seems to be no single consideration that can rationally guide one in giving up a belief in a means/end-relation *or alternatively* making a change in one's intentions (Chapter 6.6).

Jonathan Way's refined wide-scope account of instrumental rationality aims to do justice to this latter point. Way makes two claims:

*Reason claim:* The fact that  $\psi$ -ing is a necessary means to  $\phi$ -ing is a decisive reason to [not intend to  $\phi$  *or* intend to  $\psi$ ].<sup>41</sup>

*Rationality claim:* If A believes that  $\psi$ -ing is a necessary means to  $\phi$ -ing, then A is rationally required to [not intend to  $\phi$  *or* intend to  $\psi$ ].<sup>42</sup>

This interpretation of the wide-scope account is better suited to account for the guidance problem. Since the disjunctive response called for no longer involves a belief-element, at least one important obstacle for thinking that the wide-scope instrumental requirement cannot be guiding is removed. Moreover, Way makes a positive proposal regarding the content of the guiding consideration for giving the disjunctive response:

---

<sup>41</sup> See Way (2010a, 222); cf. also Way (2012b, 494). Way (2012b, 496–498) later restricts this claim to so-called 'reasons of the right kind'. We can ignore this here because I am working on the assumption that there are no wrong kinds of reasons for attitudes – an assumption that Way himself defends in a different article (Way 2012a).

<sup>42</sup> See Way (2010a, 223; 2012b, 494).

according to him, it is simply the consideration that  $\psi$ -ing is a necessary means to  $\phi$ -ing itself.

While I am looking for a view according to which rationality consists in responding to reasons, Way's view is part of an *apparent reasons account* of rationality; it entails that we may have no reason to do what rationality requires (namely in case our beliefs happen to be false).<sup>43</sup> Way's claims about reasons and rationality express exactly the kind of dualism criticised above, according to which reasons are related to facts and rationality is related to subjective states. The general starting point of the present account – namely that both reasons and rationality are connected to evidence –, gives us a way to merge Way's reason claim with his rationality claim:

*Normative instrumental rationality (NIR):* If A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, then A has decisive evidence-based reason, and is thus rationally required, to [not intend to  $\phi$  or intend to  $\psi$ ].

This would be congenial to the account I am giving here. In case one lacks sufficient evidence for the means/end-relation, one is required not to believe in it. In case one has such evidence, one is required to [not intend the end or intend the means]. It follows that anyone who believes in the means/end-relation and intends the end but not the means is irrational.

There are some reasons to be sceptical of this explanation, however. In what follows, I shall first discuss a number of problems of NIR, and then suggest a modified account that avoids these problems. The first problem is that the fundamental normative assumption of the account appears to be *ad hoc*, at least at first glance. Why exactly does the fact that  $\psi$ -ing is a necessary means to  $\phi$ -ing (or facts that sufficiently indicate that this fact occurs) provide a decisive reason to [not intend to  $\phi$  or intend to  $\psi$ ]? Certainly, this claim is in need of justification.<sup>44</sup> Second, the explanation still rests on

<sup>43</sup> One further problem is that Way's account involves implausible bootstrapping of rational requirements; see note 60 in Chapter 6.6.

<sup>44</sup> Way's own justification is not very plausible. He argues that since "intending [...] involves a [...] cost", "there is always *some* reason against intending to do something". Further, "these costs will also provide some reason against [intending to E but not intending to M]" (Way 2012b, 504). Since there is no reason *for*

the assumption that there are reasons for disjunctions of attitude-states, which raises some problems we have already encountered. To begin with, any such reason would have to satisfy what I have called the “commonality condition” in Chapter 6.6: that it could rationally guide an agent in giving each of the responses that are part of the disjunctive response favoured by the reason. Now, the consideration that  $\psi$ -ing is a necessary means to  $\phi$ -ing *by itself* does not seem to provide normative guidance for either dropping an intention to  $\phi$  or forming an intention to  $\psi$ . It can provide guidance for dropping the intention to  $\phi$  if one *also* assumes that one has reason not to  $\psi$ , and it can provide guidance for forming the intention to  $\psi$  if one *also* assumes that one has reason to  $\phi$ . But we have not yet identified a common consideration that could provide guidance for each of the responses, which is what someone who makes the distinctive claim about the existence of a disjunctive requirement or reason is committed to. A further, more general point is that it is not clear whether a reason for a disjunction of attitude-states could be object-given.<sup>45</sup> Assuming that there are only object-given reasons for attitudes, this means that it is not clear whether reasons for disjunctions of attitude-states (or, equivalently, reasons against conjunctions of attitude-states) exist in the first place. Contrast a reason against intending to [ $\phi$  and not- $\psi$ ] with a reason against [intending to  $\phi$  and not-intending to  $\psi$ ]. Clearly, the first reason might be provided by the object of the intention, namely the conjunctive action [ $\phi$ -ing and not- $\psi$ -ing]. Properties of this action, such as having bad effects or being impossible, might provide object-given reasons against intending it. It is far from clear, however, whether

---

[intending to E but not intending to M] (cf. Way 2012b, 496–498), there is always conclusive reason against it. This explanation has the awkward implication that the reason for being instrumentally rational is just the reason one has against having any intention at all. It also does not fit with Way’s earlier contention that the reason in question is the fact that M-ing is a necessary means to E-ing (Way 2010a, 222). Moreover, the explanation rests on the assumption that a reason against X-ing entails a reason against [X-ing and Y-ing], which is equivalent to the principle that a reason for X-ing entails a reason for [X-ing or Y-ing], known as *disjunction introduction*. This principle is very controversial, partly because it leads to Ross’s paradox (cf. A. Ross 1941, 61–62). I agree with those who reject it (see e.g. Broome 2007b, 19–22; Schroeder 2011b, 20–21). One disadvantage of *disjunction introduction* is that it allows us to reintroduce the belief-element into the required response: if there is a reason to [not-intend to  $\phi$  or intend to  $\psi$ ], then there will also be a reason to [not-intend to  $\phi$  or intend to  $\psi$  or not-believe that p]. In my view, this would render void one of the greatest advantages that Way’s account has in comparison to the standard wide-scope account.

<sup>45</sup> Way (2010a, 230–232) addresses this question himself, but he focuses exclusively on the question of whether an object-given reason for an intention must be provided by a reason for the action or can also be provided by some other property of the action. I think that his arguments for the latter view are convincing, but this does not resolve the more general worry mentioned in the main text.

the second reason could be object-given, for there are two intention-states with different objects at issue, the disjunction of which is supposed to be supported by one single reason. All I want to say is that the postulation of reasons for disjunctions of attitude-states comes with the theoretical burden to answer some difficult questions that such reasons pose. So far, the account I am suggesting has been free of that commitment.

Last but not least, it is arguable that NIR (as well as Way's corresponding claims about reasons and rationality) entails implausible satisfaction claims and thus involves a similar kind of overgeneralisation that I argued is involved in the standard wide-scope account. NIR entails that in every possible world in which you know that  $\psi$ -ing is a necessary means to  $\phi$ -ing, you satisfy a rational requirement if you give up your intention to  $\phi$ . But suppose that you also know that you have overwhelming reason to  $\phi$  and no reason at all to refrain from  $\psi$ -ing. In such a case, it does not seem very plausible to say that you satisfy a rational requirement if you give up your intention to  $\phi$ ; you have no rational basis for doing so.<sup>46</sup> This casts doubt on NIR.

Let me start with the first problem. Why would facts which indicate that certain means are necessary for certain ends provide reasons against having these ends without intending the means? In my view, the most plausible answer to this question lies in economic considerations. Intending an end without intending the means increases the risk that we engage in pointless activity, and this is something we have reason to avoid doing. Means/end-incoherence increases this risk because – as especially Michael Bratman has emphasised – intentions are in part constituted by various dispositions for further deliberation and planning, including the disposition to take means to the actions intended.<sup>47</sup> If one has evidence that some means is necessary for performing an action one intends, but one does not intend to take this means, then maintaining the intention increases the risk that one will take means to an action that one will not eventually perform and to deliberate and plan in unnecessary ways. One thereby risks valuable resources, which one could avoid without loss by avoiding

---

<sup>46</sup> I assume that you also have no other belief on the basis of which you can rationally drop your intention. The only belief that I cannot exclude is the belief that  $\psi$ -ing is a necessary means to  $\phi$ -ing. But I take it that this belief is no rational basis for dropping an intention to  $\phi$  if one knows that one has no reason to  $\psi$  and overwhelming reason to  $\phi$ .

