

## “Rational Agency and the Struggle to Believe What Your Reasons Dictate”

Forthcoming in Borgoni, Kindermann, and Onofri, eds., *The Fragmented Mind* (Oxford UP)

Brie Gertler, University of Virginia

*Penultimate version.*

*Abstract.* According to an influential view that I call *agentialism*, our capacity to believe and intend directly on the basis of reasons—our *rational agency*—has a normative significance that distinguishes it from other kinds of agency. Agentialists maintain that insofar as we exercise rational agency, we bear a special kind of responsibility for our beliefs and intentions, and those attitudes are truly our own. In this paper I challenge these agentialist claims. My argument centers on a case in which a thinker struggles to align her belief to her reasons, and succeeds only by resorting to non-rational methods. I argue that she is responsible for the attitude generated by this struggle; that this process expresses her capacities for rationality and agency; and that the belief she eventually arrives at is truly her own. So rational agency is not distinctive in the ways that agentialists contend.

### 1. Introduction

Our beliefs are often directly sensitive to evidence. For example, your belief that your neighbor is out of town is immediately revised when you see her on your street. But sometimes our beliefs resist the force of our reasons. Emotional factors might make it difficult for you to accept that your friend has committed a serious crime, even while you recognize that the evidence strongly points to his guilt. Similarly, our intentions are often directly sensitive to our reasons, but at times we have to struggle to commit to what we recognize as the best course of action.

The difference between believing or intending directly on the basis of reasons, and having to make an effort to align one’s attitudes with one’s reasons, is central to an influential view that I call *agentialism* (Bilgrami 2006, Boyle 2011, Burge 1996, Korsgaard 1996, Moran 2001). Agentialists maintain that our capacity to believe and intend *directly* on the basis of reasons—what I’ll call *rational agency*—has a special normative significance, a significance that distinguishes it from other kinds of agency. Insofar as we exercise rational agency, we bear a special kind of responsibility for our beliefs and intentions; and it is in virtue of our rational agency that these attitudes are truly our own.

My main objective here is to question the idea that rational agency has this distinctive normative significance. I will take it as given that having beliefs that match one’s evidence, and intentions that match one’s best reasons for acting, is usually beneficial; so an effective means of achieving this alignment between attitudes and reasons is instrumentally valuable. I will also grant that having attitudes that conform to reasons *directly* is the most reliable and efficient way to achieve alignment between attitudes and reasons. So rational agency is preferable, on instrumental grounds, to alternative ways that attitudes can come to be aligned with reasons. But I will challenge the agentialist claim that rational agency’s normative significance, relative to other ways of aligning attitudes with reasons, transcends this comparative instrumental value.

I will focus on beliefs. My argument centers on a case in which a character named Diane struggles to bring a recalcitrant belief into alignment with her evidence. This effort eventually succeeds. But it is not an exercise of rational agency, as the belief is only indirectly shaped by her reasons: she conforms her belief to her evidence only by exploiting a non-rational process. I will

argue that Diane is responsible for the attitude generated by this struggle; her success in revising her belief expresses her capacities for rationality and agency; and the belief she eventually arrives at is truly her own. So rational agency is not distinctive in the ways that agentialists contend.

The paper proceeds as follows. Section 2 explains how recalcitrant beliefs, which resist the force of reasons, constitute a kind of cognitive fragmentation that falls short of what is rationally ideal. Section 3 sketches the agentialist notion of rational agency and formulates the agentialist thesis I will challenge. Section 4 introduces the case of Diane. Section 5 considers whether Diane's recalcitrant attitude is properly classified as a *belief*. Section 6 argues that, although Diane's means of reshaping her attitude does not qualify as an exercise of rational agency, it is nonetheless rational: in this section I address the idea that Diane's approach to her attitude involves the "wrong kind of reasons". Sections 7 and 8 argue that Diane exercises responsible agency in revising her attitude, and that the revised attitude is truly her own. Section 9 examines whether, on pain of regress, all cases of deliberation must involve some direct shaping of attitudes by reasons. Section 10 responds to the claim that we cannot coherently embark on a process of deliberation unless we assume that this process will directly issue in a belief or intention.

I conclude that working to overcome recalcitrant attitudes is as rational and agency-involving as exercises of "rational agency", that we are responsible for attitudes formed through such efforts in the same way that we are responsible for those attitudes that conform to our reasons directly, and that attitudes need not be an expression of rational agency in order for the thinker to regard them as truly her own.

## 2. Recalcitrant beliefs and rationality

We all have some beliefs (or belief-like attitudes<sup>1</sup>) that influence our reasoning and affect our behavior, but resist the force of countervailing evidence and clash with explicit judgments that we readily avow. Call these *recalcitrant beliefs*. Some recalcitrant beliefs, such as beliefs in racial or gender stereotypes, are disturbing. But many are relatively mundane. Consider a tennis player who takes a detour on the way to an important match to retrieve his "lucky" socks, and who attributes some of his past losses to having not worn the socks during those matches. He might nonetheless fully recognize that the belief that those socks bring him luck is a baseless superstition, unsupported by his overall win-loss record. And he might wish that he could shake that belief even as he returns home for the socks, delaying his arrival at the match. This is a case of recalcitrant belief.

Recalcitrant beliefs involve a kind of fragmentation in that they are not integrated with the thinker's considered judgments or her sense of her reasons. This kind of fragmentation differs from believing that *p* while merely *possessing* strong evidence that not-*p*. (On fragmentation of that sort, see the contributions to this volume by Borgoni and Yalcin.) In the cases that concern me here, an attitude fails to conform to *what the thinker regards as* her reasons: e.g., the tennis player who believes the socks are lucky yet takes his evidence to favor "there are no lucky socks". The senses of "reasons" and "evidence" relevant to this discussion are limited

---

<sup>1</sup> There is no consensus as to whether the recalcitrant attitudes I describe are genuine *beliefs*. As I explain in Section 5 below, I remain neutral on that question. But for convenience I will use "recalcitrant belief" liberally, to refer to recalcitrant beliefs or belief-like attitudes.

to *what the thinker regards as* her reasons and her evidence. That is how I will use these terms throughout the paper.

### 3. Rational Agency

All of us exhibit some fragmentation of this sort. We have all had—and probably all currently *have*—recalcitrant beliefs. But each of us sometimes believes exactly what we take the totality of our evidence to support, and do so on the basis of our assessment of the evidence. Agentialism accords particular significance to our capacity to believe directly on the basis of reasons. Agentialists maintain that, in having beliefs that are directly reasons-responsive, we exercise a distinctive kind of agency, *rational agency*, that has special normative significance.

A belief expresses rational agency only if it is directly sensitive to reasons. Key passages from Moran and Boyle will help to illustrate the sense of directness at work here, and illuminate the significance that these two agentialists, in particular, attribute to rational agency.<sup>2</sup>

Moran describes the exercise of rational agency, whereby rational deliberation issues in attitudes directly, as an exercise of *authority* over one's attitudes. He contrasts this with merely *controlling* an attitude through a deliberate strategy for making one's reasons effective.

