What is (in)coherence?
Alex Worsnip

Note: this is a longer version of a paper that is forthcoming in a shorter form in <u>Oxford Studies in Metaethics</u>, vol. 13. Please cite the published version where possible.

Recent work on rationality has been increasingly attentive to "coherence requirements".[1] Coherence requirements are the requirements of rationality that have to do with the structural coherence between one's mental attitudes; about how they (in a broad sense) fit together; about which *combinations* of attitudes (and absences of attitudes) it is rational or irrational to hold *jointly*. Familiar examples of possible coherence requirements include the enkratic requirement (which requires one to intend to do what one believe one ought to do), the instrumental requirement (which requires one to intend the means to one's ends) and the noncontradiction requirement (which requires one not to believe contradictory propositions). Those are some of the most familiar examples, but in principle there could be coherence requirements on combinations of any kind of attitudinal mental states: for instance, perhaps certain combinations of beliefs and hopes are incoherent, or certain combinations of beliefs and fears. Many formal epistemologists propose that there are requirements not to have certain incoherent combinations of graded credal states; some propose that there are requirements not to combine certain graded credal states with attitudes of full belief. All of these are at least *candidate* coherence requirements in the sense I am interested in.

There is a debate to be had about whether such requirements exhaust the requirements of rationality, though some may wonder if that debate is at least partly a terminological one about how to use the term 'rationality'. Here, to be neutral on this debate, I will just talk directly about coherence. As should be obvious from the list of candidate coherence requirements above, coherence in my sense is not restricted to logical consistency or to probabilistic coherence. Indeed, I take it to be a substantive question whether putative requirements of logical consistency or of probabilistic coherence are actually genuine (as opposed to merely putative) coherence requirements. More generally, 'coherence' is not here being used in a *stipulative* fashion whereby certain combinations of states count as incoherent by stipulation; it is always open to philosophical debate whether some particular combination of states is really incoherent or not. (The account of coherence that I will eventually offer in this paper aims to illustrate how such a debate can take place in a principled, non-stipulative manner.)

The foregoing loose characterization gives us some idea of the *form* of coherence requirements. Such requirements, we've just said, pertain to the rational permissibility or impermissibility of *combinations* of mental states. This may get us at least some way to being able to tell, given some putative

[1] See, e.g., Broome (2013); Kolodny (2005, 2007, 2008a, 2008b); Scanlon (2007); Worsnip (forthcoming); Fogal (ms.).

requirement, whether it has the form of a putative coherence requirement (as opposed to some other kind of putative requirement). But what it does not give us is an account of which coherence requirements are *genuine* as opposed to *merely* putative: of the conditions under which a combination of mental states really does count as incoherent. For example, the putative "requirement" that, for any proposition p, one is rationally required not to (hope that p is true and have a credence of 0.6 in p), has the right *form* to be a coherence requirement by the foregoing account, but it is presumably not plausible that this is a genuine requirement.[2] As such, a purely formal account still leaves us with the *substantive* question: what *is* (in)coherence, really?

There are both metaphysical and epistemological questions in the neighborhood here. Metaphysically: what is it for two or more mental states to be jointly incoherent, such that they are banned by a coherence requirement? In virtue of what are some putative requirements genuine and others not? Epistemologically: how are we to *know* which of the requirements are genuine and which aren't? Typically, theorists proceed by just listing candidate requirements and considering potential counterexamples.[3] But in the absence of a general account of what coherence is, and thus of what we are looking for or having intuitions *about*, this procedure seems unguided.

These questions are made more pointed in two ways. First, the list of candidate coherence requirements is diverse, governing both doxastic and practical mental states and including everything from bans on weakness of will, to norms supplied by deductive logic, to axioms of decision theory. They include requirements on combinations of doxastic states, requirements on combinations of practical states, and requirements on combinations of doxastic and practical states. One might reasonably wonder what, if anything, all of these requirements have in common. We should hope that an account of coherence requirements will explain why these requirements all belong to a single category, by giving some general account of what it is for mental states to be jointly incoherent.