<sup>47</sup> See Bratman (1987).

means/end-incoherence. This is what I take to be the plausible normative background assumption of NIR.<sup>48</sup>

Yet this answer to the first problem only reinforces the second one. To say that a reason against a combination of attitude-states is provided by the fact that the combination in question increases the risk of pointless action is to cite a state-given rather than an object-given reason. In substance, it is like saying that the prospects of getting an award for believing something is a reason to believe it. But we cannot rationally respond to such reasons by giving the favoured response *for* those reasons. Just as we cannot believe  $p$  for the reason that our being in this state will be rewarded, we cannot intend to  $\psi$  or give up an intention to  $\phi$  for the reason that our being in this state minimises the risk of pointless action. And I agree with Derek Parfit and others who think that this calls into question that the reasons in question really are reasons for those attitudes rather than reasons for *actions* that bring these attitudes about.<sup>49</sup>

My suggestion, then, is that the reason in question is not really a reason for a disjunction of attitudes, but a reason for an action. More exactly, it is a reason to make up one's mind and decide between  $\phi$ -ing and not- $\psi$ -ing.<sup>50</sup> Now, some people use the words 'decision' and 'intention' interchangeably, and I do not want to argue about words. What is important for the view I am suggesting are two points. First, there is a difference between the attitude-state of having (or lacking) an intention, and the mental act of deciding, which has such an attitude-state as an outcome. In particular, there is a difference between being in the state of [not-intending to  $\phi$  or intending to  $\psi$ ] and making a decision between  $\phi$ -ing and not- $\psi$ -ing. There is no doubt that it is possible to

---

<sup>48</sup> Bratman (see e.g. 1987, 28–35; 2009a, 417) and Raz (2005a, 17–18) express a related idea when they argue that means/end-incoherence undermines effective agency, and effective agency is valuable. As both note, however, this does not show that one has reason to avoid means/end-incoherence in each particular case. Insofar as means/end-coherence contributes to effective agency in the particular case, its value depends on the intended end. Insofar as it contributes because it manifests a generally valuable disposition, it does not follow that there is reason to be means/end-coherent in each particular case. In contrast, it seems to me that insofar as means/end-incoherence increases the risk of pointless action, this is indeed a reason to avoid it in each particular case. (As I shall argue, however, there can be exceptional cases in which means/end-incoherence does not increase this risk.)

<sup>49</sup> Cf. Parfit (2011a, App. A).

<sup>50</sup> Note that this is in principle compatible with the view that reasons for action reduce to reasons for intention. Accordingly, the relevant reason would be a reason to intend to make a decision between  $\phi$ -ing and not- $\psi$ -ing.

be in that state without having made such a decision, so the difference cannot reasonably be denied. Second, the mental act can rationally be carried out for reasons that are provided by the benefits that some state, which would be the outcome of this act, would have. Such reasons *would* be state-given reasons if they were reasons for an attitude-state, and consequently, they would be reasons for which we could not adopt the relevant attitude-state. But since they are reasons only for the act of deciding, they do not pose the problems that state-given reasons for attitudes pose (at least as long as there are available object-given reasons for which one can adopt the attitude when making the decision). For example, it seems unproblematic to say that you made a decision between several attractive pizza options for the reason that doing so would prevent your friends from waiting much longer. But it is less clear whether you could adopt the intention to take a particular pizza for such a reason. This point, too, seems to me rather obvious. The most natural cases of reasons for making a decision between a number of options at a certain time are considerations that have to do with the benefits that making a decision would have at that time.

Given these points, it is arguable that the fact that means/end-incoherence increases the risk of wasting resources in pointless activity is a reason to make a decision between the pursuit of an end and the refusal to take the means. Even if making this decision has the relevant disjunction of attitude-states as an outcome, this reason is a reason for the mental act, not for the disjunction of attitude-states.<sup>51</sup> Since the mental act of deciding nevertheless involves the formation or revision of an intention, the question is how this act can be accomplished rationally unless we can assume the presence of object-given reasons for the attitude-states involved. But we can usually assume the presence of such reasons, if only because actions normally have costs. So at least usually there will be some reason or other for or against  $\phi$ -ing on the basis of

---

<sup>51</sup> The reason for the decision also does not *entail* the reason for the disjunction of intention-states, since the principle 'If one has reason to  $\phi$ , and one's  $\phi$ -ing entails that one  $\psi$ -s, then one has reason to  $\psi$ ' is invalid (see e.g. Broome 2007b, 19–22; Schroeder 2011b, 20–21). Note that by the same principle, we could conclude that if one has reason to decide between  $\phi$ -ing and  $\psi$ -ing, then one has reason against [intending to  $\phi$  and intending to  $\psi$  *and believing that*  $1+1=2$ ]. We would thereby not only reintroduce the problems for hybrid requirements discussed in Chapter 6.6, but also commit ourselves to the absurdity that simply because one has a reason to decide between  $\phi$ -ing and  $\psi$ -ing, one does something one has reason to do if one gives up the belief that  $1+1=2$ .

which one can rationally revise an intention to  $\phi$  or form an intention to take the necessary means to  $\phi$ -ing. The question only occurs if we assume that there is *no reason at all*, neither for nor against  $\phi$ -ing. If such a case is possible, it seems to me plausible to say that agents could rationally give up their intention on the basis of the consideration that they have no reason to perform the intended action. It follows that there will always be some object-given reason on the basis of which agents can rationally form or revise their intentions as required to make the decision that I maintain we have reason to make.

Let me turn to the third problem then, the problem of implausible satisfaction claims. To begin with, a clarification is in order. The economic considerations mentioned provide a reason to make a decision between  $\phi$ -ing and not- $\psi$ -ing only in cases where (a) one has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, (b) one intends to  $\phi$ , (c) one does not intend to  $\psi$ , and (d) one has sufficient evidence that condition C obtains. Otherwise one has no reason to make such a decision (at least not one stemming from the aforementioned considerations). It follows that one's making such a decision amounts to the satisfaction of a requirement only if these conditions are fulfilled beforehand. In view of this fact, we can now reformulate the challenge mentioned above. Suppose that an agent, in addition to satisfying these conditions, also knows that he has overwhelming reason to  $\phi$  and no reason at all to refrain from  $\psi$ -ing. If he now decides against  $\phi$ -ing and gives up the intention to  $\phi$  – is it still plausible to say that he has satisfied some requirement or done something he had reason to do? I am inclined to answer “yes”. Even though making the decision in this particular way involves irrationality, the fact that making the decision avoids the risk of engaging in pointless activity still seems to count in favour of it. Thus the satisfaction claim that follows from the suggested requirement does not strike me as implausible. It can also contribute to an explanation of the intuition discussed in Chapter 9.3, namely that escaping a structurally irrational state in the wrong direction can appear to amount to satisfying a requirement. I have argued that this phenomenon can be explained by reference to three factors: escaping incoherence in the wrong direction can manifest a generally valuable disposition, it can indicate that an agent satisfies a second-order

requirement to reflect upon and reassess his reasons, and it can qualify as correct reasoning if the attitudes are based on one another in a rationally permissible way. In principle, this explanation is available in the case of instrumental irrationality as well. But the economic considerations, which are independent of one's reasons for and against the intended action, add a further factor. According to the present proposal, escaping means/end-incoherence by making a decision against the balance of one's reason still counts as doing something one has reason to do.

It is, however, worth pointing out that there is an alternative view available for those that do not share this intuition. The alternative view is that the reason to make the decision is conditional not only on the features mentioned in (a)-(d), but additionally on the fact that the agent is in an underdetermined case. Since I have already provided an explanation of instrumental irrationality in determined cases, this weaker claim would suffice for my account.

It might be wondered whether the reason to make a decision, which I postulate here, would not be vulnerable to the transmission problems for the wide-scope account discussed in Chapter 4.7. If one cannot give up one's intention, does it not follow that intending the means is a necessary means to making the decision, and that one therefore has a reason to form this intention, no matter how stupid or crazy the intended action is? It does not follow. First, it is not clear whether intending the means in such a case is suitably described as a necessary *means* rather than a necessary side effect of making the decision, and reasons do not transmit to necessary side effects. Second, it is generally questionable that reasons for action transmit to reasons for attitudes. For example, if you have reason to act in a kindly manner towards your brother, and you can do that only if you believe that it is his birthday, it does not seem to follow that you have a reason to believe that it is his birthday. Last but not least, it is also arguable that if you cannot give up your intention to  $\phi$ , then you cannot really decide *between*  $\phi$ -ing and not- $\psi$ -ing, which would suspend the reason in question.<sup>52</sup> As far as I can see, transmission problems are not to be feared.