[I]t is because the deliberator declares the authority of reason over his thought and action that at the conclusion of his thinking there is no further thing he does to make that conclusion his actual belief or his intention. ... [I]t would be an expression of the failure of reasoning were it to terminate not in conviction itself but rather in my apprehending a particular thought, or even appraising it as best, which I then need to find some *way* to make my own, my actual state of mind. (As in other straits, having lost authority, I could only fall back on *control*.) (Moran 2001, 131)

Boyle shares this outlook. He says that “believing itself is an exercise of agency” involving “active ongoing governance”, where such governance over a state of affairs means that “I do not control it extrinsically by doing things to affect its unfolding” (Boyle 2009, 121, 143). This contrast between active governance and mere (extrinsic) control dovetails with Moran's distinction between authority and control. Boyle argues that the way that we hold each other accountable for our beliefs reflects the presumption that believers exercise direct authority over their beliefs, in contrast to mere control:

Finally, it does not seem that we merely hold [the believer] accountable in the manner of someone who might do something about a given situation, as I might be held accountable for the misbehavior of my child, or the explosion of the munitions in my basement. I am not merely accountable for allowing an unreasonable belief to persist, or for having previously brought such a belief into existence; I am myself directly accountable for now holding the belief — for presently taking things to be thus-and-so, in the context of the reasons available to me. (Boyle 2011, 11)

---

<sup>2</sup> Bilgrami, Burge, and Korsgaard hold views about the significance of rational agency that are similar to these, at least in spirit. I focus on Moran and Boyle because their characterizations of rational agency are especially clear and explicit.

Boyle calls this authority over our beliefs “doxastic self-determination”; Moran says that shaping our attitudes through deliberation amounts to “self-constitution”. For Moran and Boyle, the exercise of rational agency defines us as the kinds of creatures we are: namely, creatures who are self-determining, in that we have the authority—roughly, the right, power, and responsibility—to govern our attitudes by our reasons.<sup>3</sup>

It is important to understand exactly what Moran and Boyle are claiming here. It is not simply that the capacity to form judgments on the basis of reasoning is essential to our identity as rational thinkers. Judgments are occurrent states: ontologically speaking, they are *events*—though on some views judgments are defined as occurrent states that (typically) yield beliefs. In any case, what is at issue here are beliefs, which are at least partly constituted by dispositions to act, reason, and feel in certain ways. (I discuss the nature of belief in Section 5.)

Moran and Boyle are claiming that we are self-determining creatures in virtue of our capacity to form beliefs and intentions directly on the basis of reasoning. Recalcitrant beliefs represent a failure to exercise our capacity for self-determination. If I manage to overcome a recalcitrant belief through deliberate effort, the need to engage in effort means that my coming to believe in line with my reasons is not an exercise of self-determination, as I merely “control” the belief, and I am not truly its “author”.

Having attitudes that are *directly* reasons-responsive is an especially reliable and efficient way to secure alignment between attitudes and reasons. So rational agency is instrumentally valuable. But the agentialist claims that rational agency is not just more reliable or efficient than alternative routes to achieving alignment between attitudes and reasons; it has a special normative significance. This idea is captured in the following thesis.

### **The Normative Significance of Rational Agency:**

The normative significance of rational agency transcends its instrumental value as an especially effective means of achieving alignment between attitudes and reasons. When a thinker believes or intends *directly* on the basis of reasons, she is the responsible author of that attitude. We are not responsible authors of our attitudes (in an equally robust sense) when we manage to align them to our reasons only through deliberate effort, that is, through mere “control”.

This is the thesis I will challenge here.<sup>4</sup>

---

<sup>3</sup> Similarly, Burge ties our status as *selves* to our capacity to engage in what he calls “critical reasoning”, which involves rational agency. “Selves are critical reasoners” (Burge 2011, 331). He envisions a thinker who recognizes that he has an attitude that resists the force of his reasons, and so “must face a question of how, by what means, to make [his] reasons effective” (Burge 1996, 113). That the thinker faces this question means that the relation between reasons and attitude is not *direct* in the sense required for critical reasoning. “In critical reasoning, such questions of means and control do not arise” (*ibid.*), for attitudes are directly sensitive to reasons. A reliance on a deliberate strategy for making one’s reasons effective—that is, a failure to exercise rational agency—means that the transformation in one’s belief is not critical reasoning. By definition, selves (the kinds of rational creatures we are) must be capable of critical reasoning and therefore of rational agency.

<sup>4</sup> In previous work, I’ve challenged a related agentialist claim: namely, that rational agency has special normative significance in that it is only in virtue of our rational agency that we possess certain epistemic rights and responsibilities (Gertler 2018).

## 4. Struggling with a recalcitrant attitude: the case of Diane

### 4.1. The case of Diane

My discussion throughout the paper will center on the following case, in which a subject named Diane struggles with a recalcitrant attitude.

*Stage 1.* Diane finds the arguments against the existence of free will compelling. On the basis of these arguments, she concludes that humans are not morally responsible for their actions: they don't deserve praise or blame, reward or punishment. She tries to persuade her friends that praise and blame are misguided, and she resolves to stop condemning and lionizing others (or herself). Yet Diane remains prone to reactive attitudes such as resentment, especially as regards environmental issues. When she sees an oil company executive leaving his parked SUV one day, she feels a strong sense of moral indignation and indulges the urge to scratch the car with her keys. As she does so, she thinks to herself "he's got it coming!"

*Stage 2.* Diane reflects on this incident and on other times she's experienced condemnatory thoughts and retributive impulses. She comes to recognize that her overall attitude about moral responsibility is not in line with the skepticism that she endorses in her more contemplative moments. She resolves to work on bringing her attitude into alignment with her reasons by doing what she can to quell her moral indignation and her condemnatory thoughts and feelings. Towards this end, she takes up mindfulness meditation. In mindfulness meditation, one practices observing one's thoughts and feelings with detachment. This reportedly diminishes the tendency to regard these affective states as accurate or justified, and neutralizes their emotional force.

*Stage 3.* Now a seasoned meditator, Diane no longer feels self-righteousness or moral indignation. She is no longer inclined to blame people for their actions or to regard them as morally responsible.

At Stage 1, Diane has a recalcitrant attitude about moral responsibility: her tendency to blame and punish resists the force of her reasons. By Stage 3, she has brought her attitude into line with her reasons, which favor skepticism about moral responsibility. (I'll use "MR" to refer to moral responsibility throughout the paper.) This alignment results from an extended, deliberate process, in which Diane makes use of non-rational methods to shift her attitude. So although Diane's Stage 3 attitude about MR is shaped by her reasons, this shaping process is indirect: her attitude is not directly responsive to her reasons.

The process by which Diane's reasons shape her attitudes is not purely rational, since mindfulness meditation is not a process of deliberation or reflection on reasons. It is a non-rational cognitive practice aimed at blunting the tendency to automatically endorse our thoughts and to regard our feelings as justified. There are some indications that over time, a regular meditation practice weakens the emotional pull of passing thoughts and feelings.<sup>5</sup>

---

<sup>5</sup> Goleman and Davidson (2017) outline some of the research suggesting that a regular meditation practice enhances emotional regulation and lessens reactivity. Insofar as mindfulness meditation weakens the emotional pull of thoughts and feelings, it arguably increases the relative influence of rational factors, such as evidence, over the meditator's attitudes. Still, mindfulness meditation is itself a non-rational exercise: it does not involve evaluating thoughts or feelings to determine whether they are rationally justified or fitting.

What matters is not whether mindfulness meditation actually has this effect. What matters is that Diane takes herself to have good evidence that it does, and that after practicing mindfulness meditation her attitude about MR is aligned with her reasons. This may be because that method does work; alternatively, it may be a kind of placebo effect.