Secondly, one of the most lively debates in the literature on coherence requirements addresses the question of whether such requirements are *normative*, in the sense where a requirement is normative if one necessarily has reason to, or really ought to, comply with it.[4] But this debate often takes place against a shared assumption that there *are* coherence requirements. The assumption here is not merely that there "are" requirements in the sense that there are putative or "candidate" requirements, or things that are held to be requirements (in the same sense that there "are" principles that individuals hold, irrespective of whether those principles are true or false). Rather, the assumption is that there are coherence requirements in the sense that some putative coherence requirements are genuine and others aren't, in the sense that one can make mistakes about whether some putative coherence

---

[2] As well as this obviously non-genuine requirement, there will also be controversial cases. The already-mentioned example of requirements of logical consistency and probabilistic coherence provide examples here. Such putative requirements certainly have a form such that, if they are genuine requirements at all, they are genuine coherence requirements. But nothing we've said so far helps us to tell whether they are in fact genuine requirements or not.

[3] So, e.g., Broome (2013: 150): "How can we identify requirements of rationality? I wish I could describe a general method of doing so, but I am sorry to say I cannot. I shall defend a number of requirements one by one, on particular grounds that seem appropriate […] I find myself forced to appeal largely to our intuitions."

[4] See e.g. Kolodny (2005, 2007); Raz (2005); Broome (2005a, 2005b, 2013: ch. 11); Southwood (2008); Ross (2012); Lord (forthcoming); Kiesewetter (forthcoming).

requirement is a genuine requirement or not.[5] Thus, participants in the literature seem willing to countenance the idea that there *are* coherence requirements, but yet that these requirements are not normative. But this only intensifies our question about what coherence is, and what coherence requirements are. If an attribution of incoherence does not necessarily amount to a charge of a normative failing, what exactly does it come to? And if we cannot necessarily determine whether a coherence requirement is genuine by thinking about whether one *really ought*, normatively speaking, to satisfy it, then we seem to have even less to go on in figuring out which requirements are genuine and which aren't.

Indeed, one might worry that perhaps, the notion of a non-normative *requirement* doesn't even really make sense. As it stands, this worry is overstated. There are clear examples of genuine but non-normative requirements: for example, the requirements of British Victorian etiquette, of Mafia morality, of the grammar of the French language circa 1931, and so on. The reasons to comply with such requirements are at best both derivative and contingent: in and of themselves, they lack normative force. Yet there are *genuine* requirements of this sort, in the same that some things *are* actually requirements of Victorian etiquette, and some things aren't. In enumerating the requirements of Victorian etiquette, one can get them right or get them wrong. What is a myth is the *normativity* of these requirements, not their existence. However, all of these requirements are plausibly *conventional* in nature. They *are* requirements in virtue of conventional facts about the actual practices of (in these cases) Victorian Brits, Mafiosi, and French speakers in 1931. Yet it is less attractive, and certainly less obvious, that the requirements of coherence can be cleanly understood as being fixed by some kind of convention. So, to sharpen our worry: what we really are struggling to make sense of is not non-normative yet genuine requirements *per se*, but genuine requirements that are neither (genuinely) normative *nor* purely conventional. Again, a general account of what coherence is would be of great help here.

This paper is an attempt to give such an account. My account will be guided by three aims. First, it will aim to unify different, diverse coherence requirements, and to show what they have in common. Second, it will aim to provide us with principled criteria for determining whether coherence requirements are genuine or not (in a sense of 'genuine' that encodes more than a putative requirement's merely having the right *form* to be a coherence requirement, but less than an assumption that the requirement is robustly normative).[6] That is not to say that it will on its own definitively and

---

[5] For example, Kolodny (2005), perhaps the leading opponent of the view that coherence requirements are normative, nevertheless holds that the various variations on the enkratic requirement comprise the "core" rational (coherence) requirements, suggests tentatively that all other rational requirements might be reduced to the enkratic requirement(s) (Kolodny 2005: 557), and takes the position that to get these requirements right, we must formulate them as narrow-scope rather than wide-scope (Kolodny 2005: 518-539). These views only make sense if we understand Kolodny as holding that these enkratic requirements are genuine requirements of rationality – that rationality requires one to comply with them (Kolodny 2005: 513, 551-560) – but he nevertheless denies that one has *reason* to comply with them (Kolodny 2005: 514-551). Similarly, Broome (2013) takes seriously the idea that coherence requirements might not be normative (chapter 11), but never seriously questions the idea that they are nevertheless genuine requirements, about which there are correct answers as to which are genuine and which aren't, in some good sense.