---

<sup>52</sup> Are you not nonetheless irrational if you intend to  $\phi$  but you do not intend to  $\psi$ ? If you have seriously tried and failed to revise your intention to  $\phi$ , then it seems to me plausible to say that you are compulsive rather than irrational in a sense that is relevant for legitimate criticism.

To sum up, I have argued that considerations about the economical deployment of resources justify the following claim:

*Reason to decide:* If A intends to  $\phi$ , A has sufficient evidence that  $\psi$ -ing is a necessary means to  $\phi$ -ing, A has sufficient evidence that condition C obtains, and yet A does not intend to  $\psi$ , then A usually has a strong evidence-based reason to make a decision between  $\phi$ -ing and not- $\psi$ -ing.<sup>53</sup>

Unless there are other, comparably strong evidence-based reasons not to make such a decision, this reason is decisive and rationality requires us to make the decision. This explains why means/end-incoherence is irrational in a significant number of cases, in a way that appeals to a distinctively practical concern and is compatible with the plausible assumption that it is often permitted to change one's mind in underdetermined cases.

### 10.5 Instrumental rationality: final thoughts

If what I have argued in the last section is correct, then economical considerations usually provide a strong reason for making decisions that ensure that one is means/end-coherent. The question remains whether such a reason can fully account for the irrationality of means/end-incoherence. First, the reason in question could be outweighed by other reasons that count against making the decision. For example, one might have other very urgent things to do, or one might have been promised a large reward for not making the decision. Since the reason in question is supposed to be a reason for an action and not for an attitude-state, such examples cannot be rejected on the assumption that the outweighing reasons would not be provided by the objects of the attitude-states involved. Second, there might even be exceptional cases in which the reason is not present at all. It is a necessary constitutive truth about intentions that they involve

---

<sup>53</sup> In order for this reason to be evidence-based, we need to assume that the intention-states involved are rationally accessible. Since the claim that means/end-incoherence is irrational must itself be restricted to rationally accessible states (as I have argued above), the exclusion of inaccessible states does not pose a serious problem here, however.

dispositions to engage in certain mental activities and take further means. Evidence that a certain means is necessary for an end is therefore usually evidence that one risks engaging in pointless activity if one intends the end but not the means. But nothing guarantees that this is so in each particular case. In exceptional cases, one might have evidence that one will engage in the pointless activity anyway, independently of whether one maintains one's incoherent condition. Or one might have sufficient evidence that a particular intention will dispose one to engage only in activities that one has sufficient reason to engage in anyway, independently of whether one will successfully carry out one's intention.

On an account that ties rationality to responding to reasons, granting such exceptions means accepting that means/end-incoherence is not always irrational. As I see it, this is not an objection against the account I am proposing; exceptional cases only show that the phenomenon in question has to be interpreted in a different, and perhaps less unified, way than in terms of structural requirements. According to the present account, means/end-incoherence is a strong indicator of irrationality in the vast majority of cases. As I have shown, it is, first, necessarily irrational to have means/end-incoherent attitudes in determined cases. Second, means/end-incoherence is irrational when, in a previously underdetermined case, one has made further investments or one's adopting an intention can itself count as a significant investment. Third, means/end-incoherence is irrational if one's intention is accompanied by a belief that one ought to perform the action in question or by a belief that one will perform it. Fourth, means/end-incoherence is irrational in all cases in which maintaining the incoherent state raises the evidential probability that one will engage in pointless activity and one has no strong reason against making a decision. But I accept that if one has neither conclusive reason for nor against an intended action, and one has not made any significant investment, and one does not believe either that one ought to perform the action or that one will perform the action, and one has strong independent reason not to make a decision, it may be on balance more rational to retain a group of attitude-states despite the fact that together they exhibit a form of incoherence.

In fact, there are independent reasons for thinking that means/end-incoherence and intention inconsistency are not always irrational, for there are practical analogues

to the preface and lottery cases discussed above in the context of theoretical rationality.<sup>54</sup> Consider the following case, which is inspired by an example by Michael Bratman.<sup>55</sup> You are simultaneously playing two games of correspondence chess with one and the same person. For each of the games, it is true that you intend to win it. You know, however, that your game partner is a sore loser, and you are certain that once you have won the first game, he will not continue to play the other game. So you know that a necessary means to winning either of the games is not winning the other one first. Hence, if you intend to win each of the games, then you positively intend not to take the known necessary means to an end you intend. Intuitively, however, it does not seem irrational to play each of the games with the intention to win.<sup>56</sup> So intuitively, the claim that means/end-incoherence is irrational has exceptions.

Moreover, the proposal I have made can explain these exceptions. In the chess example, maintaining your inconsistent intentions is harmless; the activities you are disposed towards in virtue of your intention to win are the same activities you will engage in anyway if you play both games, which you have good reason to do. Even if the risk of taking means to winning a game that will eventually not be brought to a close might provide some reason for making a decision between the two games, it is not difficult to see how this reason can be outweighed by the prospects of playing both games with the intention to win.

To sum up, I have discussed a number of considerations that help to explain why means/end-incoherence is irrational. Although these considerations entail that almost all practical cases of means/end-incoherence are cases of irrationality understood

---

<sup>54</sup> See also Kolodny (2008a, §4) and McCann (1991). Bratman agrees that means/end-incoherence is not necessarily irrational. In Bratman (1987, 32), he maintains that the demands for intention consistency and means/end-coherence are “defeasible: there may be special circumstances in which it is rational of an agent to violate them”. And in Bratman (2009a, 413), he holds that means/end-incoherence is only “pro tanto irrational”. I think that the cases discussed might even show that means/end-incoherence is only *prima facie* irrational.

<sup>55</sup> Cf. Bratman (1987, 113–114). I owe thanks to Jan Gertken as this example evolved out of a discussion with him about Bratman’s video game case.

<sup>56</sup> I take it to be trivially true that if A plays the game with the intention to win, then A intends to win. This is not to be confused with what Bratman calls the “simple view”, namely the thesis that A’s intentionally  $\phi$ -ing implies that A intends to  $\phi$ . The difference is important since Bratman’s video game example is designed to show that the simple view is incompatible with the claim that inconsistent intentions are always irrational. My argument to the effect that inconsistent intentions are not always irrational relies on a structurally similar example, but it does not presuppose the simple view.

in terms of requirements that are grounded in evidence-based reasons, exceptions remain possible. But I have also argued that there are independent reasons to think that such exceptions really exist. I conclude that the considerations adduced here explain as much as there is to explain about instrumental rationality.

Philosophers (but not only philosophers) often take for granted the idea of a structural requirement of instrumental rationality – or a “hypothetical imperative”, as it is sometimes called following Kant’s terminology. To give an example, Robert Nozick holds that “the instrumental theory of rationality does not seem to stand in need of justification”. According to him, it is “the default theory, the theory that all discussants of rationality can take for granted, whatever else they think”.<sup>57</sup> Many assume that the instrumental requirement is normative, or even has “ground-level normative status” as James Dreier suggests.<sup>58</sup> Some even maintain that the normative authority of any requirement must somehow be based in or explained in terms of the instrumental requirement. This seems to be Phillipa Foot’s view when she claims that morality provides us with reasons only if it is understood as what she calls a “system of hypothetical imperatives”.<sup>59</sup> Starting from the assumption that normativity is a matter of reasons, I have argued in this thesis that the normativity of instrumental rationality is all but obvious, and that the idea of a structural requirement of means/end-coherence is quite problematic. But if the argument of this chapter is on the right track, then we can accept this conclusion and still hold on to the plausible claim that at least apart from exceptional cases, means/end-incoherence is irrational. Such irrationality can be explained in terms of requirements of reasons that are neither conditional on intentions, nor concerned with means/end-coherence as such – “categorical imperatives”, if you want to call them that. In this way, this chapter has provided the opposite of what Foot aimed to do: an explanation of the normativity of instrumental rationality in terms of categorical imperatives rather than an explanation of the normativity of anything else in terms of hypothetical imperatives.

---

<sup>57</sup> Nozick (1993, 133).

<sup>58</sup> Dreier (2001, 42).

<sup>59</sup> Foot (1972).