#### 4.2. *Diane's shift in attitude is not an expression of rational agency*

In undertaking meditation so as to conform her attitude about MR to her reasons, Diane violates the conditions on exercising rational agency given by agentialists.<sup>6</sup> As Moran describes the exercise of rational agency in deliberation, “there is no further thing [the deliberator] does to make [his] conclusion his actual belief or his intention” (Moran 2001, 131). But after deliberating about the issue of MR, there *is* a further thing Diane does to adjust her attitude so that it fits with her conclusion. Nor does Diane exercise the kind of “active ongoing governance” involved in rational believing on Boyle’s view (Boyle 2009, 143). She controls her attitude “extrinsically”, reshaping it through her meditation practice.<sup>7</sup>

Diane’s efforts do not qualify as an exercise of rational agency because her reasons don’t shape her attitudes *directly*. What makes the process indirect in the relevant sense is that, in order to bring her attitude into alignment with her reasons, she deliberately engages in a non-rational process, viz., mindfulness meditation.

Now even in exercises of rational agency, there are presumably various non-rational factors that mediate between reasons and attitudes. Let’s assume that Diane’s grasp of the evidence for MR skepticism is somehow realized in her neurophysiology. The process by which this neurophysiological event yields a belief that we lack MR may involve other neurophysiological factors that mediate between that event and the formation of the belief. But such indirectness is beside the point. Reasons may directly shape a belief in the sense required for rational agency even if non-rational processing, e.g. at the neurophysiological level, causally mediates between (the recognition of) evidence for *p* and the belief that *p*. The agentialist denies that Diane’s way of transforming her attitude is an exercise of rational agency not because there is some non-rational factor causally mediating between her reason and her attitude, but because Diane transforms her attitude only by deliberately resorting to a non-rational process.

The agentialist sees Diane’s means of transforming her attitude as problematic because it involves Diane’s deliberately engaging in a non-rational process to bring about the desired result. Now since Diane’s method for revising her attitude takes time, and is presumably less reliable than the exercise of rational agency, Diane’s indirect method is instrumentally inferior to an exercise of rational agency. But the agentialist holds that Diane is not the *responsible author* of this

---

<sup>6</sup> This is not to say that Diane doesn’t exercise rational agency at all. In working through the arguments for MR skepticism, she may well form beliefs directly on the basis of reasons: e.g., she may accept incompatibilism directly on the basis of evaluating the Consequence Argument. But that is just to say that her *judgment* about MR directly conforms to her reasons. My concern is with how she achieves alignment between her *attitude* (her belief or belief-like attitude) and her reasons. The shift in Diane’s attitude about MR from Stage 1 to Stage 3, towards greater alignment with her reasons, is not purely an exercise of rational agency.

<sup>7</sup> Diane similarly violates Burge’s conditions on critical reasoning. Her method does not qualify as critical reasoning because she “face[s] the question how, by what means, to make [her] reasons effective” (Burge 1996, 113), and relies on a deliberate strategy for accomplishing this, namely, mindfulness meditation.

attitude, as she would be if it were directly responsive to her reasons. That agentialist claim, the Normative Significance of Rational Agency thesis, is my primary concern here.

Before mounting my challenge to that thesis, I will digress to consider whether Diane's initial, recalcitrant attitude about MR is a genuine *belief*.

### 5. Is Diane's recalcitrant attitude a *belief*?

At Stage 1, Diane has a recalcitrant attitude about MR. This attitude is at odds with her considered judgment about MR. It is linked to her emotions (righteous indignation at the oil company executive) and it influences her reasoning (thinking that he deserves this ill treatment) and her behavior (scratching the car). Is this recalcitrant attitude a *belief* that we possess MR? And does Diane *believe* in MR skepticism at Stage 1, or does she acquire this belief only at Stage 3? The answers to these questions depend, of course, on how we understand belief.

I favor Schwitzgebel's approach to belief: "To believe that *p* ... is nothing more than to match to an appropriate degree and in appropriate respects the dispositional stereotype for believing that *p*" (Schwitzgebel 2002, p. 253). The dispositional stereotype for believing that *p* includes being disposed to act and to reason as if *p* is true. For example, to believe that Miguel is discreet is to be disposed to share confidences with him, to assure others that he can keep a secret, and to infer, when Miguel reports a bit of sensitive news, that he was not asked to keep this information to himself.<sup>8</sup> The stereotype also includes affective dispositions. If you believe that Miguel is discreet, you feel at ease talking with him, and would feel surprise upon learning that he had betrayed a confidence. These dispositions are linked with occurrent thoughts and judgments—at least, if the belief is not recalcitrant. Dispositions to act and reason *as if* Miguel is discreet generally include a disposition to endorse the thought that he is (when the question arises), and are generally sensitive to occurrent judgments bearing on his discretion: e.g., if you heard him gossiping and judged *Miguel is not so discreet after all*, you would no longer be disposed to feel, act, and reason as if he were discreet.

At Stage 1, Diane has some aspects of the dispositional stereotype of belief as regards *we do not possess MR*. She is disposed to assert "we are never morally responsible", and to disagree with MR realists. She is disposed, upon recognizing that retributive considerations affect criminal sentencing, to conclude that our justice system is unjust. But in some other respects, she is disposed to feel, act, and reason as if *we do* have MR: e.g., to feel indignation towards the oil company executive, to blame him and try to punish him, and to think that he deserves comeuppance. Importantly, MR realism's grip on Diane is not merely affective. She is disposed to judge that certain actions merit reward or retribution, and these judgments may be sensitive to evidence, albeit imperfectly so.

The dispositional stereotype approach to belief does not settle the question as to what, precisely, Diane believes about MR at Stage 1. Settling that question would require deciding what's needed to match the stereotype "to an appropriate degree and in appropriate respects" (*ibid.*). There is no consensus on that issue, and I will not take a position on it here. What's important is that, on the dispositional stereotype approach, "belief" is not an all-or-nothing

---

<sup>8</sup> This characterization assumes a background of relevant desires and other attitudes.

affair. Instead, there is a multiplex spectrum of possible attitudes towards  $p$ , constituted by the extent to which the thinker is disposed to feel, act, and reason in certain ways.

The table below loosely represents the spectrum of attitudes corresponding to the presence or absence of the relevant dispositions.

		“Aliefs” (Gendler 2008)	Implicit attitudes	Clear cases of belief	Beliefs as held by an ideal rational thinker
<b>A</b>	Structure	Associative	Propositional	Propositional	Propositional
<b>B</b>	Associated dispositions	Behavioral, affective	Behavioral, affective	Behavioral, affective	Behavioral, affective
<b>C</b>	Evidence sensitivity	Little sensitivity to evidence			Perfectly attuned to evidence
<b>D</b>	Inferential promiscuity	Affect few relevant inferences appropriately			Affect all relevant inferences appropriately
<b>E</b>	Relation to occurrent judgments	Frequently conflict with occurrent judgments			Always match occurrent judgments

Attitudes on the left end of this table are not beliefs, whereas attitudes on the right end *are* beliefs—that much is relatively uncontroversial. But it is controversial how much evidence sensitivity, inferential promiscuity, and alignment with occurrent judgments are minimally required for belief. I remain neutral on those questions. So I remain neutral on the question where, in rows C through E, the border between non-beliefs and beliefs falls.<sup>9</sup>

On the issue of MR, Diane’s dispositions at Stage 1 are mixed. Some correspond to MR skepticism, while others correspond to a belief in MR. Because of my neutrality about the minimally sufficient conditions for belief, I remain neutral as to how to characterize Diane’s situation at Stage 1. Depending on how one construes belief, she may have two conflicting beliefs, that is, she may both believe in MR skepticism and believe in MR realism; she may have neither of these beliefs, as her attitudes towards MR skepticism and MR realism may each fall short of belief; she may have exactly one of these beliefs, in skepticism or realism, in which case

---

<sup>9</sup> An especially illuminating example of disagreement as to what is minimally required for belief is the debate between Levy (2015) and Mandelbaum (2016). Mandelbaum argues that implicit attitudes, which fall towards the left end of the spectrum in rows C-E, are beliefs. He says that, unlike mere associations, implicit attitudes are responsive to argument and inferentially promiscuous, and can’t be eliminated by counterconditioning. Levy acknowledges that implicit attitudes have some features of belief: they are propositionally structured, and contribute to some inference-like transitions on the basis of their (propositional) content. But he thinks that their reasons-sensitivity and role in inferences are too limited and fragmented to qualify them as full-fledged beliefs. He calls them “patchy endorsements”.

that belief is diluted by its partial insensitivity to evidence or by the presence of divergent dispositions.

