



# The Real Myth of Coherence

Wooram Lee<sup>1</sup>

Received: 5 July 2019 / Accepted: 2 March 2020 / Published online: 10 March 2020  
© The Author(s) 2020

## Abstract

In this paper, I offer a novel view of the coherence (or structural) requirements on belief and intention, according to which they are not norms, but rather principles describing how your belief and intention operate. I first argue, on the basis of the unintelligibility of some relevant attitudes-reports, that there are conditions under which you simply do not count as believing or intending unless your beliefs and intentions satisfy the requirements: the conditions under which all of your relevant attitudes are occurrent or activated. I then argue that you are subject to a coherence requirement only if your relevant attitudes are all activated, for you are not necessarily subject to the charge of irrationality in violating a coherence requirement when your attitudes are not all activated. If so, however, you satisfy the coherence requirements whenever you are subject to them, which makes it plausible that the “requirements” should be seen as descriptive principles about belief and intention.

This essay develops a novel view of the nature of so-called *coherence* (or *structural*) requirements on belief and intention, such as the requirement not to hold contradictory beliefs or intentions, the requirement to believe what follows by *modus ponens* from what you believe, the requirement to intend what you believe to be a necessary means to the end you intend. I argue that the idea that there are *normative* requirements of coherence is a myth. In saying this, I mean to contrast ‘normative’ with ‘descriptive’. My view is that the requirements turn out inviolable, once we articulate the conditions under which you are subject to them and reflect on what is involved in believing and intending. They should rather be seen as constitutive principles which *describe* the ways in which your beliefs and intentions operate under certain conditions, or the conditions under which beliefs and intentions are intelligibly attributed to you.

The view consists of two theses. The first is a thesis about what is distinctive to belief and intention: your beliefs and intentions always satisfy the coherence

---

✉ Wooram Lee  
woo-ram.lee@uni-due.de

<sup>1</sup> Institute for Philosophy, Faculty of Humanities, Duisburg-Essen University, Universitätsstr. 12, 45141 Essen, Germany

requirements when they are *occurrent* or *activated* in your thought or action. I argue for this thesis on the basis of the unintelligibility of attributing an incoherent combination of occurrent beliefs and intentions to an agent. The second thesis concerns the proper target of the requirements: you are subject to the coherence requirements only if the relevant attitudes are all occurrent. The argument for this thesis, roughly, is that you would be criticizably irrational for having an incoherent combination of attitudes only if the relevant attitudes were occurrent. It follows from the two theses that the requirements are inviolable: you satisfy the requirements whenever you are subject to them. This is why the requirements are not normative but descriptive.

This view is to be distinguished from the view advanced by Niko Kolodny (2005, 2008a, b), who also calls the idea that there are “requirements of formal coherence as such” a *myth*. On Kolodny’s view, there are simply no requirements that prohibit formally incoherent combinations of beliefs and intentions (like the ones this essay focuses on). There are only *reasons* which primarily require (or forbid) a *particular* attitude rather than a *combination* of attitudes: we are only misled into the illusion of coherence requirements because complying with the requirements of reasons generally (but derivatively) guarantees having coherent patterns of attitudes.<sup>1</sup> On my view, however, coherence requirements do have a genuine subject matter: they are standards for determining whether you *are* believing or intending an object. The real myth is the idea that they are *normative* requirements you can violate.

It should be made clear from the outset that my view doesn’t imply a radical error-theory about ascriptions of *irrationality*. Recent literature on rationality features two prominent concepts of rationality: (1) *structural* rationality, which is a matter of avoiding incoherent combinations of attitudes (Scanlon 2007; Broome 2013; Worsnip 2016); (2) *substantive* rationality, which consists in responding correctly to one’s reasons (Schroeder 2009; Parfit 2011; Kiesewetter 2017; Lord 2018). My view implies only that structural irrationality is impossible, which is entirely consistent with the possibility of substantive irrationality. Indeed, as I shall argue, ordinary ascriptions of irrationality can be seen as referring to other forms of irrationality than violations of the structural requirements.

## 1 Preliminaries

This section introduces the coherence requirements this essay focuses on, lays out their basic characteristics, and elaborates on the concept of *irrationality* which the violation of a coherence requirement would entail.

An important strand of recent work on rationality understands rationality as a matter of *coherence between one’s attitudes*. The idea is that there are certain

<sup>1</sup> To be precise, Kolodny thinks that there *are* genuine requirements of rationality which tell you to hold attitudes in accordance with your (second-order) beliefs about *reasons* for those attitudes. But he denies that you always have reason to satisfy such requirements. Thus, as Kiesewetter (2017: 22) notes, Kolodny’s view concerns two myths: (1) the idea that there are requirements of formal coherence as such; and (2) the idea you always have reason to satisfy the requirements of rationality.

combinations of attitudes it would be intuitively irrational for you to have. We can explain why there would be something rationally amiss if you had such a combination of attitudes, by positing a corresponding requirement of rationality forbidding you from having it. Broome (2013: Ch. 9), for example, puts forward the following requirements:

**No Contradictory Beliefs** Rationality requires of you that you do not believe at  $t$  that  $p$  and also believe at  $t$  that not- $p$ .

**No Contradictory Intentions** Rationality requires of you that you do not intend at  $t$  that  $p$  and also intend at  $t$  that not- $p$ .

**Modus Ponens** Rationality requires of you that if you believe at  $t$  that  $p$  and you believe at  $t$  that if  $p$  then  $q$ , then you believe at  $t$  that  $q$ .<sup>2</sup>

**Means-End Coherence** Rationality requires of you that you intend  $M$  at  $t$  if you intend an end  $E$  at  $t$  and believe at  $t$  that  $M$  is a necessary means to  $E$ .<sup>3,4</sup>

One distinguishing mark of the coherence requirements is that they govern *combinations* of attitudes, without requiring you to have or lack a *particular* attitude. The basic idea is that there would be something wrong with having an incoherent combination of attitudes as such, regardless of whether you have reasons for (or against) each attitude, *taken individually* (for convenience, I shall use ‘attitudes’ to include *absences* of attitudes).<sup>5</sup> If you believe that Superman is a woman and believe that if Superman is a woman then it will rain tomorrow, you would be irrational if you did not also believe that it will rain tomorrow. This is so even if you have strong reasons against believing that it will rain tomorrow, as well as believing that Superman is a woman and believing that it will rain tomorrow if Superman is a woman.

Each requirement is agreed to be *local* rather than *global*, in the sense that they specify specific conflicts between your attitudes, as opposed to what rationality

<sup>2</sup> There are possible qualifications to this principle which I ignore for the sake of simplicity. For example, Broome (2013: 157) argues that we need to add the condition that *you care at  $t$  whether  $q$*  in order to capture the conditions under which you would be irrational. Kieseewetter (2017: 255) adds the condition that *you attend at  $t$  to  $q$* . Both are inspired by Harman’s (1986: 12) idea that rationality cannot plausibly require you to “clutter your mind” with trivialities. I simply assume that there is a correct principle in the vicinity.

<sup>3</sup> There are again possible qualifications I shall ignore. For example, if you believe that you will attain  $M$  anyway regardless of whether you intend  $M$ , or if you believe that you don’t *then* have to intend  $M$ , it doesn’t seem irrational for you to fail to intend what you believe to be a necessary means to  $E$ . See Setiya (2007), Finlay (2009), Broome (2013) for some relevant qualifications.