## 10.6 Conclusion

This concludes my account of structural irrationality. I have argued, in this chapter and the last, that structural irrationality claims such as (A)-(D) can be explained without recourse to structural requirements of rationality, by adherence to rational requirements to respond to available reasons alone. This enables us to respond to the problem that I described in Chapter 1 by adopting the *anti-structuralist view*, which rejects the assumption of structural requirements of rationality, and is thus capable of embracing the normativity of rationality while accepting the arguments to the effect that structural requirements of rationality cannot be normative.

I have motivated this anti-structuralist view independently. In Chapter 2, I argued that the denial of the normativity of rationality involves a commitment to an unattractive revisionary theory of ordinary attributions of irrationality. In Chapters 4 and 5, I showed that structural requirements of rationality cannot be normative. Chapter 6 presented arguments against the existence of structural requirements of rationality that hold even regardless of their normativity.

In Chapter 7, I began to take up one of the two main challenges for the anti-structuralist view, by developing a positive account of non-structural requirements of rationality that is compatible with the normativity of rationality. On this account, rational requirements can be understood as requirements to respond to available reasons. Finally, in Chapters 9 and 10 I have taken up the other main challenge for the anti-structuralist view, by delineating how structural irrationality can be explained in terms of such response requirements. If this explanation succeeds, the account I provide can explain structural irrationality by recourse to rational requirements that are normative in the sense that they are necessarily supported by the preponderance of available reasons. If what I have argued in Chapter 8 is correct, then these are all the reasons that there are. Hence, the account explains structural irrationality by recourse to rational requirements that are normative in the unrestricted sense that they are necessarily supported by the preponderance of reasons, *sans phrase*. This allows us to answer the main question of this thesis, without qualification, in the affirmative: Yes, we should indeed be rational.

No doubt, my account of the normativity of rationality invites a number of questions that need more careful treatment than I have given them here, or that have not even been dealt with at all – in particular, questions regarding the viability of the reason response conception of rationality as well as other assumptions on which the envisaged explanation of structural irrationality rested. The anti-structuralist approach to rationality is still in its infancy, and the best that I can hope for is that this thesis has taken some steps towards its coming of age.



## Bibliography

---

- Adams, Robert Merrihew. 1985. "Involuntary Sins." *The Philosophical Review* 94 (1): 3–31.
- Alvarez, Maria. 2010. *Kinds of Reasons. An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- Andrić, Vuko. 2013a. "Objective Consequentialism and the Licensing Dilemma." *Philosophical Studies* 162: 547–566.
- . 2013b. "The Case of the Miners." *Journal of Ethics and Social Philosophy* (Discussion Note): 1–8.
- Anscombe, G.E.M. 1957. *Intention*. 2nd ed. Cambridge, MA: Harvard University Press (repr. 2000).
- Aristoteles. *Politik*. Edited by Franz F. Schwarz. Stuttgart: Reclam (1989).
- . *Nikomachische Ethik*. Edited by Ursula Wolf. Reinbek bei Hamburg: Rowohlt (2006).
- . Magna Moralia. In *Aristoteles. Werke in Deutscher Übersetzung*, edited by Ernst Grumach. Vol. 8. Berlin: Akademie (1958).
- Arpaly, Nomy. 2000. "On Acting Rationally Against One's Best Judgment." *Ethics* 110 (3): 488–513.
- Baier, Kurt. 1978. "Moral Reasons and Reasons to Be Moral." In *Values and Morals*, edited by Alvin I. Goldman and Jaegwon Kim, 231–256. Dordrecht: Reidel.
- Bedke, Matthew S. 2009. "The Iffiest Oughts: A Guise of Reasons Account of End-Given Conditionals." *Ethics* 119 (4): 672–698.
- Bittner, Rüdiger. 1980. "Hypothetische Imperative." *Zeitschrift für Philosophische Forschung* 34 (2): 210–226.
- . 2001. *Doing Things for Reasons*. Oxford: Oxford University Press.
- Björnsson, Gunnar, and Stephen Finlay. 2010. "Metaethical Contextualism Defended." *Ethics* 121 (1): 7–36.
- Bradley, Raymond, and Norman Swartz. 1979. *Possible Worlds*. Indianapolis: Hackett.
- Bratman, Michael E. 1981. "Intention and Means-End Reasoning." *The Philosophical Review* 90 (2): 252–265.
- . 1987. *Intention, Plans, and Practical Reason*. Stanford: CSLI (repr. 1999).
- . 1991. "Cognitivism About Practical Reason." *Ethics* 102 (1): 117–128.

- . 2001. "Two Problems About Human Agency." Reprinted in: *Structures of Agency*, 89–105. Oxford: Oxford University Press (2007).
- . 2009a. "Intention, Practical Rationality, and Self-Governance." *Ethics* 119 (3): 411–443.
- . 2009b. "Intention Rationality." *Philosophical Explorations* 12 (3): 227–241.
- . 2009c. "Intention, Belief, Practical, Theoretical." In *Spheres of Reason. New Essays in the Philosophy of Normativity*, edited by Simon Robertson, 29–61. Oxford: Oxford University Press.
- . 2009d. "Intention, Belief, and Instrumental Rationality." In *Reasons for Action*, edited by David Sobel and Steven Wall, 13–36. Cambridge: Cambridge University Press.
- Bridges, Jason. 2009. "Rationality, Normativity, and Transparency." *Mind* 118 (470): 353–367.
- Broome, John. 1991. *Weighing Goods*. Oxford: Blackwell.
- . 1999. "Normative Requirements." *Ratio* 12 (4): 398–419.
- . 2001a. "Are Intentions Reasons? And How Should We Cope with Incommensurable Values?" In *Practical Rationality and Preference. Essays for David Gauthier*, edited by Christopher W. Morris and Arthur Ripstein, 98–120. Cambridge: Cambridge University Press.
- . 2001b. "Normative Practical Reasoning." *Proceedings of the Aristotelian Society, Supplementary Volumes* 75: 175–193.
- . 2002. "Practical Reasoning." In *Reason and Nature: Essays in the Theory of Rationality*, edited by José Luis Bermúdez and Alan Millar, 85–111. Oxford: Clarendon Press.
- . 2004a. "Reasons." In *Reason and Value. Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith, 28–55. Oxford: Oxford University Press.
- . 2004b. *Weighing Lives*. Oxford: Oxford University Press.
- . 2005a. "Does Rationality Give Us Reasons?" *Philosophical Issues* 15 (1): 321–337.
- . 2005b. "Have We Reason To Do As Rationality Requires? - A Comment on Raz." *Journal of Ethics and Social Philosophy* 1 (1): 1–8.
- . 2006. "Reasoning with Preferences?" *Royal Institute of Philosophy Supplement* 59: 183–208.
- . 2007a. "Is Rationality Normative?" *Disputatio* 2 (23): 161–178.
- . 2007b. "Requirements." In *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*, edited by Toni Rønnow-Rasmussen, Björn Petersson, Jonas Josefsson, and Dan Egonsson, 1–41.

- URL = <http://www.fil.lu.se/hommageawlodek/>.
- . 2007c. “Wide or Narrow Scope?” *Mind* 116 (462): 359–370.
- . 2007d. “Does Rationality Consist in Responding Correctly to Reasons?” *Journal of Moral Philosophy* 4 (3): 349–374.
- . 2008. “Reply to Southwood, Kearns and Star, and Cullity.” *Ethics* 119 (1): 96–108.
- . 2009. “The Unity of Reasoning?” In *Spheres of Reason. New Essays in the Philosophy of Normativity*, edited by Simon Robertson, 62–92. Oxford: Oxford University Press.
- . 2010. “Rationality.” In *A Companion to the Philosophy of Action*, edited by Timothy O’Connor and Constantine Sandis, 285–292. Chichester: Wiley-Blackwell.
- . 2013. *Rationality Through Reasoning*. Chichester: Wiley-Blackwell.
- Brunero, John. 2005. “Two Approaches to Instrumental Rationality and Belief Consistency.” *Journal of Ethics and Social Philosophy* 1 (1): 1–20.
- . 2009a. “Against Cognitivism About Practical Rationality.” *Philosophical Studies* 146 (3): 311–325.
- . 2009b. “Reasons and Evidence One Ought.” *Ethics* 119 (3): 538–545.
- . 2010a. “The Scope of Rational Requirements.” *The Philosophical Quarterly* 60 (238): 28–49.
- . 2010b. “Self-Governance, Means-Ends Coherence, and Unalterable Ends.” *Ethics* 120 (3): 579–591.
- . 2012. “Instrumental Rationality, Symmetry and Scope.” *Philosophical Studies* 157 (1): 125–140.
- Bykvist, Krister. 2011. “How to Do Wrong Knowingly and Get Away with It.” In *Neither/Nor. Philosophical Papers Dedicated to Erik Carlson on the Occasion of His Fiftieth Birthday*, 31–47. Uppsala Philosophical Studies 58. Uppsala: Department of Philosophy, Uppsala University.
- Capes, Justin A. 2012. “Blameworthiness Without Wrongdoing.” *Pacific Philosophical Quarterly* 93 (3): 417–437.
- Chang, Ruth, ed. 1997. *Incommensurability, Incomparability, and Practical Reason*. Cambridge, MA: Harvard University Press.
- . 2009. “Voluntarist Reasons and the Sources of Normativity.” In *Reasons for Action*, edited by David Sobel and Steven Wall, 243–271. Cambridge: Cambridge University Press.
- Cheng-Guajardo, Luis. “The Normative Requirement of Means-End Rationality and Modest Bootstrapping.” Forthcoming in: *Ethical Theory and Moral Practice*: 1–17. doi:10.1007/s10677-013-9455-y.