What interests me is Diane's overall attitude towards *we possess MR*—that is, the entire set of varied, sometimes competing dispositions that she has, relevant to *we possess MR*—and how that overall attitude shifts with her deliberate effort to transform it. The key point is this. At Stage 1, Diane's overall attitude concerning the question of MR is far from that of an ideal rational thinker, one whose dispositions mesh perfectly and are perfectly attuned to her evidence—represented by the rightmost column of **Table 1**. (To sidestep certain controversies unrelated to my central point here, I will stipulate that both Diane and the “ideal rational thinker” are in situations that are ideal from the perspective of rationality in that they do not face the kind of problem introduced by higher-order evidence.<sup>10</sup>) At Stage 3, Diane's overall attitude towards *we possess MR* is far more cohesive and more aligned with her evidence than it was at Stage 1. Her attitude more closely approximates a belief held by an ideal rational thinker.

My point is that, by deliberately engaging in a non-rational process, Diane shifts her attitude about MR so that she comes to approximate the rational ideal with respect to this attitude. Her attitude about MR shifts into alignment with her evidence. Moreover, Diane's belief in MR skepticism matches her occurrent judgments about that issue, and it is poised to contribute to inferences and thereby to serve as a reason that shapes other attitudes. And Diane is no longer disposed to act, reason, and feel as if MR realism were true. So at Stage 3, Diane's MR skepticism at least approximates the profile of an attitude in an ideal rational thinker.

Some agentialists will balk at the dispositional approach to beliefs, claiming that it assimilates all beliefs to what are in fact marginal cases. According to these agentialists, the core notion of belief is the notion of a judgment-sensitive doxastic *commitment*. Coliva nicely outlines what this notion involves. In her view, doxastic commitments (a) result from the assessment of evidence, (b) are constrained by norms of reasoning, (c) are regarded by the subject as constrained by norms of reasoning, and (d) are states for which the subject is held rationally responsible (paraphrased from Coliva 2012, 219). Clearly, Diane's overall attitude towards MR at Stage 1 does not qualify as a doxastic commitment, as some of its elements resist the force of evidence and it is not adequately constrained by norms of reasoning, in particular by the norm that beliefs should align with evidence.

But Diane's attitude at Stage 3, which approximates a belief held by an ideal rational thinker, arguably does qualify as a doxastic commitment. (The requirement that doxastic commitments are shaped *directly* by reasons would of course beg the question in this context.) I will argue that Diane's revision of her attitude—the process she uses to shift from Stage 1 to Stage 3—is guided by her assessment of the evidence and her application of the norm that beliefs should align with evidence, and that she is rationally responsible for applying this norm

---

<sup>10</sup> On the problem of higher-order evidence, see Roush 2017. Although I don't have the space to discuss this here, I believe that the spirit of agentialist views like Moran's and Boyle's favors what Horowitz calls the Non-Akrasia Constraint: “it can never be rational to have high confidence in something like, ‘P’, but my evidence doesn't support P’.” (Horowitz 2014, 718) If that is correct, my stipulation that an ideal rational thinker does not face the problem of higher-order evidence is unneeded in the current context, for the agentialist will grant that, for an ideally rational thinker, evidence and rational beliefs will not diverge in the way envisioned by those who reject the Non-Akrasia Constraint.

and for her Stage 3 attitude. So although Diane makes use of non-rational means to revise her attitude, that revision process has whatever normative significance attaches to exercises of rational agency.

## 6. Diane's attitude revision is rational

### 6.1 *Diane's attitude revision appears to be rational*

Diane's decision to reshape her attitude, and the overall process by which she accomplishes this, are *rational* in various respects. Careful deliberation has led her to conclude that the evidence—in the form of philosophical arguments—favors MR skepticism. Her commitment is to satisfying the rational norm that beliefs should conform to evidence is what drives her to take up mindfulness meditation when she comes to recognize that her attitude about MR is in the grip of affective factors (given her evidence that mindfulness meditation loosens the grip of affective factors). She comes to wholeheartedly embrace MR skepticism because this is the position her reasons dictate. Her efforts eliminate dispositions that are in tension with her evidence, such as the disposition to blame people for their actions. Finally, the connection between Diane's evidence and her attitude doesn't involve any deviant causal links of the sort that would prevent her evidence from rationalizing her attitude (Davidson 1973). So Diane's evidence for MR skepticism both explains and rationalizes her belief in MR skepticism.

### 6.2 *The role of practical reasoning*

Agentialists can agree that Diane's reshaping of her attitude is rational in these respects. But they will deny that it is an exercise of rational agency, since she manages to conform her belief to her reasons only by deliberately engaging in the non-rational exercise of mindfulness meditation. In Diane's case, the indirect route from reasons to belief involves the use of *practical* reasoning. Her decision to take up mindfulness meditation is precisely a response to the problem Moran describes: the "need to find some *way* to make [my conclusion] my own, my actual state of mind" (Moran *ibid.*). And meditation is precisely a means of controlling her mental state "extrinsically", that is, outside the space of reasons (Boyle *ibid.*). That practical reasoning plays this role in Diane's efforts means that the belief resulting from those efforts does not express rational agency.

What differentiates Diane's method from the exercise of rational agency is the specific role that practical reasoning plays; it is not simply that practical reasoning is involved. One can exercise rational agency even when relying on practical reasoning as part of a larger deliberation aimed at discovering what is true. Suppose I'm asked where I'll be at 6 pm. I consult the weather forecast, and on that basis I decide to leave the office by 5 pm, so as to arrive home before an expected thunderstorm. This bit of practical reasoning about when to depart informs my theoretical reasoning, leading me to judge that I'll be home at 6 pm. So long as that judgment directly results from my theoretical reasoning, it is an expression of rational agency.<sup>11</sup>

---

<sup>11</sup> Similarly, one can exercise rational agency even when relying on theoretical reasoning in a larger deliberation aimed at deciding what to do. In deciding when to leave, I may rely on theoretical reasoning about my chance of getting caught in a thunderstorm if I were to leave at 5:15 given that the forecast predicts that the storm will begin at 5:30. This bit of theoretical reasoning then informs my practical reasoning, leading me to decide to leave by 5 pm. So long as that decision results directly from my practical reasoning, it is an expression of rational agency.

In the case just described, practical reasoning's role in generating or sustaining belief is limited to providing reasons bearing on the truth of the belief. My decision to leave the office by 5 pm provides evidence (via my awareness of this decision) about where I will be at 6 pm. Since my practical reasoning serves only as a source of evidence in this case, it does not mediate between my evidence and my belief. By contrast, Diane's practical reasoning mediates between her evidence and her belief, for it concerns *how to bring her attitude into line with her reasons*. That issue—and hence, the result of her practical deliberation—does not bear on the truth of her belief, which concerns whether we possess moral responsibility.

### 6.3 *The wrong kind of reasons?*

As the case of Diane illustrates, a belief can be rationalized and justified by reasons, and held on the basis of those reasons, even when the reasons' influence on the belief is indirect. Why think that believing indirectly on the basis of suitable reasons is somehow problematic, rather than simply less efficient and reliable than having beliefs that are directly reasons-responsive?