<sup>4</sup> To this list, Broome (2013) adds “Enkrasia”, which requires that you intend to  $A$  if you believe that you ought to  $A$ . However, I shall set aside Enkrasia in this essay, since the status of Enkrasia as a coherence requirement is relatively obscure, as Broome (2013: 173–175) himself notes and Reisner (2013) has argued. As noted in footnote 1, Kolodny also distinguishes between requirements of formal coherence and enkrasia-style requirements and thinks that only the latter are genuine requirements of rationality.

<sup>5</sup> Scanlon (1998, 2007) also distinguishes between structural claims about the relations between one’s attitudes that must hold if one is to be rational and substantive claims about reasons for one’s attitudes, and suggests that we restrict the charge of irrationality to the failure to meet the structural requirements. This distinction is implicitly invoked by Davidson (2004: Ch. 12).

requires of you in light of your entire set of attitudes. The idea is that you would be irrational if you violated a local requirement, regardless of your overall psychology. Thus, the satisfaction of such a requirement guarantees local coherence between specific attitudes, rather than the coherence of one's attitude-states as a whole.<sup>6</sup>

In being a *requirement*, a coherence requirement is said to be *strict*: it is irrational, always, everywhere, and for anyone, to violate a coherence requirement. This means that the requirements are *insensitive* to your particular situation: what you are required to do remains the same across every possible situation. In this respect, the coherence requirements contrast with what you have *most reason* to (or *ought* to) believe or intend. For what you have most reason to believe is generally a function of your epistemic situation, including the body of evidence available to you, and what you have reason to intend is likewise a function of the relevant features of your situation. Moreover, the strictness of a requirement means that you always go wrong if you violate it. In this respect, being required to  $\varphi$  contrasts with *being recommended* or *having a reason to  $\varphi$* .<sup>7</sup> For you are not always subject to criticism for failing to do what you are recommended or have a reason to do. You might, for example, have a reason (say, pleasure or relief) to smoke a cigarette and yet have a stronger reason (say, health) not to smoke at the same time, in which case you would not be subject to criticism for not smoking.

Let me finally articulate the concept of *irrationality* I shall work with, which I believe is well entrenched in the philosophical literature and seems to be borne out by ordinary attributions of irrationality. On this conception, irrationality implies *personal* criticizability, and in this respect analogous with *blameworthiness*: being irrational, just like being blameworthy, reflects badly on you as a person.<sup>8</sup> To be

<sup>6</sup> For the notion of local requirement of rationality, see Kolodny (2005), Broome (2007), Bruner (2010), Way (2011). An anonymous referee notes that this contrast between *local* and *global* requirements overlooks what one might call *regional* requirements, or what rationality requires of you in light of some *not-so-small-but-not-global* set of attitudes. It might be that local, regional, and global requirements naturally form a spectrum, and a good account of rational coherence should generalize beyond the small number of local requirements. While I cannot fully address this concern here, I would justify the focus on *local* requirements by pointing out that they are, unlike other requirements, *exceptionless*: violating such a requirement is, in every possibility, sufficient for your irrationality. For example, while it is not so controversial that it would always be irrational for anyone to hold a *pair* of contradictory beliefs or intentions, it is hard to find a stronger consistency requirement involving a larger number of attitudes, the violation of which guarantees irrationality. See Broome (2013) and Kiesewetter (2017) for discussions of exceptions to consistency requirements on belief and see McCann (1991) and Kolodny (2008b) for discussions of exceptions to consistency requirements on intention. Admittedly, this picture of local coherence requirements cannot be taken for granted. Fogal (2019), for example, denies that there are requirements that are strict in this sense and defends a view according to which coherence is a matter of responding to *gradable* and *pro tanto* "pressure" generated by an attitude to form or give up another attitude. On his view, there are no principles the violation of which would necessitate irrationality, and there are effectively no "requirements" of coherence in the strict sense. I will have to set aside such a view for the purpose of this paper.

<sup>7</sup> Broome (1999) was the first to explicitly draw the contrast, calling the requiring relation 'strict' and the reasons relation 'slack'. This distinction is accepted by, for example, Setiya (2007), Schroeder (2009), Way (2018).

<sup>8</sup> Parfit (2011: 163) notes, "questions about rationality are, in several ways, like questions about blameworthiness" and claims, "we ought [...] to use the words 'rational' and 'irrational' to express certain kinds of praise or criticism."

clear, my aim in this section is to articulate, and provide some basic motivation for, a prominent conception of (ir)rationality, not to argue that this is the only conception of rationality available to us, or that we should accept it as the *right* conception. Rather, I shall take this conception for granted and see where it leads with respect to the normative status of the coherence requirements. While I shall address an alternative conception of rationality towards the end of the essay (Sect. 4), it is beyond the scope of this essay to adjudicate between competing conceptions of rationality. Thus, the reader could regard my argument as having a conditional form: *if* the conception of irrationality as personal criticizability is correct, *then* there are no violable coherence requirements of rationality.

It is not difficult to see what motivates the idea that irrationality implies personal criticizability. Both judgments about irrationality and judgments about blameworthiness are sensitive to your epistemic position (e.g. beliefs and available evidence) and your capacities: you are not necessarily irrational or blameworthy for failing to do what you (objectively) ought to do, if the relevant facts are simply beyond your ken or if you lack the relevant capacity to be responsive to such facts.

The conception of irrationality as personal criticizability has been explicitly defended in the literature. Setiya (2004), for example, argues that irrationality is akin to *moral culpability*, in that whenever you are irrational, you fail to avoid what you could have been *legitimately expected* to avoid, where what you could be legitimately expected to avoid is a function of your relevant capacities.<sup>9</sup> Kiesewetter (2017) also defends this conception of irrationality. On his view, ordinary ascriptions of irrationality are naturally taken to express criticism, and agents who get called irrational “do not merely understand this to mean that they fell short of some evaluative standards”, but rather “feel personally criticized for their responses.” (Kiesewetter 2017: 39)<sup>10</sup> Lord (2018: 3) agrees: “[the] sort of credit or blame one is open to when one is rational or irrational is thus a very personal evaluation.” Ridge (2014: 244) seems to concur: “*ceteris paribus*, to call someone irrational is to insult, or at least criticize them.” This motivates the following hypothesis:

**Personal Criticizability** Necessarily, if you are irrational, you are personally criticizable.

Personal Criticizability explains the plausible idea that irrationality is not just a matter of falling short of some evaluative standard. As Way (2018) writes, the charge of irrationality seems to be distinct from the kind of criticism that “we might express

<sup>9</sup> Way (2018) endorses this point. See Finlay (2009) for a similar conception of irrationality, according to which you are rationally required to  $\varphi$  in a given situation only if you could be reasonably expected to  $\varphi$ . Wedgwood (2013) also argues that whenever you are rationally required to  $\varphi$ , you have the *specific* ability to  $\varphi$ , where this is a matter of having the relevant rational capacity and the opportunities to exercise it.