- Chisholm, Roderick M. 1963. "Contrary-to-duty Imperatives and Deontic Logic." *Analysis* 24 (2): 33–36.
- . 1964. "The Ethics of Requirement." *American Philosophical Quarterly* 1 (2): 147–153.
- Coates, Allen. 2013. "The Enkratic Requirement." *European Journal of Philosophy* 21 (2): 320–333.
- Conee, Earl. 2004. "First Things First." In *Evidentialism. Essays in Epistemology*, by Earl Conee and Richard Feldman. Oxford: Clarendon Press.
- Coons, Christian, and David Faraci. 2010. "First-Personal Authority and the Normativity of Rationality." *Philosophia* 38 (4): 733–740.
- Cullity, Garrett. 2008. "Decisions, Reasons, and Rationality." *Ethics* 119 (1): 57–95.
- Dancy, Jonathan. 1977. "The Logical Conscience." *Analysis* 37 (2): 81–84.
- . 2000a. *Practical Reality*. Oxford: Oxford University Press (rev. ed. 2002).
- . 2000b. "Should We Pass the Buck?" *Royal Institute of Philosophy Supplement* 47 (1): 159–173.
- . 2002. "Prichard on Duty and Ignorance of Fact." In *Ethical Intuitionism: Re-evaluations*, edited by Philip Stratton-Lake, 229–247. Oxford: Clarendon Press.
- . 2004a. *Ethics Without Principles*. Oxford: Oxford University Press.
- . 2004b. "Enticing Reasons." In *Reason and Value. Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith, 91–118. Oxford: Oxford University Press.
- . 2009. "Reasons and Rationality." In *Spheres of Reason. New Essays in the Philosophy of Normativity*, edited by Simon Robertson, 93–112. Oxford: Oxford University Press.
- Darwall, Stephen L. 1983. *Impartial Reason*. Ithaca: Cornell University Press (repr. 1985).
- Davidson, Donald. 1963. "Actions, Reasons, and Causes." Reprinted in: *Essays on Actions and Events*, 2nd ed., 3–19. Oxford: Clarendon Press (2001).
- . 1969. "How Is Weakness of the Will Possible?" Reprinted in: *Essays on Actions and Events*, 2nd ed., 21–42. Oxford: Clarendon Press (2001).
- . 1971. "Agency." Reprinted in: *Essays on Actions and Events*, 2nd ed., 43–61. Oxford: Clarendon Press (2001).
- . 1978. "Intending." Reprinted in: *Essays on Actions and Events*, 2nd ed., 83–102. Oxford: Clarendon Press (2001).
- . 1982. "Paradoxes of Irrationality." Reprinted in: *Problems of Rationality*, 169–187. Oxford: Oxford University Press (2004).
- . 1985. "Incoherence and Irrationality." Reprinted in: *Problems of Rationality*, 189–198. Oxford: Oxford University Press (2004).

- De Lazari-Radek, Katarzyna, and Peter Singer. 2010. "Secrecy in Consequentialism: A Defence of Esoteric Morality." *Ratio* 23 (1): 34–58.
- Dorsey, Dale. 2012. "Objective Morality, Subjective Morality, and the Explanatory Question." *Journal of Ethics and Social Philosophy* 6 (3): 1–24.
- Dreier, James. 2001. "Humean Doubts About Categorical Imperatives." In *Varieties of Practical Reasoning*, edited by Elijah Millgram, 27–47. Cambridge, MA: MIT Press.
- . 2009. "Practical Conditionals." In *Reasons for Action*, edited by David Sobel and Steven Wall, 116–133. Cambridge: Cambridge University Press.
- Dretske, Fred I. 1970. "Epistemic Operators." *The Journal of Philosophy* 67 (24): 1007–1023.
- Enoch, David. 2006. "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action." *The Philosophical Review* 115 (2): 169.
- Evers, Daan. 2011. "Two Objections to Wide-scoping." *Grazer Philosophische Studien* 83 (1): 251–255.
- Ewing, A.C. 1947. *The Definition of Good*. New York: Macmillan.
- . 1953. *Ethics*. London: English Universities Press.
- Feldman, Fred. 2012. "True and Useful: On the Structure of a Two Level Normative Theory." *Utilitas* 24 (2): 151–171.
- Feldman, Richard. 1988a. "Having Evidence." Reprinted in: *Evidentialism. Essays in Epistemology*, edited by Earl Conee and Richard Feldman, 219–241. Oxford: Clarendon Press (2004).
- . 1988b. "Subjective and Objective Justification in Ethics and Epistemology." *The Monist* 71 (3): 405–419.
- Fink, Julian. 2010. "Asymmetry, Scope, and Rational Consistency." *Croatian Journal of Philosophy* 10 (29): 109–130.
- . 2011. "Are There Process-Requirements of Rationality?" *Organon F* 18 (4): 475–487.
- . 2012. "The Function of Normative Process-Requirements." *Dialectica* 66 (1): 115–136.
- Finlay, Stephen. 2006. "The Reasons That Matter." *Australasian Journal of Philosophy* 84 (1): 1–20.
- . 2009. "Against All Reason? Skepticism About the Instrumental Norm." In *Hume on Motivation and Virtue*, edited by Charles R. Pigden, 155–178. Basingstoke: Palgrave Macmillan.
- . 2010. "What *Ought* Probably Means, and Why You Can't Detach It." *Synthese* 177 (1): 67–89.

- Fitelson, Branden. 2010. "Pollock on Probability in Epistemology." *Philosophical Studies* 148 (3): 455–465.
- . 2011. "Evidence of Evidence Is Not (Necessarily) Evidence." *Analysis* 72 (1): 85–88.
- Føllesdal, Dagfinn, and Risto Hilpinen. 1971. "Deontic Logic: An Introduction." In *Deontic Logic: Introductory and Systematic Readings*, edited by Risto Hilpinen, 1–35. Dordrecht: Pallas Edition (repr. 1981).
- Foot, Philippa. 1972. "Morality as a System of Hypothetical Imperatives." *The Philosophical Review* 81 (3): 305–316.
- . 1978. "Are Moral Considerations Overriding?" Reprinted in: *Virtues and Vices*, 181–188. Oxford: Clarendon Press (2002).
- . 2001. *Natural Goodness*. Oxford: Clarendon Press.
- Forrester, James William. 1984. "Gentle Murder, or the Adverbial Samaritan." *The Journal of Philosophy* 81 (4): 193–197.
- Frankfurt, Harry G. 1987. "Identification and Wholeheartedness." Reprinted in: *The Importance of What We Care About*, 159–176. Cambridge: Cambridge University Press (2007).
- Gauthier, David P. 1986. *Morals by Agreement*. Oxford: Clarendon Press.
- Geach, P. T. 1977. *The Virtues*. Cambridge: Cambridge University Press.
- . 1982. "Whatever Happened to Deontic Logic?" *Philosophia* 11 (1): 1–12.
- Gensler, Harry J. 1985. "Ethical Consistency Principles." *The Philosophical Quarterly* 35 (139): 156–170.
- Gert, Bernard. 2005. *Morality. Its Nature and Justification*. Oxford: Oxford University Press (rev. ed.).
- Gettier, Edmund L. 1963. "Is Justified True Belief Knowledge?" *Analysis*: 121–123.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings. A Theory of Normative Judgment*. Oxford: Clarendon Paperbacks (repr. 2002).
- . 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
- . 2005. "Truth and Correct Belief." *Philosophical Issues* 15 (1): 338–350.
- Gibbons, John. 2006. "Access Externalism." *Mind* 115 (457): 19–39.
- . 2009. "You Gotta Do What You Gotta Do." *Noûs* 43 (1): 157–177.
- . 2010a. "Missing the Obvious: Reply to Moon." *Mind* 119 (473): 153–158.
- . 2010b. "Things That Make Things Reasonable." *Philosophy and Phenomenological Research* 81 (2): 335–361.
- Graham, Peter A. 2010. "In Defense of Objectivism About Moral Obligation." *Ethics* 121 (1): 88–115.