One familiar concern is that, when practical reasoning contributes to a belief-forming process, the resulting belief may be based on the “wrong kind of reasons”.<sup>12</sup> Diane's eventual wholehearted belief in MR skepticism would face this problem if, for example, her reason for endorsing MR skepticism (and thus for taking steps to reshape her attitude) was that she determined that this would improve her tenure chances, since her colleagues heavily favor MR skepticism. But in the scenario as described Diane's eventual wholehearted belief in MR skepticism does not rest on such “wrong kind” reasons. Her reasons exclusively concern the issue of MR, and those reasons bear on the *truth* about that issue.<sup>13</sup>

Still, I suspect that the “wrong kind of reasons” worry may be playing a more subtle role here. The question of how to bring an attitude into alignment with your reasons arises only when your attitudes are not perfectly integrated with your reasons—that is, only when you fail to exercise rational agency. In wrestling with that question, the guiding principle “fit attitudes to reasons” becomes more or less explicit. And awareness of that principle might naturally lead you to consider whether there are alternative principles you might adopt in shaping your attitudes: e.g., a pragmatic principle such as “strive for attitudes that maximize happiness”. In short, a failure of rational agency, and the resulting recognition of “fit attitudes to reasons” as a guiding principle, can open the conceptual space for challenges to the assumption that conformance to reasons is the *sine qua none* of attitudes such as beliefs and intentions.

However, although reflection on the goal of fitting attitudes to reasons may reveal the conceptual possibility of alternative principles for attitude formation, it does not cast doubt on the presumption that conformance to reasons (of the right kind) is the sole factor by which attitudes should be evaluated. So it does not support the idea that an alternative, possibly pragmatic principle could be legitimate. Engaging in practical reasoning as part of the process of aligning attitudes to reasons, as Diane does, could perhaps inspire the thinker to entertain alternative principles for shaping attitudes. But that is a psychological phenomenon, not grounds

---

<sup>12</sup> For a useful discussion of the “wrong kinds of reasons” debate, see Gertken and Kiesewetter 2017.

<sup>13</sup> For a forceful challenge to the idea that consequentialist and other considerations *are* wrong kinds of reasons, see Reisner (2018).

for an objection to the idea that beliefs should be shaped exclusively by evidence. The need to figure out how to align one's attitudes to one's reasons does nothing to undermine the idea that attitudes should be evaluated solely by their conformance to the right kind of reasons.

#### 6.4 *Practical vs. epistemic rationality*

The agentialist may object that while Diane's revision of her belief is rational, the kind of rationality it exhibits is *practical* rationality, whereas the agentialist's concern is with *epistemic* rationality. Tristram McPherson has raised this objection in conversation, using an example he calls Bad Diane. Bad Diane's attitude about MR is exactly similar to Diane's attitude at Stage 1: it is the same mix of dispositions corresponding to skepticism and realism about MR. Like Diane, Bad Diane recognizes that her attitude about MR is not aligned with her evidence. But this recognition does not move Bad Diane to action. She feels no concern about it whatsoever.

According to the objection at hand, Diane and Bad Diane are on a par as regards *epistemic* rationality. They conform to epistemic norms in precisely the same way, and to precisely the same degree, in evaluating the arguments for MR skepticism and in assessing the relation between their attitude about MR and their evidence. Diane is more rational only as regards practical rationality: she forms intentions that conform to what her reasons dictate, namely, to (try to) align her beliefs with her evidence. Bad Diane has equal reason to try to align her beliefs with her evidence, so her failure to form the appropriate intentions is a failure of practical rationality.

This point constitutes a challenge to my Diane example only insofar as that example requires that Diane's way of aligning her attitude with her reasons—her shift from Stage 1 to Stage 3—exhibits purely epistemic rationality. But my example does not require this. The agentialist thesis targeted by my example concerns the normative significance of rational agency, where the rationality involved in “rational agency” encompasses both epistemic and practical rationality: that is, the capacities to reason about what is true and about what to do. To exercise rational agency is to respond in a certain way to reasons—where these reasons may be epistemic or practical reasons, and where the response may be a belief or an intention. The case of Diane is designed to challenge the normative significance of the *directness* involved in rational agency. Diane's means of aligning her attitude with her reasons is a *rational* (albeit indirect) response to reasons, and issues in the appropriate attitude (a belief that fits her evidence).

McPherson's objection is a way of interpreting the agentialist claim that the “need to find some *way* to make [my conclusion] my own, my actual state of mind” is “an expression of the failure of reasoning” (Moran 2001, 131). The suggestion is that the need to engage in practical reasoning about how to align one's beliefs with one's evidence expresses a failure of epistemic (theoretical) reasoning. I readily grant that the need to engage in practical reasoning is a sign of *imperfect* rationality, since ideally rational beings always exercise rational agency, and so they always believe directly on the basis of their evidence. The issue at hand is whether ideally rational responses to reasons have a non-instrumental normative significance that is absent from exercises of imperfect rationality like Diane's.

#### 6.5 *Why Diane's attitude revision is rational*

Diane's reshaping of her attitude is *rational*. It is undertaken with the goal of conforming her attitude to her evidence, and it succeeds in achieving that goal by the use of methods that

(let's stipulate) she has good reason to trust.<sup>14</sup> Diane's reasons for endorsing MR skepticism—the reasons she acquires through deliberation—both explain and rationalize her belief. The process Diane uses to revise her attitude about MR takes some time, so her attitude is out of alignment with her reasons for an extended period; and her method is less reliable than rational agency. Still, Diane believes in MR skepticism because that is the position that her reasons dictate. Moreover, we can suppose that if she were to become persuaded by a new argument for MR realism, she would come to embrace that position, either immediately or through a deliberate process of revising her attitude.

Moran describes a case like Diane's as a "failure of reasoning" (Moran 2001, 131). But it seems to me that reasoning succeeds when it generates and rationalizes an appropriate belief, even if the path from reasoning to belief is indirect. Diane's Stage 3 belief about MR is explained and rationalized by her evidence, and it is sensitive to new evidence. Although her means of achieving and maintaining alignment between her evidence and her reasons is non-optimal, her belief is nevertheless rationally formed and rationally maintained.

## **7. Diane is responsible for her revised belief**

### *7.1 Accounts of responsible agency*

For current purposes, we should grant the agentalist idea that believing and intending are exercises of agency, for which the thinker bears responsibility. (Of course, the skeptical arguments about MR that Diane finds persuasive apply equally to non-moral responsibility, but in order to engage with the agentalist I'll assume that those arguments aren't sound.)

Diane appears to satisfy plausible conditions for responsibility, relative to the shift in her attitude about MR. To make this case, I'll briefly explain how Diane meets the conditions for responsible agency given by the three leading approaches to understanding responsibility: hierarchical accounts, reasons-responsive accounts, and Strawsonian accounts. These approaches are usually associated with compatibilism. But most libertarians adopt one or another of these approaches, supplementing their favored approach with the further requirement that indeterminism affects the decision or action for which the agent is responsible, or some step leading up to that decision or action.

Hierarchical accounts take responsibility for  $\phi$ -ing to consist in the fact that one's  $\phi$ -ing, or deciding to  $\phi$ , is guided by one's values or by what one endorses as good reasons. By this standard, Diane is responsible for her revised attitude. Her revision of her attitude is guided by her values: in particular, her commitment to believing what the evidence supports. It is because of this commitment that she goes to such lengths to bring her belief into alignment with her evidence. And she accepts MR skepticism because she endorses certain arguments as providing good reason to do so.