<sup>10</sup> Kiesewetter (2017: 38–39) credits this point to Nadeem Hussain, who often debates with his friends whether their attitudes towards God are irrational. Hussain imagines himself saying ‘I think you are irrational, but this is not meant as a criticism. I don’t think there is anything *wrong* with being irrational’, and notes that this would undermine the point of their disagreement. For in such a case his friends could appropriately reply, ‘If you don’t think there’s anything wrong with us, then stop calling us irrational’.

by saying ‘he made a mistake’, ‘he did the wrong thing’”.<sup>11</sup> You might fall short of such standards because you are ignorant of or mistaken about the relevant facts, or you lack the relevant capacity to respond to such facts, in which case you would not be personally criticizable. Personal Criticizability explains why you would not be irrational in such cases.

Personal Criticizability also explains the widely accepted idea that rational requirements on attitudes should not be *too demanding* for ordinary agents like us. For example, it has been argued that rationality doesn’t require that *all* of our beliefs be logically consistent, given that we are not logically omniscient (Broome 2013). It has been argued that rationality doesn’t require that we believe every logical consequence of what we believe, given our limited cognitive capacities and abilities (Harman 1986; Schroeder 2004; Broome 2013). It has also been argued that rationality doesn’t issue any *synchronic* requirements, because it takes a considerable amount of time for finite beings like us to update our attitudes to satisfy the (putative) requirements on our attitudes (Podgorski 2017). Personal Criticizability explains why a (putative) rational requirement should not be too demanding: if it is too demanding, you could not be legitimately expected to satisfy it, and you would not be personally criticizable for your failure to satisfy it. Personal Criticizability entails that it is not irrational for you to fail to do so.

## 2 Constitutivism About Coherence Requirements

As a first step towards the thesis that the coherence requirements are not normative, this section argues that there are conditions under which you don’t count as believing or intending unless you satisfy those requirements. The argument is based on the *unintelligibility* of attributing such combinations of beliefs and intentions to someone who meets those conditions. In particular, I shall defend the following thesis:

**Constitutivism** It is partly constitutive of belief and intention that your occurrent (or activated) beliefs and intentions satisfy the coherence requirements.

### 2.1 What Constitutivism Says

Constitutivism is a version of the more general idea that compliance with rational requirements is *in some way* constitutive of having relevant attitudes, which has been endorsed by a number of philosophers.<sup>12</sup> It states that your occurrent (or activated) beliefs and intentions necessarily satisfy the requirements.<sup>13</sup> By an *occurrent* or *activated* attitude, I mean an attitude currently being *employed* or *utilized* in your

<sup>11</sup> Way draws on Parfit’s (1984, 2011) idea that to call someone irrational is to express a serious kind of criticism that would also be expressed by calling someone ‘stupid’, ‘senseless’, or ‘crazy’, and to express a weaker kind of criticism we can use phrases like ‘less than fully rational’. See also Scanlon (1998).

<sup>12</sup> See, for example, Dennett (1989), Blackburn (1998), Davidson (2001, 2004), Wedgwood (2009).

<sup>13</sup> Finlay (2009) has advanced such a view regarding Means-End Coherence. Worsnip (2018) makes a similar point regarding coherence requirements in general.

cognitive task or problem-solving. When your attitude is so employed, it plays its characteristic role in your mental economy.<sup>14</sup> Here is a brief sketch of the characteristic roles played by occurrent belief and intention. Plausibly, what is distinctive about belief and intention is that they are *settling* attitudes: belief settles the question of what is true; and intention settles the question of what to do.<sup>15</sup> When you *judge* (or occurrently believe) a proposition  $p$ , for example, your belief is employed in answering the question whether  $p$  is true. When you utilize your belief in  $p$  in *reasoning* or *action*, you treat  $p$  as true without questioning its truth: if your belief that Alice arrives at the airport around 2 pm is occurrent, for example, you might *conclude* in your reasoning that you need to leave home around 1 pm to pick her up and *act* accordingly, treating it as true that she will arrive around 2 pm. Likewise, when you decide (and occurrently intend) to  $\varphi$ , your intention is employed in answering the question *whether to  $\varphi$* . When your intention to  $\varphi$  is utilized in your reasoning or action, you take it to be true that you will  $\varphi$  and plan accordingly (in the case of future-directed intention), or simply try to  $\varphi$  (in the case of present-directed intention).

While the word ‘occurrent’ is sometimes treated as a synonym for ‘conscious’ in the literature, an attitude’s being occurrent is different from its being *conscious*, on my distinction. For being consciously aware that you have an attitude is neither necessary nor sufficient for the attitude’s being employed. You might, when performing reasoning, employ your belief in  $P$  as a premise without being aware that you have that belief, in which case your belief is activated but not conscious. Likewise, you might, through some psychotherapy, become aware that you believe  $Q$ , but that belief might not be employed in any of your activities, in which case your belief in  $Q$  is conscious but not occurrent.

You might possess an attitude that is not occurrent, which I shall call a *dispositional* attitude. One intuitive notion of attitudes like belief and intention is that they are standing, dispositional states, stored in your cognitive system, which can be truly ascribed to you even when they are not occurrent, as when you are asleep or not thinking about anything. I take it that the following is true about the relationship between dispositional and occurrent attitudes: if you dispositionally believe (or intend)  $p$ , you are disposed to occurrently believe (or intend)  $p$ , under appropriate triggering conditions.<sup>16</sup>

<sup>14</sup> See Bartlett (2018) for an understanding of occurrent states which is close to one that I endorse.

<sup>15</sup> This point about settledness as a distinguishing feature of intention has been duly made in some classic work on the nature of intention. See, for example, Harman (1976), Bratman (1987), Velleman (1989), Mele (1992). As for the settledness involved in belief, see Friedman (2019). See Hieronymi (2005, 2006) for a unifying account of belief and intention as question-settling attitudes.

<sup>16</sup> As Audi (1994) notes, the converse doesn’t hold, because of the distinction between a *dispositional belief* and a *disposition to believe*.



## 2.2 A Defense of Constitutivism

To see why Constitutivism is plausible, consider the following dialogue: Suppose I tell you that, upon being asked whether I am going to visit my friend at 3 pm, that I intend to visit him, that I believe that the only available way to do so is to get a cab now. Suppose, however, that I don't get a cab. You ask me, 'Is anyone getting a cab for you?' and I answer, 'No—I don't intend to get one. I hate spending money on a cab.' You ask, 'So now you don't intend to visit your friend?' and I answer, 'Of course I do! It's just that I don't intend to get a cab.' As Worsnip (2018: 229) notes regarding a similar case, "the most natural way to hear [my] speech is as a joke". But if you assume my sincerity, you can legitimately cast doubt on the intelligibility of my attitudes-report: either I don't really intend to visit my friend, or I don't really believe that I need to get a cab now, in which case I don't violate Means-End Coherence.<sup>17</sup>

By contrast, no such puzzle would arise and my attitudes-report would be fully intelligible if the attitudes I claim to have were some other conative attitude. Suppose I tell you that, instead of intending, I *hope*, *wish*, or *want* to visit my friend. Suppose I have exactly the same beliefs as above, and again fail to get a cab, saying that I dislike spending money on cabs. This time, the claim about the attitudes I have seems perfectly intelligible, which would be clearer if I backed up my claim as follows: 'Look, I really *want* to visit my friend, but I haven't actually made my mind up about whether to do so. You know, I need to get myself on a cab now, but I really don't *want* to spend that much money, too.'