[6] Kathryn Lindeman and Sarah Stroud pushed me on the question of whether I need to make a further distinction here, between a set of states being genuinely *incoherent*, and the requirement than bans these states being a genuine *requirement*. Both of these notions, according to Lindeman and Stroud, are more than merely formal, but less than robustly normative. I find myself torn as to whether I can see the distinction that Lindeman and Stroud see here. I am assuming at least *extensional* equivalence between the two notions they identify. If a set of states is *genuinely* incoherent, then I

3

transparently settle every controversial case, but it will at least show how the debate is to proceed. Third, it will aim to assign coherence requirements an important philosophical role (again, irrespective of their normative status). If we can find an account of coherence and of coherence requirements that satisfies these three aims, then coherence requirements will, I believe, earn their ontological keep. I do not aim to be giving a conceptual analysis of whatever is being picked out by all uses of the word 'coherent' in English. Nor will I treat our pre-theoretical, intuitive list of genuine coherence requirements (if there be such a list) as unrevisable (though no doubt the account shouldn't make extensional predictions that deviate too wildly from this list). Rather, I will be looking for an interesting, well-regimented, philosophically important notion in the neighborhood of what we are talking about when we talk about coherence.

In asking questions like "why care about coherence?", it is easy to slip from the point of view of the *agent* asking whether coherence requirements matter *normatively*, to the point of view of the *theorist* asking whether coherence requirements matter *philosophically*. But not every philosophically important phenomenon is important normatively. In particular, the account I will give assigns coherence requirements an important and highly distinctive role in our (descriptive) *philosophical psychology*. So I hope that my account will make coherence requirements interesting even for those who are skeptical about their normativity. But equally, I think that my account could be accepted by someone who thinks that coherence requirements *are* normative. I myself will take no firm stand on the debate about the normativity of coherence requirements, though I will make a few programmatic remarks about this issue at the end of the paper, in light of the account that I have offered.

The account of coherence (and of coherence requirements) that I am offered here might be thought of as a kind of naturalistic, reductive realism about this property, though the proper application of these labels is a notoriously fraught matter. In any case, I do *not* take the account developed here for coherence requirements to generalize to other, more "substantive" reasons and requirements. If anything, the particular version of the view I develop – especially its claims about the role of coherence requirements in our philosophical psychology – reinforces the metanormative *disunity* of these coherence requirements on one hand and substantive (moral, prudential, epistemic, etc) norms on the other.[7]

I.   The view

No point being shy; here is the view. A set of attitudinal mental states is jointly incoherent if and only if it is (partially) constitutive of the mental states in question that, for any agent that holds these attitudes, the agent is disposed, when conditions of full transparency are met, to give up at least one of the attitudes. That is, human agents are disposed such that they are (at least normally) not able to (or at least find it difficult to) psychologically sustain such combinations of attitudes under conditions

---

assume that there is a *genuine* coherence requirement that bans holding them jointly (bearing in mind that neither use of 'genuine' here is supposed to entail robust normativity). I can see that one might think that more needs to be said to fully vindicate this assumption, but I must simply rely on it here.

[7] There is little to no existing work on the *metanormative* (as opposed to *normative*) status of coherence requirements. Hussain (ms.: esp. 10-11) draws attention to the importance of this topic, but does not give a detailed account of it.

of full transparency. A putative coherence requirement is genuine if and only if every combination of states that it forbids is jointly incoherent.[8]

"Attitudinal mental states" as I use the term includes both mental attitudes and *absences* of mental attitudes, such as the absence of a particular belief or intention. Some coherence requirements (for instance, the instrumental requirement) effectively ban one from having some attitude while *lacking* some other attitude, so this broadness is required for full generality.

By "conditions of full transparency", I mean conditions under which the agent knows, and explicitly and consciously believes, that she has the states in question, without self-deception, mental fragmentation, or any failure of self-knowledge (pertaining to those attitudes). Notice that it is *not* required for these conditions to be met that the agent acknowledge *that her mental states violate a requirement* as such. It is merely required that she acknowledge that she has the states that (perhaps unbeknownst to her) violate the requirement.

The present account makes coherence a matter of whether (or how easily) agents can psychologically sustain the states in question under conditions of full transparency. However, notice that for a combination of states to be incoherent, it has to be true that *any* human agent would be unable (or at least find it difficult) to sustain the states under full transparency. So it won't suffice for incoherence that some individual agent has a psychological quirk such that they are disposed not to hold two attitudes jointly: the disposition has to be present in all agents. Moreover, this fact has to be constitutive of the mental states in question. There may be some states that all agents will find it hard to simultaneously sustain, but where this is not constitutive of the states in question;[9] the account does not count such combinations of states as incoherent.