Reasons-responsive accounts take responsibility for  $\phi$ -ing to consist in the capacity to recognize and respond to reasons to  $\phi$  (or not to  $\phi$ ). Diane has this capacity: she can recognize

---

<sup>14</sup> We can stipulate this without assuming that mindfulness meditation is actually effective in neutralizing reactive emotions, since a false claim can be backed by good evidence. Diane's evidence could derive from scientific articles in reputable journals, describing (what she reasonably takes to be) well-conducted studies of meditation.

and respond to reasons to believe in—and, we can suppose, to reject—MR skepticism.<sup>15</sup> These reasons are generated by her evaluation of philosophical arguments about MR.

Strawsonian accounts take responsibility for  $\phi$ -ing to consist in its being appropriate for others to hold one responsible for  $\phi$ -ing, that is, to react to one's  $\phi$ -ing in certain ways (including admiration and resentment). "To be responsible for something ... is to be open to certain sorts of assessment on account of that thing, and ... to be the appropriate target of certain sorts of reactions on account of it" (Hieronymi 2014, 9). As Diane brings her attitude into alignment with her reasons, it seems appropriate to admire her for believing what her evidence supports. Depending on whether one takes Diane to be accountable for her tendency towards indignation and related affective dispositions, the fact that achieving this alignment requires a struggle may increase or diminish the degree of admiration that is appropriate.

Diane meets the conditions for responsibility associated with the three leading approaches to this issue. So if any of these leading accounts of responsible agency are correct, Diane is responsible for revising her attitude and for the result of that revision, her acceptance of MR skepticism.

### *7.2 Is this the right kind of responsible agency?*

The agentialist may object that the accounts of responsible agency just canvassed concern voluntaristic agency—or "mere control"—whereas rational agency is a matter of "authority", a distinctive, non-voluntaristic type of agency. However, those accounts are intended to be quite general. In fact, some accounts of responsibility for belief adopt the reasons-responsive approach, the Strawsonian approach, or some combination thereof (e.g. Hieronymi 2014, McHugh 2014, Wedgwood 2013). Most importantly, these accounts are intended to reflect our basic conception of responsible agency.

### *7.3 Why Diane's attitude revision is agential*

According to the leading approaches to responsible agency, Diane is responsible both for revising her attitude about MR and for her Stage 3 belief in MR skepticism.

The Normative Significance of Rational Agency thesis implies that Diane's way of revising her attitude is not only instrumentally sub-optimal but normatively problematic in some more basic sense. However, it is difficult to see why her process should be thought to be problematic. Diane's belief in MR skepticism is guided by her values—in particular, her commitment to rational norms. Her belief is generated and sustained by a reasons-responsive mechanism, and she is the appropriate target of reactive attitudes in virtue of it. Diane's successful struggle to align her belief to her evidence thus exemplifies responsible agency, according to leading accounts of responsible agency. So her efforts to revise her attitude, and the revised attitude that results from those efforts, have whatever normative significance attaches to exercises of responsible agency.

---

<sup>15</sup> Reasons-responsive accounts of responsibility do not generally require, for responsible agency, that the reasons to which the agent responds are the best reasons in an objective sense. If they did require this, then agents would be responsible for only those actions that are guided by an ideal and correct evaluation of reasons. But some responsibility theorists do take responsibility to require an ability to recognize and respond to *good* reasons (Wolf 1990, Nelkin 2011).

## 8. Diane's attitude revision is first-personal

### 8.1 Control and the first person

According to agentialism, one is estranged from those attitudes that are not directly sensitive to one's reasons: such attitudes are, in some sense, not truly one's own. Insofar as an attitude is truly *mine*, I am responsible for it, and I exert rational agency relative to it. In describing the possibility of shaping an attitude by a means other than reflection on reasons, Moran says "This 'external' sort of activity and responsibility is not essentially first-personal at all" (Moran 2001, 118).

The previous section made clear that Diane bears first-person responsibility for her Stage 3 belief. That belief is shaped and sustained by *her* values, including her commitment to rational norms; it is responsive to *her* reasons; and *she* is an appropriate target of reactive attitudes on account of it.

Still, the agentialist seems concerned with the possibility that Diane's approach—engaging in a non-rational process to revise her attitude—could be used to shape someone else's attitude.<sup>16</sup> Suppose that Diane manages to brainwash her friend into believing MR skepticism. In that case, Diane would bear responsibility for her friend's belief. But that belief would not be rational in the way that Diane's belief in MR skepticism is: it is not explained or rationalized by *the believer's* sense of the evidence. Diane's revision of her attitude about MR is first-personal because it is rationalized by *her* evidence for MR skepticism, and explained by *her* commitment to believing in accord with the evidence. Bringing someone else to believe that *p* cannot be rationalized (simply) by my evidence for *p*, and the effort to transform someone else's belief cannot be explained by my commitment to rational norms. These normative elements are first-personal: that Diane undertakes a deliberate process to arrive at a rationally-held belief, and thereby to satisfy the corresponding rational norm, does not make the attitude any less her own. The particular reasons and values that drive Diane to revise her belief, and that make that revision *rational*, also make it first-personal.

### 8.2 Considering a belief as a belief

Here is another reason one might think that only attitudes that express rational agency are truly our own. Assessing an attitude as *recalcitrant* requires seeing it as an object rather than simply seeing the world through its lens, so to speak. And it is the latter—e.g., Diane's seeing the world as containing persons possessing MR, as opposed to considering "we possess MR" as a psychological state—that makes the state truly one's own. (Hence the metaphor of transparency in the claim that we can grasp what we believe about *p* by simply considering whether *p*. In this metaphor, the belief is the transparent lens through which we view the world.)

Considering a belief as a belief—that is, as a psychological state—certainly differs from simply seeing the world through its lens. But this doesn't mean that considering it as a psychological state makes it less one's own. Although Diane regards elements of her Stage 1 attitude about MR as recalcitrant, she can nonetheless see the world through the lens of those elements: when she feels moral indignation towards the oil company executive, he seems to her

---

<sup>16</sup> Obviously, one could use a *rational* process to shape someone else's belief, by presenting them with an argument. When this attempt succeeds, the shift in belief is best explained by the believer's own deliberation.

*blameworthy* and *deserving of retribution*. The ability to recognize a belief—whether it is recalcitrant or not—requires thinking about beliefs as such. But the ability to recognize a belief as a belief does not hinder us from seeing the world through its lens.

In fact, considering a psychological state as such can arguably *enhance* one’s agency relative to that state. This idea is related to a point that McGeer (2008) develops in critiquing Moran’s agentialist view. Moran holds that we have agential responsibility for our attitudes insofar as those attitudes are directly reasons-responsive. McGeer argues that limiting first-person agency to this kind of reasons-responsiveness is unduly restrictive. Given that we are not perfectly rational, she says, fulfilling our responsibility for our own attitudes sometimes requires adopting a detached, third-person perspective on our cognitive lives. From this perspective, we can better identify biases, impulses, and other non-rational factors affecting our sense of our reasons (McGeer *ibid.*, 102).

I think McGeer is entirely right about this. Moreover, the case of Diane supports and extends her point. It shows that adopting a detached, third-person perspective is required not only for recognizing the influence of non-rational factors on our sense of our reasons, but also for blunting that influence.<sup>17</sup> Diane is able to effect the shift in her attitude only by adopting a detached perspective—that is, by refusing to identify with her retributive impulses, condemnatory thoughts, and other recalcitrant elements of her attitude about MR, regarding them as unfortunate features of her psychology.