The same goes for No Contradictory Intentions. Suppose I say, upon being asked whether I'll be at the gym this evening, 'I intend to go to the gym, but I also intend not to'. You ask, 'But if you succeed in not going, doesn't that mean that you won't be there?' and I answer, 'I know that, of course, but I still intend to go to the gym, too.' Again, assuming my sincerity, you can legitimately cast doubt on the intelligibility of my attitudes-report: either I don't really intend to go to the gym or I don't intend not to do so, in which case I don't violate No Contradictory Intentions. No such unintelligibility would be present, however, if my claim were about what I *want* to do.

Now consider a putative case of contradictory *beliefs*. It would likewise be difficult to take what I say at face value if I said, upon being asked whether God exists, 'I believe God exists but also believe that God doesn't.' You ask, 'So you think God exists?' and I answer in the affirmative. You again ask, 'But doesn't that mean that it cannot be true that God doesn't exist?' and I answer, 'Sure, but I believe that God doesn't exist, too.' Again, it seems that you could legitimately doubt that I really have the attitudes I claim to have: either I don't really believe that God exists, or I

<sup>17</sup> We could of course continue further. You might ask, 'but didn't you say that you intend to visit him at 3 pm? How are you going to accomplish that?', and none of the following possible answers would be intelligible: (1) 'I believe I won't, but I still intend to'; (2) 'I'm going to take a cab, of course, but I don't intend to do so.' See Wallace (2001) and Finlay (2009) for a similar doubt as to whether you can be intelligibly seen as intending the end in such a case.



don't believe that God doesn't exist. By contrast, it would make perfect sense if I replaced 'believe' with some other cognitive attitude-verb, such as 'assume', 'imagine', '50% confident', etc. For example, the following seems perfectly intelligible: 'Look, I'm assuming that God exists and also that God doesn't exist, only for the sake of argument. I haven't made up my mind about whether God really exists—I don't have decisive evidence either way.'

Consider, finally, a putative violation of Modus Ponens. Suppose I say, 'I'm now interested in whether there's heaven. I believe that God exists, and I also believe that heaven exists if God exists. But I don't believe that there is heaven.' You ask, 'Well, why not? What you just said sounds like: *God exists and God's existence implies the existence of heaven*. Have you just given up one of your beliefs?' I answer, 'Not at all! I still believe that God exists, and I also believe that if there's God, there's heaven. But I don't believe that there's heaven.' You ask, as a last-ditch attempt, 'Are you sure you want to come to a position on this?' and I answer yes. At this point, you could again legitimately doubt that I really have the attitudes I claim to have, assuming my sincerity.

Here is the datum that calls for an explanation. In each case, I seem to fail to meet some basic conditions for *intelligibility*: it is unclear how to make sense of the mental states I ascribe myself.<sup>18</sup> Importantly, the datum is not just that the attitudes I claim myself to have seem *bad*, *unreasonable*, or *unjustified*. Intelligibility is one thing and badness is another, for a violation of a norm doesn't necessarily make things unintelligible in this way. When we assess people's attitudes as bad, unreasonable, or unjustified, we do so precisely on the grounds that they *have* those attitudes: people violate a moral norm when they *do* intend heinous deeds and violate an epistemic norm when they *do* believe against their available evidence. Moreover, we don't have difficulty understanding what it is to be in such a state: we understand what it is for a person to intend to torture innocent people and predict how such a person would behave on the basis of this intention. Intelligibility, in short, is a *precondition* for normative assessment, and the intuitive judgment about the above cases seems to be the judgment that I fail to meet such a precondition, rather than the judgment that I *have* attitudes that violate some norm.

Constitutivism nicely explains the datum. In each case, it is naturally assumed that all of my relevant beliefs and intentions are occurrent, with respect to the question to which I am attending. It follows from Constitutivism, however, that I fail to count as believing or intending unless these occurrent attitudes satisfy the relevant requirement. This explains why my attitudes-reports are baffling: I claim to have the attitudes that I don't count as having.

<sup>18</sup> Worsnip (2018) has recently argued that we observe exactly the same kind of unintelligibility when we imagine transparent violations of other coherence requirements, such as the requirement to have a *transitive* pattern of preferences ("Do not prefer A to B, prefer B to C, and prefer C to A!"), and the requirement to harmonize one's first-order and higher-order doxastic attitudes ("Do not believe *p* while also believing that you lack sufficient evidence for *p*!"). Friedman (2019) also proposes the requirement not to simultaneously believe that *p* and *inquire* into whether *p* and notes that violating the requirement in full consciousness would be unintelligible.

So much for what Constitutivism declares impossible. However, Constitutivism *allows* for the possibility of having an incoherent combination of beliefs and intentions, understood as *dispositional* states. The basic idea is that there is a good sense in which you possess combinations of attitudes even when they aren't simultaneously occurrent or activated. Indeed, a number of philosophers have invoked the idea that your mind consists of several distinct *compartments* (or *fragments*) in explaining how you could hold incoherent patterns of attitudes: when your attitudes belong to different compartments (or fragments), they might be activated under different conditions without being integrated into a unified system (Lewis 1982; Stalnaker 1984; Davidson 2004; Egan 2008). Modeling an agent's attitudinal states as dispositional states with various triggering or activation conditions provides us with a key to understanding how you could have combinations of attitudes that appear to run afoul of the requirements at issue.

Consider, for example, contradictory beliefs and intentions. Even if you believe (or intend) that  $p$  and also believe (or intend) that  $\textit{not-p}$ , those beliefs (or intentions) might be activated under different types of circumstances. For example, Broome offers the following case, in the context of illustrating the possibility of contradictory beliefs and intentions, respectively:

Suppose you used to keep the whisky under the bed, but recently you started keeping it behind the fridge. When you instinctively go for whisky, you head for the bedroom, but if you first think about it, you go to the kitchen. [...] You might form the intention of being in London on 23 February, to go to a seminar. Then, later, you might learn that there will be a tube strike in London on 23 February, and form the intention of not being in London that day, without noticing you already have the opposite intention. (Broome 2013: 156)

In this case, you could be naturally seen as holding a dispositional belief that the whisky is under the bed, which would be activated under some triggering conditions (e.g. getting a desire for it), but also a belief that it is not under the bed, which would be activated under others (e.g. carefully thinking about where it is). Likewise, you could be seen as having a dispositional intention to be in London which would be activated in one context (e.g. thinking about the seminar), and also a conflicting dispositional intention not to be there, which would be activated in another (e.g. thinking about the tube strike). If so, Broome's case illustrates the possibility of holding a pair of dispositional beliefs or intentions towards contradictory contents.

We can also understand how someone could have combinations of attitudes that seem to violate Modus Ponens and Means-End Coherence. Even when you believe  $p$  and also believe *if  $p$  then  $q$* , you might not activate both of them simultaneously and fail to put them together to believe  $q$ . Likewise, even when you have a dispositional intention to  $E$ , and a dispositional belief that doing  $M$  is the only available means to  $E$ , you might fail to intend to  $M$ , because the belief and intention are never brought together at the same time.