To clear one potential objection out of the way: the present account does not make violations of coherence requirements impossible.[10] For the avoidance of all doubt: the view I am defending does *not* say that combinations of mental states are incoherent only when they are held under conditions of full transparency. Rather, it says that some combination of attitudes is incoherent in cases where the agent is disposed such that, *were* conditions of full transparency to be met, she would at least find it difficult to sustain the attitudes together.[11] These attitudes are still incoherent when these conditions

---

[8] It's worth noting that in principle, my analysis allows that the set of attitudinal mental states that is incoherent might have only one member. I take this to be a virtue of the account, since some single attitude might be on its own incoherent. Most obviously, a single belief in a contradictory conjunction, (p and not-p), seems incoherent (cf. Broome 2013: 153). A purely formal account of coherence (cf. the introduction above) actually has difficulty explaining how this requirement is of a piece with other coherence requirements, since it does not pertain to *multiple* attitudes not fitting together. My substantive account of coherence, by contrast, can explain why this requirement belongs with the others. Thanks to Selim Berker, among others, for discussion here.

[9] Harvey Lederman suggested the example of working on a complex math problem while being in excruciating pain. I'm not sure either of these things are mental *attitudes* in my sense, but it gives a sense of the general sort of example.

[10] Confusions about what or whom the incoherence is being predicated *of* can confuse matters here. When I say that an agent is incoherent, I mean to predicate a property – incoherence – to that agent, or that agent's mental states. But it's easy to confuse this with the claim that the very idea of such an agent existing is incoherent. When we predicate incoherence of an idea, we mean that the idea is internally incoherent in a way that means that is does not represent a genuine possibility. So if what was incoherent was not the agent's mental states but the very idea that an agent could have such mental states, we really would be saying that such an agent cannot possibly exist.

[11] In this respect the view is like reductive subjectivist analyses of value that identify something's being valuable with it's being the case that an agent *would* value it under certain ideal conditions (cf., e.g., Lewis 1989). Such analyses obviously don't say that the thing in question ceases to be valuable when the ideal conditions aren't met: the whole point of the analysis is to identify something's being *valuable* now (in perhaps non-ideal conditions) with its being *valued* in ideal

of full transparency are not *in fact* met (and when, consequently, they may not *in fact* be difficult to sustain). Moreover, conditions of full transparency are very often not met.[12] Consequently, my view allows that coherence requirements can often be violated. It does say that such violations will tend to involve some kind of failure of transparency. But I do not think that this is a bad consequence: in the next section, I will argue that, for paradigmatic coherence requirements, our making sense of violations of them relies on a tacit assumption that conditions of full transparency fail to obtain.[13]

Though this naïve objection fails as it stands, some philosophers may worry that the status of a coherence requirements could not be *requirements* in virtue of descriptive facts about which states agents can or cannot psychologically sustain, even under conditions of full transparency. Equally, some may worry that *no* combination of mental states is such that *all* agents would be disposed not to sustain it under conditions of full transparency, and that the present approach smacks of *a priori* armchair psychology. I will work to dispel these objections as I develop my account. But before I develop the theory, let's consider some examples that help to pave the way for it.

II.   Helpful illustrative cases

We begin with some helpful cases that (I hope) make my view more plausible. We'll come to a harder case later.

(a) Instrumental irrationality

Consider first the instrumental requirement. This requirement, roughly, says that the following combinations of attitudinal mental states is incoherent: intending an end, believing some means is necessary for that end, but not intending the means.

Suppose you know that your friend's partner is cheating on her, and that she will discover this soon. You believe that it would be better if she heard it from you, both for her and for you (since she will also find out that you knew). So you intend to be the one who tells her about the infidelity. You also know (and hence, believe) that today is the last day on which you have the opportunity to tell her, and that the only way to do so is to call her. So, you would violate the instrumental requirement if you lacked the intention to call her today.

Might you lack that intention? Certainly.[14] But let's contrast two ways in which we might try to tell the story about how you do so. The easiest way to make that possibility clear and intelligible is to reach for some story on which your mental states are not fully transparent to you. So, perhaps your

---

conditions. However, I introduce this analogy just to make the structure and commitments of the view clear. As I will say in section V, I do not think that accepting my reductive account of coherence provides much reason to accept a reductive account of moral properties or of value.

[12] For arguments that such conditions are often not met, see e.g. Williamson 2000: ch. 4; Schwitzgebel 2008; Srinivasan 2015.

[13] See also Friedman (ms.: 10-11), who argues that something like this is true of the incoherent combination of believing p and taking certain "interrogative attitudes" toward p.