- . “A Sketch of a Theory of Moral Blameworthiness.” Forthcoming in: *Philosophy and Phenomenological Research*: 1–22. doi:10.1111/j.1933-1592.2012.00608.x.
- Greenspan, P. S. 1975. “Conditional Oughts and Hypothetical Imperatives.” *The Journal of Philosophy* 72 (10): 259–276.
- Hailperin, Theodore. 1996. *Sentential Probability Logic. Origins, Development, Current Status, and Technical Applications*. Bethlehem: Lehigh University Press.
- Hájek, Alan. 2002. “Interpretations of Probability.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, 1–71 (rev. ed. 2011). URL = <http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>.
- Hampton, Jean E. 1998. *The Authority of Reason*. Cambridge: Cambridge University Press.
- Harman, Gilbert. 1973. “Review of Roger Wertheimer, The Significance of Sense: Meaning, Modality, and Morality.” *The Philosophical Review* 82 (2): 235–239.
- . 1976. “Practical Reasoning.” *The Review of Metaphysics* 29 (3): 431–463.
- . 1986. *Change in View. Principles of Reasoning*. Cambridge, MA: MIT Press.
- Henning, Tim. “Normative Reasons Contextualism.” Forthcoming in: *Philosophy and Phenomenological Research*: 1–32. doi:10.1111/j.1933-1592.2012.00645.x.
- Heuer, Ulrike. 2010. “Reasons and Impossibility.” *Philosophical Studies* 147: 235–246.
- Hill, Thomas E. 1973. “The Hypothetical Imperative.” *The Philosophical Review* 82 (4): 429–450.
- Horty, John. 2001. *Agency and Deontic Logic*. Oxford: Oxford University Press.
- Howard-Snyder, Frances. 2005. “It’s the Thought That Counts.” *Utilitas* 17 (3): 265–281.
- Huemer, Michael. 2001. *Skepticism and the Veil of Perception*. Lanham: Rowman & Littlefield.
- . 2006. “Phenomenal Conservatism and the Internalist Intuition.” *American Philosophical Quarterly* 43 (2): 147–158.
- . 2007. “Compassionate Phenomenal Conservatism.” *Philosophy and Phenomenological Research* 74 (1): 30–55.
- Husi, Stan. 2013. “Why Reasons Skepticism Is Not Self-Defeating.” *European Journal of Philosophy* 21 (3): 424–449.
- Hussain, Nadeem J.Z. “The Requirements of Rationality”. Unpublished manuscript, Stanford University, August 2007, version 2.4.
- Jackson, Frank. 1991. “Decision-theoretic Consequentialism and the Nearest and Dearest Objection.” *Ethics* 101 (3): 461–482.
- Jackson, Frank, and Robert Pargetter. 1986. “Oughts, Options, and Actualism.” *The Philosophical Review* 95 (2): 233–255.

- Jackson, Frank, and Michael Smith. 2006. "Absolutist Moral Theories and Uncertainty." *The Journal of Philosophy* 103 (6): 267–283.
- Joyce, Richard. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.
- Kamm, Frances M. 2000. "The Doctrine of Triple Effect and Why a Rational Agent Need Not Intend the Means to His End." *Proceedings of the Aristotelian Society, Supplementary Volumes* 74 (1): 21–39.
- Kant, Immanuel. 1785. *Grundlegung zur Metaphysik der Sitten*. In *Immanuel Kant Werkausgabe*, edited by Wilhelm Weischedel. Vol. VII. Frankfurt a.M.: Suhrkamp (1974).
- Kavka, Gregory S. 1983. "The Toxin Puzzle." *Analysis* 43 (1): 33–36.
- Kearns, Stephen, and Daniel Star. 2008. "Reasons: Explanations or Evidence?" *Ethics* 119 (1): 31–56.
- . 2009. "Reasons as Evidence." *Oxford Studies in Metaethics* 4: 215–42.
- . 2013. "Weighing Reasons." *Journal of Moral Philosophy* 10 (1): 70–86.
- Kelly, Thomas. 2006. "Evidence." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, 1–58.  
URL = <http://plato.stanford.edu/archives/fall2008/entries/evidence/>.
- Kiesewetter, Benjamin. 2011a. "Willensschwäche." In *Aristoteles-Handbuch*, edited by Christof Rapp and Klaus Corcilius, 389–393. Stuttgart: Metzler.
- . 2011b. "'Ought' and the Perspective of the Agent." *Journal of Ethics and Social Philosophy* 5 (3): 1–24.
- . 2012. "A Dilemma for Parfit's Conception of Normativity." *Analysis* 72 (3): 466–474.
- . "Instrumental Normativity". Unpublished manuscript, Humboldt-Universität zu Berlin, December 2012.
- . "Can the Lottery Paradox Be Solved by Identifying Epistemic Justification with Epistemic Permissibility?" Unpublished manuscript, Humboldt-Universität zu Berlin, May 2013.
- Kolodny, Niko. 2005. "Why Be Rational?" *Mind* 114 (455): 509–563.
- . 2007a. "How Does Coherence Matter?" *Proceedings of the Aristotelian Society* 107 (3): 229–263.
- . 2007b. "State or Process Requirements?" *Mind* 116 (462): 371–385.
- . 2008a. "The Myth of Practical Consistency." *European Journal of Philosophy* 16 (3): 366–402.
- . 2008b. "Why Be Disposed to Be Coherent?" *Ethics* 118 (3): 437–463.
- . 2009. "Reply to Bridges." *Mind* 118 (470): 369.

- . 2011a. "Aims as Reasons." In *Reasons and Recognition. Essays on the Philosophy of T.M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman, 43–78. Oxford: Oxford University Press.
- . 2011b. "Raz's Nexus." *Jurisprudence* 2 (2): 333–351.
- . "Instrumental Reasons." Forthcoming in: *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star.
- . "Lectures on Reasons and Rationality". Unpublished manuscript, University of California, Berkeley, January 2011.
- Kolodny, Niko, and John MacFarlane. 2010. "Ifs and Oughts." *The Journal of Philosophy* 107 (3): 115–143.
- . "Ought: Between Objective and Subjective". Unpublished manuscript, University of California, Berkeley, September 2009.
- Kornblith, Hilary. 2002. *Knowledge And Its Place In Nature*. Oxford: Oxford University Press.
- Korsgaard, Christine M. 1986. "Skepticism About Practical Reason." *The Journal of Philosophy* 83 (1): 5–25.
- . 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- . 1997. "The Normativity of Instrumental Reason." In *Ethics and Practical Reason*, edited by Garrett Cullity and Berys Gaut, 215–254. Oxford: Clarendon Press.
- . 2009a. "The Activity of Reason." *Proceedings and Addresses of the APA* 83 (2): 23–43.
- . 2009b. *Self-Constitution. Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Kroedel, Thomas. 2012. "The Lottery Paradox, Epistemic Justification and Permissibility." *Analysis* 72 (1): 57–60.
- Kyburg, Henry E. 1961. *Probability and the Logic of Rational Belief*. Middletown: Wesleyan University Press.
- Littlejohn, Clayton. 2012. "Lotteries, Probabilities, and Permissions." *Logos & Episteme* 3 (3): 509–514.
- Lord, Errol. 2010. "Having Reasons and the Factoring Account." *Philosophical Studies* 149 (3): 283–296.
- . 2011. "Violating Requirements, Exiting from Requirements, and the Scope of Rationality." *The Philosophical Quarterly* 61 (243): 392–399.
- Makinson, D. C. 1965. "The Paradox of the Preface." *Analysis* 25 (6): 205–207.
- Mason, Elinor. 2013. "Objectivism and Prospectivism About Rightness." *Journal of Ethics and Social Philosophy* 7 (2): 1–21.