### 3 Why You Cannot Violate the Requirements

In the previous section, I argued for Constitutivism, according to which there are conditions under which your beliefs and intentions cannot violate the coherence requirements: the conditions under which they are all occurrent. In doing so, however, I have granted that there are also conditions under which you might have combinations of attitudes that go against the requirements: the conditions under which at least one of the attitudes remains dispositional. Thus, one might think that I have only made it harder for myself to establish the main thesis: why not simply say that the requirements can be violated precisely under those conditions? Indeed, many philosophers who recognize that we have a general tendency to meet rational principles on our attitudes appeal to the fact that some of our attitudes could sometimes remain purely dispositional, in making room for the violability of those principles.<sup>19</sup>

More recently, Worsnip (2018) has advocated a view similar in spirit with Constitutivism. He also argues, on the basis of the unintelligibility of the kind of attitudes-reports considered in the previous section, that a combination of attitudes is incoherent just in case “it is constitutive of the [attitudes] in question that for any agent that holds these attitudes, the agent is disposed, when the conditions of full transparency are met, to give up at least one of the attitudes”, where the “conditions of full transparency” refers to the “conditions under which the agent knows, and explicitly and consciously believes, that she has the states in question, without self-deception.” (Worsnip 2018: 188) This again is an instance of the view on which there are certain conditions under which one would satisfy the requirements. At the same time, however, Worsnip makes it explicit that his account doesn’t make violations of the requirements impossible, precisely because what he calls the *conditions of full transparency* are often not met (Worsnip 2018: 189).

This section argues that one cannot appeal to such conditions to make room for the violability of the requirements, because you are not *subject* to the requirements under such conditions. The basic idea is that, once we consider concrete cases where (some of) your attitudes are dispositional, violations of the requirements no longer necessitate *irrationality*. If your violation of a principle doesn’t make you irrational, however, it cannot be a *requirement*.

Here, then, is a principled argument for this view:

---

<sup>19</sup> We have already seen Broome (2013)’s attempt to allow for the violability of the coherence requirements. Both Railton (1997: 312–313) and Scanlon (2007: 61) argue that while intending an end E surely disposes one to give some weight to E in one’s subsequent practical reasoning, this is consistent with one’s giving no weight to E on a particular occasion, because the disposition can fail to be activated. Scanlon makes the same point about believing a proposition P and using P in one’s subsequent reasoning. Similarly, Sinhababu (2011) offers a Humean response to Korsgaard’s argument (1997) that the Humean account of motivation makes instrumental irrationality impossible, by invoking the notion of “dispositional desire strength”. What Korsgaard calls “true irrationality” can occur, on his view, when one’s dispositional desire for A is stronger than one’s dispositional desire for B, but one chooses to B because vivid representations of B, or the nearness of B in space or time, increases the violence of the desire for B, and motivates one to pursue B rather than A.

**P1** If your dispositional attitudes are subject to a coherence requirement  $R$ , then necessarily, it is irrational for you to have a combination of dispositional attitudes prohibited by  $R$ .

**P2** Necessarily, if it is irrational for you to  $\varphi$ , then it is personally criticizable for you to  $\varphi$ . [Personal Criticizability]

**P3** If your dispositional attitudes are subject to  $R$ , then necessarily, it is personally criticizable for you to have a combination of dispositional attitudes prohibited by  $R$ . [from P1 and P2]

**P4** It is possible for you to have a combination of dispositional attitudes prohibited by  $R$  without being personally criticizable.

**C** Your dispositional attitudes are not subject to  $R$ . [from P3, P4]

P1 follows from the *strictness* of the requirements: on the assumption that the requirements apply to dispositional attitudes, it *must* be irrational for you to hold an incoherent combination of dispositional attitudes, regardless of your particular situation. P2 is just a restatement of Personal Criticizability from Sect. 1. The crucial premise is P4, which I shall defend by considering the following cases:

*Party* It's Monday and you are told that there will be a party on Saturday. You are with Jane, who tells you that she hates going out on rainy days. So you come to believe that if it rains on Saturday, Jane won't come to the party. A few days later, you hear from a weather forecast that it will rain on Saturday and come to believe it. At the moment, however, you are having an important conversation with your partner. So you don't bother at all to consider, in light of this information, whether Jane will come, and fail to believe that Jane won't come.

*Omelet* It's midnight and you feel like having an omelet for breakfast tomorrow, and so come to intend to make an omelet. Of course, you know that you've got to crack a few eggs to make an omelet. You wake up the next morning, check your email, only to find out that you are about to miss the deadline for your favorite conference. You work on your paper all morning, without thinking about breakfast at all, and fail to (intend to) break eggs.

*London* You form the intention to go to a seminar on 23 February, upon noticing that it features some excellent speakers. Then, later, you learn from news that there will be a tube strike in London on 23 February and form the intention not to be in London. Your prior intention to be in London simply doesn't occur to you, and you have no time to check your calendar at the moment, because you are preparing for a talk.<sup>20</sup>

*Neighbor* You are disposed to judge that Xiao is the person living in Apartment 502 upon hearing the name 'Xiao', because you have strong evidence

<sup>20</sup> This is a variant of the case discussed in Sect. 2, which Broome (2013: 154) offers.

that he lives in 502. But when asked, ‘who is the person living in Apartment 502?’ you draw blank. You are now carefully going over an important document on renewing your lease with your apartment manager, who tells you in passing that the person living in 502 is not Xiao. You nod and continue reading. As a result, you are now disposed to judge that Xiao is the guy living in 502 upon hearing the name ‘Xiao’, but also disposed to judge that it is not Xiao who lives in 502 upon being asked, ‘who is the person living in Apartment 502?’.<sup>21</sup>

In each case, you begin with a rationally permissible state of mind and undergo a change in your attitude, which leaves you with a combination of *dispositional* attitudes that apparently violates a coherence requirement: Modus Ponens in Party; Means-End Coherence in Omelet; No Contradictory Intentions in London; and No Contradictory Beliefs in Neighbor.

My intuitive judgment is that you are not personally criticizable in any of the above cases: your failure, if there is any, doesn’t seem to be criticizable in a way that is akin to blameworthiness. Correspondingly, none of the above cases seems to be a clear case of irrationality. This is because, I argue, that it is possible for the *difficulty* of exercising your relevant capacity to make it the case that you are not liable to personal criticism, and each of the above cases illustrates such a possibility. In each case, it is too difficult for you to exercise your relevant rational capacity at the moment when you acquire the attitude that results in the incoherent combination of dispositional attitudes. As a result, you could not, at that time, be legitimately expected to avoid having such a combination of attitudes. Plausibly, however, you are personally criticizable for failing to satisfy a requirement at *t* only if you could be legitimately expected to do satisfy it at *t*.

To see this, consider the following case:

*Headphones* You live in an apartment alongside a beautiful lake. Stressed out with work, you decide to listen to music. As a responsible tenant, you wear headphones, enjoying the music in full volume. Alas, there happens to be a child drowning in the lake, screaming desperately for help. You have the normal hearing ability to notice the child: if you were not wearing headphones, you would clearly hear the scream and notice the child. At the moment, however, the scream just strikes you as a faint noise. Since the lake is usually quiet and peaceful, you have no reason to expect there to be a child drowning in the lake. So you don’t bother to cast your eyes at the lake.