[14] *Pace* Finlay (2009), who argues that violation of the requirement of instrumental rationality is impossible *simpliciter*. I think that much of what motivates Finlay's argument is right, but that it overreaches. Violations of instrumental rationality require non-transparency, but are not impossible *simpliciter*.

mental states are fragmented.[15] "Somewhere" in your mind intend to tell your friend, and "somewhere" in your mind you know that to do this you must call her. But you never put these two mental states together or reflect on what they jointly commit you to. By putting at least one of the two states out of your occurrent reach – perhaps subconsciously motivated by the awkwardness of calling your friend – you never come to intend to call her. This is a familiar kind of failure. It wouldn't be correct for you, if pressed after the fact, to deny that you intended to tell your friend about the infidelity, nor for you to deny that you knew that you had to call her. You just avoided simultaneous, conscious consideration of the fact that you had both of those states. And so you managed never to form the intention to call her.

Suppose now, however, that we try to tell the story so that your mental states are fully transparent to you. So here, you explicitly say: "Absolutely: I intend to tell my friend about the infidelity. And the only way to do that is to call her today. But I have no intention whatsoever to call her today." The most natural way to hear your speech here is as a joke. Why is that? Because if you were to sincerely utter these words, you would be confused about what it is to have an intention. Why is *that*? Because someone who is in this position *doesn't really count* as genuinely intending to tell her friend about the infidelity. That is: it is part of *what it is* to intend an end that one also be disposed, *under conditions of full transparency*, to form corresponding intentions to intend the means that one believes to be necessary to that end (or to give up the intended end). Whether one will do this is a crucial litmus test of whether one really does intend the end.[16]

That isn't to say that one cannot be in various, weaker states with respect to the end. One can certainly *desire* or *wish* that one tells one's friend about the infidelity, or think that it would be good if one did so, while not forming the intention to call one's friend. On any account, we need some way of distinguishing these weaker states from intention proper. On the present account, it is part of what distinguishes these weaker states from the stronger state of intention that they can persist even in the face of conscious, reflective recognition that one is not following through on the (believed) means.

On my view, the incoherence of violations of the instrumental requirement consists in the fact that we are disposed not to sustain them jointly under conditions of full transparency. That is what makes the instrumental requirement a genuine requirement, whereas the following requirements are bogus:

> Rationality requires that if one desires that one Φ's, and one believes that to Φ one must Ψ, then one intends to Ψ.

Or even:

> Rationality requires that if one desires that one Φ's, and one believes that to Φ one must Ψ, then one desires to Ψ.

---

[15] On fragmentation see, e.g., Stalnaker (1984), Davidson (2004: ch. 11), Elga (2005), Egan (2008), and Greco (2014).
[16] Wallace (2001: esp. 26) suggests something similar.

Violations of these putative requirements are easy even under conditions of full transparency. Thus, such combinations of states are not genuinely incoherent. So the above requirements are bogus.

(b) Transitivity

It would be good to show that the present account can *unify* apparently disparate coherence requirements, showing how violations of them can each be incoherent in a single, core sense. So let's next turn to a rather different sort of coherence requirement, often found in decision theory and economics: that of transitivity of preference. This requirement bans one from simultaneously preferring A to B, preferring B to C, and preferring C to A.

Here is a case of violation of transitivity which is easy to imagine. Consider the three following things that a philosopher might do with his Saturday: working on his new article, volunteering at the homeless shelter, or re-watching series 4 of *Friday Night Lights*.

- Attending to the options of working on his article and volunteering at the homeless shelter, working on the article seems like an important project that he can justifiably pick over volunteering, and which allows him to stay in his pajamas and not have to interact with anyone. So he prefers working on his article to volunteering.
- Attending to the options of volunteering at the homeless shelter and re-watching series 4 of *Friday Night Lights*, choosing to do something so trivial as watch TV rather than volunteering seems callous. So he prefer volunteering to watching TV.
- But, attending to the options of watching *Friday Night Lights* and working on his new article, the writing of the article seems difficult and energy-consuming after his long week, and it's not like it's *morally* required. So he prefers watching TV to working on his article.

These preferences are intransitive. Yet, again, what is crucial in the telling of the story here is that the philosopher only *thinks* about the pairwise comparisons one at a time. The point here is delicate. We shouldn't say on this basis that the philosopher doesn't *have* the preferences simultaneously and at one point in time. That would neglect the way in which one can have non-occurrent, dispositional preferences. It would also rob requirements like transitivity of their applicability, since violations of such requirements are paradigmatically revealed by a *series* of choices. If a series of choices at different points in time revealed nothing about stable, dispositional, simultaneous underlying attitudes, then it would reveal no violation of any synchronic transitivity requirement. So I do not think we should say that this isn't a violation of transitivity.