Of course, Diane would not need to resort to methods that Moran and Boyle regard as mere “control” if her attitude about MR conformed directly to her assessment of the arguments. But that is just to say that she wouldn’t need to resort to an indirect process if a direct one was operative.

### 8.3 *Why Diane’s attitude revision is first-personal*

The agentialist holds that only those attitudes that express rational agency truly belong to the thinker. I have suggested that attitudes *belong* to the thinker so long as they are explained and rationalized by the thinker’s reasons, including a commitment to rational norms—regardless of whether the thinker believes directly on the basis of those reasons or exploits a non-rational process to bring her attitude into line with her reasons. (My claim here is that this suffices for belonging to the thinker, not that it is necessary.) And given that we are imperfect thinkers, and therefore have some recalcitrant attitudes, adopting a detached perspective can enhance our agency, enabling us to identify and correct recalcitrant attitudes.<sup>18</sup>

---

<sup>17</sup> Levy (2016) makes a related point about addiction. Alcoholics Anonymous encourages participants to describe themselves third-personally, with the famous phrase “My name is X and I’m an alcoholic”. Levy argues that the practice of regarding oneself in this way can be helpful in combatting addiction, since for an addict the reasons to drink or take a drug can appear, at the moment of choice, stronger than the reasons not to do so. In such cases, deciding what to do by directly considering what you have most reason to do can lead to backsliding on a previous commitment to quit.

<sup>18</sup> Borgoni (2015) addresses a related concern, arguing that we have privileged first-person access to our recalcitrant attitudes.

## 9. Is rational agency required to avoid a regress?

It might be argued that, on pain of regress, one can undertake to bring one's attitudes into alignment with one's reasons only by *some* exercise of rational agency. For in order to try to adopt attitude B on the basis of reasons R, one must recognize that R favor B. But that recognition—the judgment that reasons R favor B, e.g. that the evidence points to *p*—is, in the best case, a response to reasons. If it's a direct response, then rational agency is involved. If it's not a direct response, this means that the thinker must take up the task of bringing her attitude—in this case, towards *R favor B*—into alignment with her reasons. And that in turn requires recognizing that her reasons favor that judgment, a recognition that (in the best case) is a response to reasons. On pain of regress, at some point in the deliberative process an attitude—e.g., a judgment as to where the evidence points, or what one's reasons favor—must be shaped by one's reasons *directly*.

I'm happy to grant this claim for current purposes. It means that rational agency plays a crucial role in Diane's case, in that it makes her efforts to conform her attitude to her reasons possible. More generally, the regress argument shows that rational agency will contribute to any case of *believing on the basis of reasons*. But this is just to say that some *direct* connection between reasons and attitudes is a necessary component of any process in which reasons shape attitudes indirectly.

It's worth noting that the regress argument does not imply that there must be a direct connection between reasons favoring a belief that *p* and the judgment that *p*. Consider the following case. I recognize that my reasons favor *my friend is guilty of a serious offense*, but my emotional investment in her innocence makes it difficult for me to embrace this conclusion even in the way required for a momentary judgment. I acknowledge that I *should* judge her to be guilty, yet I'm unable to see her as guilty. I might then make recourse to non-rational means to help me embrace this conclusion: e.g., by engaging in some non-rational exercise aimed to blunt the influence of my emotions on my judgment. In this case, my evaluation of my reasons directly shapes my judgment that *I should consider her guilty*, but the connection between my reasons and my eventual judgment that *she is guilty* is indirect in that it rests on a non-rational process.

## 10. Must we assume that deliberation will directly shape our attitudes?

### 10.1 Moran's requirement

I have just conceded that rational agency may always play a crucial role in an agent's conforming her attitudes to her reasons, even in cases like Diane's where the attitude in question is only indirectly shaped by reasons. This concession may appear to surrender one motivation for challenging the Normative Significance of Rational Agency thesis. That motivation is to avoid the idea that we cannot coherently embark on rational deliberation unless we make what Moran describes as “a presupposition of rational thought”: the assumption that “what I believe about something [is] the expression of my sense of the reasons relating to the content of that belief” (Moran 2012, 232). Moran elsewhere calls this “a Transcendental assumption of Rational Thought”, and aptly describes it as being “as venerable and familiar as it is obscure” (Moran 2003, 406). This suggests that (at least our presumption of) rational agency has a special normative significance.

[If a thinker takes] his belief or his intentional action to be up to something other than his sense of the best reasons, ... then there's no point in his deliberating about what to do. Indeed, there is no point in calling it 'deliberation' any more, if he takes it to be an open question whether this activity will determine what he actually does or believes. (Moran 2001, 127)

There are various ways to unpack the idea that beliefs' conforming to reasons is a "presupposition of rational thought". On a weak reading, it means that I will not be motivated to deliberate unless I assume that deliberation will issue in a *judgment* as to whether  $p$  (or a judgment to the effect that, given my current evidence, I should withhold belief as to whether  $p$ ). But it's not entirely clear that even on this weak interpretation Moran's claim is correct. It seems possible to embark on deliberation about whether my friend is guilty while acknowledging that I may be incapable of judging her guilty even if I conclude that there is strong evidence of her guilt.

In any case, Moran appears to be making a stronger claim here, namely that I will not be motivated to deliberate unless I assume that my deliberation will issue in a *belief* as to whether  $p$ : that is, a relatively stable set of dispositions fitting the stereotype for a belief that  $p$  or that not- $p$  (or dispositions corresponding to withholding from a belief as to  $p$ ).

#### 10.2 *Why we needn't assume that deliberation directly shapes our beliefs or intentions*

It seems to me that we can coherently engage in deliberation without assuming that this process will directly shape our beliefs or intentions. Consider the following scenario. Carly hasn't read Hume's skeptical argument about induction. But she has heard that the argument concludes that our inductive beliefs are not epistemically justified. She also knows that Hume predicted that even those persuaded by his argument would continue to rely on induction—that is, to act and reason as if past events were a reliable guide to what will occur in the future. Before working through Hume's argument, Carly might reflect as follows.

Although my deliberations about induction (guided by my evaluation of Hume's argument) will presumably shape my judgment as to whether our inductive beliefs are justified, they may not shape my inductive beliefs themselves. So I may continue to believe that the sun will rise tomorrow, even if I judge that I have no reason (of the "right" kind) for that belief—that is, even if my "sense of the best reasons" (Moran *ibid.*) dictate withholding belief as to whether the sun will rise tomorrow.

These reflections do not prevent Carly from deliberating about induction by evaluating Hume's skeptical argument. This scenario casts doubt on Moran's claim that to engage in deliberation one must presuppose that "what I believe about something [is] the expression of my sense of the reasons relating to the content of that belief" (Moran 2012, 232).

Similarly, suppose that Diane initially approaches the question of MR with an open mind, and with no expectation as to whether she will be persuaded by the arguments for MR skepticism or will instead favor MR realism. Prior to working through these arguments, Diane might recognize that the idea that we are morally responsible permeates her habits of acting, reasoning, and responding to others, habits that are deeply rooted and difficult to break. So she could acknowledge that her deliberation about MR may not directly result in the corresponding revision of her dispositions. It is nonetheless perfectly reasonable—and, I think, even

admirable—for her to embark on the project of determining whether those dispositions *should* be altered, even as she recognizes that this alteration might require extended effort.

### 10.3 *Why we deliberate*

As a conceptual matter, it seems possible to embark on deliberation even if one doubts that it will have *any* effect on one's dispositional beliefs. Some might engage in deliberation purely because they find it enjoyable. Moran would likely deny that this process deserves the name “deliberation”. But that is a verbal issue: if evaluating arguments purely for fun doesn't qualify as “deliberation”, that is presumably for the trivial reason that the notion of deliberation is conceptually connected to the aim of shaping beliefs and intentions.