Let us assume that you have a duty to notice when there is a person in the vicinity screaming for help, which you violate in this case. Plausibly, however, you are not personally criticizable for failing to notice the child, since you could not be legitimately expected to notice the child. Although you have the requisite hearing ability,

<sup>21</sup> This case is based on Elga and Rayo’s (2015: 3) example of a person who can produce his neighbor’s apartment number when supplied with her name, but not the other way around.

you are, at the moment, listening to a loud music, which makes it extremely difficult for you to exercise it.

Just as you have the general ability to notice the child in Headphones, you have the relevant rational capacity to satisfy the requirements in question in the above cases. Were you to attend to the question of whether Jane will come in Party, you would easily conclude that Jane won't come (or revise one of your antecedent beliefs) and satisfy Modus Ponens. Were you to recall your intention to make an omelet for breakfast in Omelet, you would easily give up the prior intention to cook an omelet or come to intend the means, and thereby satisfy Means-End Coherence. Likewise, you would easily avoid having contradictory intentions and beliefs in London and Neighbor, if you were to consider your plan for 23 February, and ask yourself whether Xiao lives in 502, respectively.

Crucially, however, you are in the middle of a task which requires a significant amount of cognitive resources when you acquire the attitude that puts you in an apparently incoherent state of mind. This makes it difficult for you to exercise your rational capacity, just as the loud music in Headphones makes it difficult for you to notice the child. This is because you are not an ideally rational being with unlimited cognitive resources and abilities. While it is surely part of our rational capacity to satisfy the relevant requirements when we *attend* to relevant questions or *notice* a tension between our attitudes, it isn't plausibly part of our rational capacity to instantaneously update our attitudes in response to every relevant change in our attitudes. Thus, to the extent we are in a situation which makes it extremely difficult for us to attend to a relevant question or to notice any conflict among our attitudes, it is difficult for us to exercise our relevant rational capacity.

Let me now address some possible objections. First, one might think that personal criticizability (and irrationality) comes in *degrees*: to be criticizable *simpliciter* is to be criticizable *to some degree*, but most of our judgments about whether someone is personally criticizable (or irrational) are judgments concerning whether they are criticizable to a *contextually-relevant degree*. On this basis, one might argue that although you aren't personally criticizable in the contextually relevant sense in the above cases, you are still personally criticizable to *some (but not very high) degree*: while the difficulty of exercising the relevant capacity in the relevant circumstances significantly *mitigates* your criticizability, it doesn't fully eliminate your criticizability.<sup>22</sup>

My response is that you are personally criticizable for  $\varphi$ -ing to *some degree* only if  $\varphi$ -ing is *attributable* to you as an agent, in the sense that your  $\varphi$ -ing can be sensibly appraised as *faulty* in some way that reflects poorly on you.<sup>23</sup> However, it is unclear that your failure to exercise your rational capacity in each case is even

<sup>22</sup> I am grateful to an anonymous referee for asking me to address this objection.

<sup>23</sup> See Watson (1996) for the concept of responsibility as *attributability*, where an action is attributable if it counts as a sensible target of *aretaic appraisal*, which basically concerns the agent's excellences and faults. Fischer and Tognazzini (2011) plausibly argue that the question about attributability is the question that is conceptually prior to any other questions about responsibility, and that other, richer notions of responsibility have attributability as a basic condition. So, if criticizability (plausibly) entails responsibility, it also entails, at minimum, attributability.

attributable to you, and so it is unclear that you could be held criticizable *at all*. The (alleged) failure doesn't consist in *incorrectly* exercising your capacity (you don't exercise the capacity at all!), which means that it doesn't reflect *how* you exercised the capacity. Nor does the failure seem to reflect your *choice* or *values*: it is not as if the difficulty of exercising your capacity is self-imposed, or you choose not to exercise the capacity. Nor does the failure seem to reflect your *character*: it is not as if you are generally negligent person who is (deliberately or not) inattentive to relevant features of your situation. If anything, your failure reflects your *cognitive limitations* which explain why you fail to satisfy the relevant requirements. However, when one's failure is explained simply by one's lack or possession of a capacity, it cannot plausibly be attributed to one as a ground for criticism.

Second, one might object that my argument relies on an overly narrow conception of rational capacities, on which we can only satisfy the relevant requirements by attending to a relevant question or noticing a tension between our attitudes. It maybe that we are also equipped with some automatic mechanism, operating at the *sub-personal* level, which updates our attitudes without being mediated by our consideration or attention. For example, we often automatically lose our belief that *p* as soon as we learn that not-*p*, without explicitly considering or attending to any question (cf. Broome 2013: 206). If so, it may be that we could be expected, on the basis of such a mechanism, to satisfy the requirements even in the above cases.

I find it plausible that we are equipped with such a sub-personal mechanism. I deny, however, that the presence of such a mechanism can ground *personal* criticizability. For even if you failed to satisfy a requirement due to the malfunctioning of such a mechanism, it would not plausibly reflect badly on you as a *person*, just as catching the flu due to an abrupt failure of your immune system doesn't reflect badly on you as a person. It would rather count as a psychological failure, or a failure of some sub-personal system, which by itself doesn't warrant any normative criticism. You would be personally criticizable for violating a requirement only if it were due to the failure to exercise the capacity at the *personal level*. However, it is plausible that any exercise of such a capacity requires, at its minimum, your notice or attention: a process or an event that happens to you even without your attending or noticing could not be plausibly seen as the personal-level exercise of your capacity.<sup>24</sup>

Finally, one might object that there is still some sense in which you *go wrong* when you come to have incoherent combinations of dispositional attitudes in the above cases, which my argument seems to rule out. However, I don't deny that there are other ways in which your states of mind could be negatively assessed. For example, there is something *unfortunate* or *less than ideal* about your overall state of mind in each case. You might feel like a fool, admit that you were confused, that you have fallen short, etc., if you later realize that you were in such a state of mind. My point is that the propriety of such assessments doesn't entail *personal* criticizability, since they could be appropriate even when you suffer from a mere *psychological*

<sup>24</sup> Kolodny also notes "irrationality is a failing of the person, not of some sub-personal system" in arguing against the idea that rationality is a matter of proper functioning of a particular kind of system (Kolodny 2005: 521–522).



failure through no fault of your own. Consider a memory lapse, for example. Suppose that you are taking a test, but you momentarily forget the answer (which you know) because you happen to be too tired or nervous for no reason. If you end up giving the wrong answer, there is something unfortunate or less than ideal about your state of mind in such a case. Still, it doesn't follow that you are personally criticizable.

This concludes my defense of P4, the claim that it is possible for you to hold an incoherent combination of dispositional attitudes without being personally criticizable. Together with Personal Criticizability, this implies that you are not subject to the coherence requirements when your relevant attitudes are not simultaneously occurrent.

If *all* of your attitudes were occurrent, of course, you *would* be subject to the requirements, and it *would* indeed be irrational for you to violate them. For there could not be clearer cases of irrationality than, say, occurrently believing that Xiao lives in 502 and also that he doesn't, or failing to answer whether Jane will come to the party, while employing the belief that she won't if it rains and the belief that it will rain, or occurrently intending to make an omelet but failing to intend to break eggs. However, Constitutivism denies that you can violate the requirements under such conditions: you satisfy the requirements whenever you are subject to them.