- McCann, H. J. 1991. "Settled Objectives and Rational Constraints." *American Philosophical Quarterly* 28 (1): 25–36.
- McDowell, John. 1995. "Might There Be External Reasons?" Reprinted in: *Mind, Value, and Reality*, 95–111. Cambridge, MA: Harvard University Press (1998).
- Mele, Alfred R. 1987. *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. Oxford: Oxford University Press.
- Moore, G.E. 1912. *Ethics*. Oxford: Oxford University Press (repr. 1969).
- Moran, Richard. 1988. "Making Up Your Mind." *Ratio* 1 (2): 135–151.
- . 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- Nagel, Thomas. 1970. *The Possibility of Altruism*. Oxford: Clarendon Press.
- Nelson, Mark T. 2010. "We Have No Positive Epistemic Duties." *Mind* 119 (473): 83–102.
- Nielsen, Kai. 1984. "Why Should I Be Moral? Revisited." *American Philosophical Quarterly* 21 (1): 81–91.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Belknap Press.
- . 1993. *The Nature of Rationality*. Princeton: Princeton University Press (repr. 1995).
- Oddie, Graham, and Peter Menzies. 1992. "An Objectivist's Guide to Subjective Value." *Ethics* 102 (3): 512–533.
- Olson, Jonas. 2004. "Buck-Passing and the Wrong Kind of Reasons." *The Philosophical Quarterly* 54 (215): 295–300.
- Owens, David. 2012. *Shaping the Normative Landscape*. Oxford: Oxford University Press.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press (rev. ed. 1987).
- . 1997. "Reasons and Motivation." *Proceedings of the Aristotelian Society, Supplementary Volumes* 71: 99–130.
- . 2001. "Rationality and Reasons." In *Exploring Practical Philosophy: From Action to Values*, edited by Dan Egonsson, Jonas Josefsson, Björn Petersson, and Tonio Rønnow-Rasmussen, 17–39. Aldershot: Ashgate.
- . 2011a. *On What Matters*. Vol. 1. Oxford: Oxford University Press.
- . 2011b. *On What Matters*. Vol. 2. Oxford: Oxford University Press.
- Piller, Christian. 2006. "Content-Related and Attitude-Related Reasons for Preferences." *Royal Institute of Philosophy Supplements* 81 (59): 155–181.
- . 2007. "Ewing's Problem." *European Journal of Analytic Philosophy* 3 (1): 43–65.

- Price, A. W. 2008. "Against Requirements of Rationality." *Proceedings of the Aristotelian Society* 108 (2): 157–176.
- Prichard, H.A. 1912. "Does Moral Philosophy Rest on a Mistake?" Reprinted in: *Moral Writings*, 7–20. Oxford: Clarendon Press (2002).
- . 1932. "Duty and Ignorance of Fact." Reprinted in: *Moral Writings*, 84–101. Oxford: Clarendon Press (2002).
- . 1949. "Moral Obligation." Reprinted in: *Moral Writings*, 163–225. Oxford: Clarendon Press (2002).
- Putnam, Hilary. 1973. "Meaning and Reference." *The Journal of Philosophy* 70 (19): 699–711.
- Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen. 2004. "The Strike of the Demon: On Fitting Pro-attitudes and Value." *Ethics* 114 (3): 391–423.
- . 2006. "Buck-Passing and the Right Kind of Reasons." *The Philosophical Quarterly* 56 (222): 114–120.
- Raz, Joseph. 1975a. *Practical Reason and Norms*. 2nd ed. Oxford: Oxford University Press (repr. 1999).
- . 1975b. "Reasons for Action, Decisions and Norms." *Mind* 84 (1): 481–499.
- . 1990. "Rethinking Exclusionary Reasons." Postscript to the Second Edition of: *Practical Reason and Norms*, 178–199. Oxford: Oxford University Press (repr. 1999).
- . 1998. "Incommensurability and Agency." Reprinted in: *Engaging Reason. On the Theory of Value and Action*, 46–66. Oxford: Oxford University Press (1999).
- . 1999a. "Explaining Normativity: On Rationality and the Justification of Reason." Reprinted in: *Engaging Reason. On the Theory of Value and Action*, 67–89. Oxford: Oxford University Press (1999).
- . 1999b. *Engaging Reason. On the Theory of Value and Action*. Oxford: Oxford University Press.
- . 2005a. "The Myth of Instrumental Rationality." *Journal of Ethics and Social Philosophy* 1 (1): 2–28.
- . 2005b. "Instrumental Rationality: A Reprise." *Journal of Ethics and Social Philosophy* 1 (1): 1–19.
- . 2009a. "Reasons: Explanatory and Normative." Reprinted in: *From Normativity to Responsibility*, 13–35. Oxford: Oxford University Press (2011).
- . 2009b. "Reasons: Practical and Adaptive." Reprinted in: *From Normativity to Responsibility*, 36–58. Oxford: Oxford University Press (2011).
- . 2010a. "Reason, Reasons and Normativity." *Oxford Studies in Metaethics* 5: 5–24.

- . 2010b. “On the Guise of the Good.” Reprinted in: *From Normativity and Responsibility*, 59–84. Oxford: Oxford University Press (2011).
- . 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.
- Regan, Donald. 1980. *Utilitarianism and Co-Operation*. Oxford: Oxford University Press.
- Reisner, Andrew. 2009. “Unifying the Requirements of Rationality.” *Philosophical Explorations* 12: 243–260.
- . 2011. “Is There Reason to Be Theoretically Rational?” In *Reasons for Belief*, edited by Andrew Reisner and Asbjørn Steglich-Petersen, 34–53. Cambridge: Cambridge University Press.
- Ridge, Michael. 1998. “Humean Intentions.” *American Philosophical Quarterly* 35 (2): 157–178.
- Rippon, Simon. 2011. “In Defense of the Wide-Scope Instrumental Principle.” *Journal of Ethics and Social Philosophy* 5 (2): 1–21.
- Rivera-López, Eduardo. 2006. “Can There Be Full Excuses for Morally Wrong Actions?” *Philosophy and Phenomenological Research* 73 (1): 124–142.
- Robertson, Simon. 2011. “Epistemic Constraints on Practical Normativity.” *Synthese* 181 (1): 81–106.
- Rosen, Gideon. 2001. “Brandom on Modality, Normativity and Intentionality.” *Philosophy and Phenomenological Research* 63 (3): 611–623.
- Ross, Alf. 1941. “Imperatives and Logic.” *Theoria* 7 (1): 53–71.
- Ross, Jacob. 2006. *Acceptance and Practical Reason*. PhD dissertation, Rutgers University.
- . 2009. “How to Be a Cognitivist About Practical Reason.” *Oxford Studies in Metaethics* 4: 243–281.
- . 2010. “The Irreducibility of Personal Obligation.” *Journal of Philosophical Logic* 39 (3): 307–323.
- . 2012a. “Rationality, Normativity, and Commitment.” *Oxford Studies in Metaethics* 7: 138–181.
- . 2012b. “Actualism, Possibilism, and Beyond.” *Oxford Studies in Normative Ethics* 2: 74–96.
- Ross, W. David. 1930. *The Right and the Good*. Edited by Philip Stratton-Lake. Oxford University Press (repr. 2002).
- . 1939. *Foundations of Ethics*. Oxford: Clarendon Press (repr. 1968).
- Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press.
- . 2004. “Reasons: a Puzzling Duality?” In *Reason and Value. Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith, 231–246. Oxford: Clarendon Press.