In any case, such examples are outliers. Most of us deliberate with the aim of shaping our beliefs and intentions. But we sometimes recognize that the route from reaching a conclusion as to what we should believe or do, and forming the corresponding belief or intention, may be indirect. Prior to deliberating as to whether he should quit drinking, a heavy drinker will surely recognize that giving up alcohol would be challenging. He may even recognize that the prospect of quitting would be so daunting that, if his reasoning leads him to judge that it would be best to quit, he will need to struggle to form the intention dictated by his reasons. This case is not exceptional: we often deliberate about whether to  $\phi$  while recognizing that powerful reason to think that  $\phi$ -ing is best will not directly lead to our  $\phi$ -ing. One can engage in practical reasoning with the expectation (and hope) that such reasoning will affect one's decisions and actions, while recognizing that conforming one's intentions to one's best reasons may involve an extended, tortuous process.

As a general matter, we probably would not be motivated to deliberate unless we assumed that the outcome of our deliberation would shape our actions and beliefs. The case of the heavy drinker demonstrates that the assumption motivating a bit of practical reasoning may be only that engaging in practical reasoning will raise the likelihood of doing what's best. And as the case of Carly illustrates, a truly open-minded thinker may embark on deliberation while recognizing that her beliefs may well resist the force of her evidence. So it seems perfectly coherent to deliberate without assuming that one's intentions and beliefs will conform to the outcome of one's deliberations immediately and directly.

## CONCLUSION

The agentialist view envisions a stark contrast between two types of cases. The first are exercises of rational agency, in which a belief directly conforms to reasons. According to the agentialist, these involve a distinctive kind of agency with a special normative significance. The second are cases like Diane's, in which the thinker struggles to bring her attitude into alignment with her reasons and succeeds only by resorting to non-rational methods. The agentialist sees those cases as deficient, lacking the normative significance that attaches to rational agency. I've argued that beliefs indirectly shaped by reasons are rationally held; that such beliefs are exercises of agency for which the thinker is responsible; that rational beliefs truly belong to the person whose reasons rationalize them; and that acknowledging a belief as a psychological state is compatible with acknowledging it as one's own, and may even enable one to exercise greater agency relation to it.

It is easy to appreciate the value of rational agency. Believing and intending directly on the basis of what one regards as one's best reasons is the maximally efficient and reliable way to avoid the kind of fragmentation that seems rationally problematic: having beliefs (or intentions) that diverge from the beliefs (or intentions) that one takes one's reasons to favor. So rational agency is preferable, on instrumental grounds, to the indirect process Diane uses to align her attitude with her reasons. If the regress argument succeeds, then rational agency is also instrumentally significant in that it makes a crucial contribution to believing and intending on the basis of reasons even when the route from reasons to attitude is indirect.

But it is hard to see why rational agency should be thought to have a normative significance not shared by the efforts of conscientious thinkers like Diane, when they diligently struggle to fit their attitudes to their reasons. We should recognize and respect the variety of strategies by which imperfect thinkers like us manage to bring our attitudes into closer alignment with our reasons, and thereby become less fragmented.<sup>19</sup>

### Works cited

- Bilgrami, A. (2006), *Self-Knowledge and Resentment* (Harvard University Press).
- Borgoni, C. (2015), 'On Knowing One's Own Resistant Beliefs', in *Philosophical Explorations* 18: 212-225.
- Borgoni, C. THIS VOLUME
- Boyle, M. (2009), 'Active Belief', in *Canadian Journal of Philosophy* 39: 119-147
- Burge, T. (1996), 'Our Entitlement to Self-Knowledge', in *Proceedings of the Aristotelian Society* 96: 91-116.
- Burge, T. (2011), 'Self and Self-Understanding', in *The Journal of Philosophy* 108: 287-383.
- Coliva, A. (2012) 'One Variety of Self-Knowledge: Constitutivism as Constructivism', in A. Coliva, ed., *The Self and Self-Knowledge* (Oxford University Press).
- Davidson, Donald (1973), 'Freedom to Act', in T. Honderich (ed.), *Essays on Freedom of Action*. (Routledge).
- Gendler, Tamar Szabó (2008), 'Alief and Belief', in *Journal of Philosophy* 105: 634-663.
- Gertken, Jan & Kiesewetter, Benjamin (2017), 'The right and the wrong kind of reasons', in *Philosophy Compass* 12 (5):e12412.
- Gertler, B. (2018), 'Self-Knowledge and Rational Agency: A Defense of Empiricism', in *Philosophy and Phenomenological Research* 96: 91-109.
- Goleman, D. and Davidson, R. (2017), *Altered Traits: Science Reveals how Meditation Changes Your Mind, Brain, and Body* (Penguin Random House).

---

<sup>19</sup> I am deeply grateful to Eric Schwitzgebel and an anonymous reviewer for their feedback on an earlier draft of this paper. Their thoughtful questions and suggestions improved the paper significantly. I also owe thanks to an audience at The Ohio State University, where I presented a version of this paper, and especially to Justin D'Arms, Nathan Dowell, Tristram McPherson, and Declan Smithies.

- Hieronymi, P. (2014), 'Reflection and Responsibility', in *Philosophy and Public Affairs* 42: 3-41.
- Horowitz, S. (2014), 'Epistemic Akrasia', in *Noûs* 48: 718-744.
- Korsgaard, C.M. (1996), *The Sources of Normativity* (Cambridge University Press).
- Levy, Neil. (2015), 'Neither Fish nor Fowl: Implicit Attitudes as Patchy Endorsements', in *Noûs* 49: 800-823.
- Levy, Neil. (2016), '“My Name is Joe and I'm an Alcoholic”: Addiction, Self-knowledge and the Dangers of Rationalism', in *Mind and Language* 31: 265-276.
- Mandelbaum, E. (2016), 'Attitude, Inference, Association: On the Propositional Structure of Implicit Bias', in *Noûs* 50: 629-658.
- McGeer, V. (2008), 'The Moral Development of First-Person Authority', in *European Journal of Philosophy* 16: 81-108.
- McHugh, C. (2014), 'Exercising Doxastic Freedom', in *Philosophy and Phenomenological Research* 88: 1-37.
- Moran, R. (2001), *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton University Press).
- Moran, R. (2003), 'Responses to O'Brien and Shoemaker', in *European Journal of Philosophy* 11: 402-419.
- Moran, R. (2012), 'Self-Knowledge, “Transparency”, and the Forms of Activity', in D. Smithies and D. Stoljar, eds., *Introspection and Consciousness* (Oxford University Press), 211-236.
- Nelkin, D. (2011), *Making Sense of Freedom and Responsibility*, (Oxford University Press).
- Reisner, A. (2018), 'Pragmatic Reasons for Belief', in D. Star (ed.), *The Oxford Handbook of Reasons and Normativity* (Oxford University Press), 705-728.
- Roush, S. (2017) 'Epistemic Self-Doubt', in *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2017/entries/epistemic-self-doubt/>](https://plato.stanford.edu/archives/win2017/entries/epistemic-self-doubt/).
- Schwitzgebel, E. (2002), 'A Phenomenal, Dispositional Account of Belief', in *Noûs* 36: 249-275.
- Wedgwood, R. (2013), 'Rational “Ought” Implies “Can”', in *Philosophical Issues* 23: 70-92.
- Wolf, S. (1990), *Freedom Within Reason* (Oxford University Press).
- Yalcin, S. THIS VOLUME