#### 4 Tying Up the Loose Ends

The idea that the coherence requirements at issue are inviolable principles that simply describe how beliefs and intentions operate might seem incredible. Surely it cannot be that every ascription of irrationality is false? This section clarifies the overall picture of coherence requirements that emerges out of my argument and addresses some possible misgivings about it.

First, my argument has focused only on local coherence requirements, which apply to *combinations* of attitudes rather than individual attitudes. However, there is yet another important dimension in our rational assessment: correctly *responding to one's reasons* by adopting or giving up particular attitudes (Schroeder 2009; Parfit 2011; Kiesewetter 2017; Lord 2018; Fogal 2019), which goes under the name of *substantive rationality*. People fail to be substantively rational (1) when they believe or intend what they have no reason to believe or intend, or (2) when they fail to believe or intend what they have strong reason to, and they are criticized as irrational in many such cases.

My argument is entirely consistent with the possibility of substantive irrationality. For you might fail to respond appropriately to your reasons while maintaining perfect coherence among your attitudes: you might avoid having contradictory beliefs by believing that you can travel faster than light and not believing that you cannot; you might avoid having contradictory intentions by intending to torture an innocent person and not intending not to do so; you might satisfy Modus Ponens by giving up one of your antecedent beliefs when you have most reason to form a belief in their consequence; and you might satisfy Means-End Coherence by giving up the

end (say, in the face of the difficulty of pursuing the means) when you have most reason to attain that end.

Second, my argument has focused only on *synchronic* requirements, which specify how our attitudes ought to be at a *given time*. Plausibly, however, there are *diachronic* requirements of rationality, prescribing how you should update your attitudes over time, or how mental processes like reasoning should unfold over time. Broome (2013: Ch.10), for example, argues that there are: (1) requirements of *persistence* on attitudes, which tell us to maintain our beliefs and intentions unless some “cancelling event” occurs; and (2) *basing prohibitions*, which tell us not to give up or acquire a type of attitude *on the basis of* some other type of attitude. Podgorski (2017: 15) argues that there are norms which tell us when it is appropriate to initiate and terminate a process of reasoning, what kinds of considerations are to be taken into account in such a process, etc.

My argument is also consistent with the violability of the diachronic (or process) requirements: the fact that you cannot violate the synchronic requirements by no means rules out the violability of such *diachronic* norms. You might randomly change your beliefs or intentions over time and thereby break some requirement of persistence, consistently with the fact that you cannot hold any contradictory (occurrent) beliefs or intentions at a given time. Likewise, you might begin or end a process of reasoning when you should not, or you might take inappropriate considerations into account in such a process and end up with terrible beliefs or intentions, consistently with the fact that you cannot break any of the synchronic requirements. So ordinary ascriptions of irrationality, by and large, can turn out true even on my account.

I have defended the thesis that the coherence requirements on attitudes, if at all, govern occurrent rather than dispositional attitudes on the basis of Personal Criticizability: any principle whose violation would guarantee personal criticizability would primarily concern occurrent attitudes. Moreover, this view naturally results from the consistent application of the idea that rational requirements should not be too demanding for ordinary agents like us, which is explained by Personal Criticizability (Sect. 1): if rationality doesn't demand that we always have logically consistent set of beliefs because we are not logically omniscient, and rationality doesn't demand that we have a deductively closed system of beliefs because our memory is limited, then rationality should not similarly demand that we always have coherent sets of *dispositional* attitudes because we are unable to update dispositional attitudes instantaneously and comply with the requirements in every situation.

Importantly, this doesn't mean that we cannot subject dispositional attitudes to criticism at all. First, as I have emphasized, dispositional attitudes you have can be criticized as *unfortunate* or *less than ideal* (Sect. 3). Second, we can understand the assessment of dispositional attitudes as *derivative* from the assessment of occurrent attitudes, for it is often through the formation or revision of occurrent attitudes that you form or revise dispositional attitudes. For example, if you come to occurrently believe or intend what you have reason not to believe or intend, then normally you will acquire a corresponding dispositional attitude that fails to be responsive to your reasons, and you could be derivatively criticized for having such an attitude even when the attitude is not activated.

At this point, one might wonder why we should accept Personal Criticizability in the first place, which I have assumed throughout the essay. Consider, for example, an *idealized* conception of rationality, according to which rational requirements primarily concern how ideally rational beings would maintain and update their attitudes (cf. Christensen 2007; Hedden 2016). Presumably, ideally rational beings enjoying unlimited cognitive resources and abilities would satisfy the requirements in any situation. If the attitudes held by such ideally rational beings set the standards for evaluating our attitudes, we would simply count as irrational to the extent we fail to approximate such beings, but mostly without being personally criticizable.

I shall neither dispute such an idealized conception of rationality, nor deny that, as Christensen (2007: 5) puts it, “while there are certainly some projects in epistemology that must take careful account of human limitations, they do not exhaust interesting epistemology.” I shall only suggest that it is difficult to do without Personal Criticizability entirely. In Sect. 1, I have noted that it does justice to ordinary ascriptions of irrationality, which seem to be sensitive to our cognitive limitations, and explains why irrationality seems to be distinct from mere *mistakes* or failure to meet some evaluative standards. If one denies Personal Criticizability and opts for an alternative conception of (ir)rationality, one is still left with the task of accommodating ordinary judgments of irrationality and explaining the distinctive force that the charge of irrationality seems to possess. This, of course, might not move the readers who take the requirements of coherence to be rational *ideals* in the strong sense. Still, my argument shows that there are no coherence requirements whose violation entails *personally criticizable* irrationality, or that a widely shared view of rationality commits us to denying the existence of violable requirements of coherence.

## 5 Conclusion

In this essay, I first argued for Constitutivism, the hypothesis that it is constitutive of belief and intention that occurrent beliefs and intentions always satisfy the coherence requirements. While Constitutivism is consistent with having incoherent combinations of dispositional attitudes, I have also argued that your dispositional attitudes are not subject to the requirements. This completes my argument the they are inviolable, and so are not normative principles: they should rather be seen as *descriptive* principles about how occurrent beliefs and intentions operate.

The view proposed here has two payoffs. First, it dissolves a puzzle raised by the natural idea that there are coherence (or structural) *requirements* of rationality. Rationality, on its face, seems to be an important source of normativity, which determines what we ought to do, unlike rules of chess or etiquette. But some writers, such as Broome (2005, 2007, 2013) and Kolodny (2005, 2008a, b), have forcefully argued that rationality understood as coherence is not normative: it is difficult to see what reason there is always to have coherent combinations of attitudes. But this gives rise to a puzzle: how could rationality be a source of normativity without giving us reason to meet its demands? My view simply dissolves this question, since there are no coherence *requirements* in the first place. There are only descriptive principles

about how our activated attitudes combine together, but no *normative* requirements we ought to satisfy.

Second, it has an interesting implication for the relationship between *structural* and *substantive* rationality. The principles of structural rationality are descriptive principles concerning the intelligibility of our attitude-ascriptions, whereas the principles of substantive rationality give us genuine norms for belief and intention we can violate. Thus, rationality as a normative (as opposed to descriptive) domain consists only of substantive rationality. Some writers, such as Kiesewetter (2017) and Lord (2018), have recently attempted to subsume structural rationality under substantive rationality, which arguably leads to independent philosophical problems (cf. Way 2018). My view implies that there is simply no need for such a unifying account, for we don't have two different kinds of *normative* principles to begin with.