- . 2007. "Structural Irrationality." In *Common Minds. Themes from the Philosophy of Philip Pettit*, edited by Geoffrey Brennan, Robert Goodin, Frank Jackson, and Michael Smith, 84–103. Oxford: Oxford University Press.
- . 2008. *Moral Dimensions. Permissibility, Meaning, Blame*. Cambridge, MA: Belknap Press.
- Schroeder, Mark. 2004. "The Scope of Instrumental Reason." *Philosophical Perspectives* 18 (1): 337–364.
- . 2005a. "The Hypothetical Imperative?" *Australasian Journal of Philosophy* 83 (3): 357–372.
- . 2005b. "Instrumental Mythology." *Journal of Ethics and Social Philosophy* 1 (1): 1–12.
- . 2007a. *Slaves of the Passions*. Oxford: Oxford University Press.
- . 2007b. "Reasons and Agent-neutrality." *Philosophical Studies* 135 (2): 279–306.
- . 2007c. "The Humean Theory of Reasons." *Oxford Studies in Metaethics* 2: 195–219.
- . 2007d. "Weighting for a Plausible Humean Theory of Reasons." *Noûs* 41 (1): 110–132.
- . 2008. "Having Reasons." *Philosophical Studies* 139 (1): 57–71.
- . 2009. "Means-End Coherence, Stringency, and Subjective Reasons." *Philosophical Studies* 143 (2): 223–248.
- . 2011a. "What Does It Take to 'Have' a Reason?" In *Reasons for Belief*, edited by Andrew Reisner and Asbjørn Steglich-Petersen, 201–222. Cambridge: Cambridge University Press.
- . 2011b. "Ought, Agents, and Actions." *The Philosophical Review* 120 (1): 1–41.
- . 2012. "The Ubiquity of State-Given Reasons." *Ethics* 122 (3): 457–488.
- . "What Makes Reasons Sufficient?" Unpublished manuscript, University of Southern California, August 2010.
- Searle, John R. 2001. *Rationality in Action*. Cambridge, MA: MIT Press.
- Setiya, Kieran. 2007a. "Cognitivism About Instrumental Reason." *Ethics* 117 (4): 649–673.
- . 2007b. *Reasons Without Rationalism*. Princeton: Princeton University Press.
- Shah, Nishi. 2003. "How Truth Governs Belief." *The Philosophical Review* 112 (4): 447–482.
- . 2008. "How Action Governs Intention." *Philosophers' Imprint* 8 (5): 1–19.
- Sher, George. 2006. *In Praise of Blame*. Oxford: Oxford University Press.
- Shpall, Sam. 2013. "Wide and Narrow Scope." *Philosophical Studies* 163 (3): 717–736.

- Sinhababu, Neil. "The Desire-Belief Account of Intention Explains Everything." Forthcoming in: *Noûs*: 1–19. doi:10.1111/j.1468-0068.2012.00864.x.
- Sinnott-Armstrong, Walter. 1987. "Moral Dilemmas and 'Ought and Ought Not'." *Canadian Journal of Philosophy* 17 (1): 127–139.
- Skorupski, John. 2007. "What Is Normativity?" *Disputatio* 2 (23): 247–69.
- . 2010. *The Domain of Reasons*. Oxford: Oxford University Press.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.
- . 1995. "Internal Reasons." *Philosophy and Phenomenological Research* 55 (1): 109–131.
- . 2004. "The Structure of Orthonomy." *Royal Institute of Philosophy Supplement* 55: 165–193.
- . 2006. "Moore on the Right, the Good, and Uncertainty." In *Metaethics after Moore*, edited by Terry Horgan and Mark Timmons, 133–148. Oxford: Oxford University Press.
- . 2008. "Consequentialism and the Nearest and Dearest Objection." In *Minds, Ethics, and Conditionals. Themes from the Philosophy of Frank Jackson*, edited by Ian Ravenscroft, 237–266. Oxford: Clarendon Press.
- . 2013. "The Ideal of Orthonomous Action, or the How and Why of Consistent Buck-Passing." In *Thinking About Reasons. Essays in the Honour of Jonathan Dancy*, edited by David Bakhurst, Brad Hooker, and Margaret Little, 50–75. Oxford: Oxford University Press.
- Southwood, Nicholas. 2008. "Vindicating the Normativity of Rationality." *Ethics* 119 (1): 9–30.
- Stocker, Michael. 1976. "The Schizophrenia of Modern Ethical Theories." *The Journal of Philosophy* 73 (14): 453–466.
- Stratton-Lake, Philip. 2000. *Kant, Duty and Moral Worth*. London: Routledge.
- . 2005. "How to Deal with Evil Demons: Comment on Rabinowicz and Rønnow-Rasmussen." *Ethics* 115 (4): 788–798.
- Strawson, P. F. 1962. "Freedom and Resentment." Reprinted in: *Freedom and Resentment and Other Essays*, 1–28. London: Routledge (2008).
- Street, Sharon. 2008. "Constructivism About Reasons." *Oxford Studies in Metaethics* 3: 207–245.
- Streumer, Bart. 2007. "Reasons and Impossibility." *Philosophical Studies* 136: 351–384.
- . 2010. "Reasons, Impossibility and Efficient Steps: Reply to Heuer." *Philosophical Studies* 151: 79–86.
- Svavarsdóttir, Sigrún. 2003. "Evaluations of Rationality." *The Southern Journal of Philosophy* 41: 121–136.

- Thompson, Michael. 2008. *Life and Action. Elementary Structures of Practice and Practical Thought*. Cambridge, MA: Harvard University Press.
- Thomson, Judith Jarvis. 1983. "Imposing Risks." Reprinted in: *Rights, Restitution, and Risk*, 173–191. Cambridge, MA: Harvard University Press (1986).
- . 1990. *The Realm of Rights*. Cambridge, MA: Harvard University Press.
- . 2008. *Normativity*. Chicago and La Salle: Open Court.
- Van Fraassen, B.C. 1973. "Values and the Heart's Command." *The Journal of Philosophy* 70 (1): 5–19.
- Velleman, J. David. 1989. *Practical Reflection*. Stanford: CSLI (repr. 2007).
- Vogelstein, Eric. 2012. "Subjective Reasons." *Ethical Theory and Moral Practice* 15 (2): 239–257.
- Vranas, Peter B. M. 2008. "New Foundations for Imperative Logic I: Logical Connectives, Consistency, and Quantifiers." *Noûs* 42 (4): 529–572.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- . 2001. "Normativity, Commitment, and Instrumental Reason." *Philosophers' Imprint* 1 (3): 1–26.
- . 2006. *Normativity and the Will. Selected Essays on Moral Psychology and Practical Reason*. Oxford: Oxford University Press.
- Watson, Gary. 1975. "Free Agency." *The Journal of Philosophy* 72 (8): 205–220.
- Way, Jonathan. 2009. "Two Accounts of the Normativity of Rationality." *Journal of Ethics and Social Philosophy* (Discussion Note): 1–8.
- . 2010a. "Defending the Wide-scope Approach to Instrumental Reason." *Philosophical Studies* 147: 213–233.
- . 2010b. "The Normativity of Rationality." *Philosophy Compass* 5 (12): 1057–1068.
- . 2011. "The Symmetry of Rational Requirements." *Philosophical Studies* 155: 227–239.
- . 2012a. "Transmission and the Wrong Kind of Reason." *Ethics* 122 (3): 489–515.
- . 2012b. "Explaining the Instrumental Principle." *Australasian Journal of Philosophy* 90 (3): 487–506.
- Wedgwood, Ralph. 2002. "Internalism Explained." *Philosophy and Phenomenological Research* 65 (2): 349–369.
- . 2003. "Choosing Rationally and Choosing Correctly." In *Weakness of Will and Practical Irrationality*, edited by Sarah Stroud and Christine Tappolet, 201–229. Oxford: Oxford University Press.

- . 2006. "The Meaning of 'Ought'." *Oxford Studies in Metaethics* 1: 127–160.
- . 2007. *The Nature of Normativity*. Oxford: Oxford University Press.
- White, Roger. 2005. "Epistemic Permissiveness." *Philosophical Perspectives* 19 (1): 445–459.
- Williams, Bernard. 1965. "Ethical Consistency." *Proceedings of the Aristotelian Society, Supplementary Volumes* 39: 103–138.
- . 1979. "Internal and External Reasons." Reprinted in: *Moral Luck*, 101–113. Cambridge: Cambridge University Press (1981).
- . 1980. "Ought and Moral Obligation." Reprinted in: *Moral Luck*, 114–123. Cambridge: Cambridge University Press (1981).
- . 1989. "Internal Reasons and the Obscurity of Blame." Reprinted in: *Making Sense of Humanity*, 35–45. Cambridge: Cambridge University Press (1995).
- . 1993. "Nietzsche's Minimalist Moral Psychology." Reprinted in: *Making Sense of Humanity*, 65–76. Cambridge: Cambridge University Press (1995).
- . 1995. "Replies." In *World, Mind, and Ethics. Essays on the Ethical Philosophy of Bernard Williams*, edited by J.E.J. Altham and Ross Harrison, 185–224. Cambridge: Cambridge University Press.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Zimmerman, Michael J. 1996. *The Concept of Moral Obligation*. Cambridge: Cambridge University Press.
- . 2008. *Living with Uncertainty. The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.