**Acknowledgements** Open Access funding provided by Projekt DEAL. For extremely helpful comments and conversations, I thank Louis Gularte, Neil Roughley, Alex Worsnip, audiences at the 12th Mark L. Shapiro Graduate Philosophy Conference and the 6th Seoul Graduate Philosophy Conference, and two anonymous referees for this journal. I am especially grateful to Steve Finlay, Mark Schroeder, and Ralph Wedgwood for valuable discussions and detailed comments on earlier versions of this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Audi, R. (1994). Dispositional beliefs and dispositions to believe. *Noûs*, 28(4), 419–434.
- Bartlett, G. (2018). Occurrent states. *Canadian Journal of Philosophy*, 48(1), 1–17.
- Blackburn, S. (1998). *Ruling passions*. Oxford: Oxford University Press.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Broome, J. (1999). Normative requirements. *Ratio*, 12(4), 398–419.
- Broome, J. (2005). Does rationality give us reasons? *Philosophical Issues*, 15(1), 321–337.
- Broome, J. (2007). Is rationality normative? *Disputatio*, 2(23), 161–178.
- Broome, J. (2013). *Rationality through reasoning*. Oxford: Wiley-Blackwell.
- Brunero, J. (2010). The scope of rational requirements. *Philosophical Quarterly*, 60(238), 28–49.
- Christensen, D. (2007). Does Murphy's law apply in epistemology?: Self-doubt and rational ideals. In T. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (Vol. 2, pp. 3–30). Oxford: Oxford University Press.
- Davidson, D. (2001). *Essays and actions and events*. Oxford: Clarendon Press.
- Davidson, D. (2004). *Problems of rationality*. Oxford: Clarendon Press.
- Dennett, D. (1989). *The intentional stance*. Cambridge, MA: MIT Press.
- Egan, A. (2008). Seeing and believing: Perception, belief formation, and the divided mind. *Philosophical Studies*, 140(1), 47–68.
- Elga, A., & Rayo, A. (2015). Belief fragmentation. Unpublished draft of March 1, 2015.
- Finlay, S. (2009). Against all reason?: Skepticism about the instrumental norm. In C. Pigden (Ed.), *Hume on motivation and virtue* (pp. 155–178). Cambridge: Palgrave MacMillan.
- Fischer, J. M., & Tognazzini, N. (2011). The physiognomy of responsibility. *Philosophy and Phenomenological Research*, 82(2), 381–417.

- Fogal, D. (2019). Rational requirements and the primacy of pressure. *Mind*. <https://doi.org/10.1093/mind/fzz038>.
- Friedman, J. (2019). Inquiry and belief. *Noûs*, 53(2), 296–315.
- Harman, G. (1976). Practical reasoning. *Review of Metaphysics*, 29(3), 431–463.
- Harman, G. (1986). *Change in view*. Cambridge, MA: MIT Press.
- Hedden, B. (2016). Mental processes and synchronicity. *Mind*, 125(499), 873–888.
- Hieronymi, P. (2005). The wrong kind of reason. *Journal of Philosophy*, 102(9), 437–457.
- Hieronymi, P. (2006). Controlling attitudes. *Pacific Philosophical Quarterly*, 87(1), 45–74.
- Kiesewetter, B. (2017). *The normativity of rationality*. Oxford: Oxford University Press.
- Kolodny, N. (2005). Why be rational? *Mind*, 114(455), 509–563.
- Kolodny, N. (2008a). Why be disposed to be coherent? *Ethics*, 118(3), 437–463.
- Kolodny, N. (2008b). The myth of practical consistency. *European Journal of Philosophy*, 16(3), 366–402.
- Korsgaard, C. M. (1997). The normativity of instrumental reason. In G. Cullity & B. Gaut (Eds.), *Ethics and practical reason* (pp. 215–254). Oxford: Oxford University Press.
- Lewis, D. (1982). Logic for equivocators. *Noûs*, 16(3), 431–441.
- Lord, E. (2018). *The importance of being rational*. Oxford: Oxford University Press.
- McCann, H. J. (1991). Settled objectives and rational constraints. *American Philosophical Quarterly*, 28(1), 25–36.
- Mele, A. (1992). *Springs of action*. Oxford: Oxford University Press.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.
- Parfit, D. (2011). *On what matters* (Vol. 1). Oxford: Oxford University Press.
- Podgorski, A. (2017). Rational delay. *Philosopher's Imprint*, 17(5), 1–19.
- Railton, P. (1997). On the hypothetical and non-hypothetical in reasoning about belief and action. In G. Cullity & B. Gaut (Eds.), *Ethics and practical reason* (pp. 53–79). Oxford: Oxford University Press.
- Reisner, A. (2013). Is enkratic principle a requirement of rationality? *Organon F*, 20(4), 436–462.
- Ridge, M. (2014). *Impassioned belief*. Oxford: Oxford University Press.
- Scanlon, T. M. (1998). *What we owe to each other*. Cambridge, MA: Harvard University Press.
- Scanlon, T. M. (2007). Structural irrationality. In G. Brennan, et al. (Eds.), *Common minds: Themes from the philosophy of Philip Pettit* (pp. 84–103). Oxford: Clarendon Press.
- Schroeder, M. (2004). The scope of instrumental reason. *Philosophical Perspectives*, 18, 337–364.
- Schroeder, M. (2009). Means-end coherence, stringency, and subjective reasons. *Philosophical Studies*, 143(2), 223–248.
- Setiya, K. (2004). Against internalism. *Noûs*, 38(2), 266–298.
- Setiya, K. (2007). Cognitivism about instrumental reason. *Ethics*, 117(4), 649–673.
- Sinhababu, N. (2011). The Humean theory of practical irrationality. *Journal of Ethics and Social Philosophy*, 6(1), 1–13.
- Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.
- Velleman, D. (1989). *Practical reflection*. Princeton, NJ: Princeton University Press.
- Wallace, R. J. (2001). Normativity, commitment, and instrumental reason. *Philosopher's Imprint*, 1(4), 1–26.
- Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, 24(2), 227–248.
- Way, J. (2011). The symmetry of rational requirements. *Philosophical Studies*, 155(2), 227–239.
- Way, J. (2018). Reasons and rationality. In D. Star (Ed.), *Oxford handbook of reasons and normativity* (pp. 485–503). Oxford: Oxford University Press.
- Wedgwood, R. (2009). The normativity of the intentional. In A. Beckermann & B. P. McLaughlin (Eds.), *The Oxford handbook of the philosophy of mind* (pp. 421–436). Oxford: Clarendon Press.
- Wedgwood, R. (2013). Rational 'ought' implies 'can'. *Philosophical Issues*, 23(1), 70–92.
- Worsnip, A. (2016). Moral reasons, epistemic reasons, and rationality. *Philosophical Quarterly*, 66(263), 341–361.
- Worsnip, A. (2018). What is (in)coherence? In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 13, pp. 184–206). Oxford: Oxford University Press.