But again, consider what it would be for such a violation to be fully transparent to the agent. Imagine the philosopher having all three options vividly before his mind, and sincerely declaring, "I prefer working on my article to volunteering, I prefer volunteering to watching TV, and I prefer watching TV to working on my article." Again, this sounds like a joke. Once the philosopher vividly attends to the intransitivity, he will feel a pressure to resolve it. If he does not, these are not all-things-considered preferences, but only *pro tanto* desires.

Yet he can get away with never vividly attending to it. Here's one way that it's depressingly likely to go: by focusing first on the choice between the article and the volunteering, he rules out the volunteering and puts that out of his mind. Then he compares the article and the TV, and picks the TV. So, he ends up watching TV, never attending to the comparison between volunteering and watching TV.

(c) Inter-level coherence

For our third example, let us turn to a requirement on doxastic states only, that I will call, as I have elsewhere, "inter-level coherence".[17] Inter-level coherence bans incoherent combinations of first-order and higher-order doxastic attitudes, where the latter are judgments about which first-order attitudes one's evidence supports. For instance, it forbids believing p while also believing that one lacks adequate evidence for p. Some think of this as a requirement forbidding "epistemic akrasia" or as a doxastic variant of the enkratic requirement, which (in its practical guise) requires one to intend to do what one believes one ought to do.[18] But (at least as I am understanding it) inter-level coherence refers not to one's judgments about what one ought to believe but rather to one's judgments about what one's evidence supports. At least given this specification, I think there are important disanalogies between inter-level coherence and the enkratic requirement, as our discussion will eventually bring out. So I prefer to give it its own name.

Again, we can make sense of violations of inter-level coherence.[19] For example, suppose that Fabian considers himself to be extremely attractive to most members of the opposite sex. Suppose also that Fabian is aware of a body of psychological research that shows that people like him tend to systematically overestimate their attractiveness to the opposite sex, and that the women he tries to seduce often ask him to leave them alone. When Fabian reflects on all of this, he is inclined to admit that his evidence that he is extremely attractive to most members of the opposite sex is pretty lousy. But he doesn't like to dwell on that. When he starts to think like that, he just jumps in his sports car, rolls down the windows, turns the volume on his stereo up to 11 and goes for a spin, and very soon he stops thinking about it. His belief that he is extremely attractive to most members of the opposite sex survives.

Again, one could argue that there is never a single moment where Fabian both believes that he is extremely attractive to most members of the opposite sex and believes that this belief of his not well-supported by the evidence. On this diagnosis, in his moments of reflection his belief that he is extremely attractive is suspended, so that he no longer counts as believing, while at all other times he does not count as believing that his belief is well-supported by the evidence. But while we can perhaps imagine the case that way, I also think that saying that this is the only way of making sense of the case is *ad hoc*. Most of the beliefs that we have at any particular point in time are not occurrent at that exact

---

[17] See Worsnip (forthcoming) for an explication and defense.
[18] Cf., e.g., Horowitz (2014).
[19] Interestingly, a number of philosophers have been tempted to suggest that violations of what I'm calling inter-level coherence are impossible *simpliciter*: see Hampshire (1965); Pettit & Smith (1996: 448); Setiya (2008: 43); Adler (2002a, 2002b); Hurley (1989: 130-5, 159-70); Wallace (2001: 12-13). I think this over-reaches, for the reasons I'm about to explain.

moment; we can put many of our beliefs out of our minds while still counting as believing them. I believe that when I was nine years old a very embarrassing incident occurred at my birthday party, but fortunately I manage to put it out of my mind most of the time. When I do so, I don't count thereby as suspending judgment about whether the incident occurred. I don't see why Fabian can't be the same. Most of the time, he manages to put the meagre evidential basis for his belief in his own attractiveness out of his mind. He doesn't thereby cease to count as believing that the evidential basis for his belief is meagre.[20] Meanwhile, his first-order belief that he is extremely attractive continues to play a role in explaining his behavior, in a way that makes it right to attribute that belief to him.

But – again – what is hard to make sense of in Fabian is a persistent, stable state whereby he consciously and transparently violates inter-level coherence. There is something incredibly odd about an utterance like "all my evidence suggests that I'm not very attractive to members of the opposite sex. Nevertheless, in fact I am very attractive to members of the opposite sex." Again, it sounds like a kind of joke. There is a strong pressure to interpret the agent either as not really believing that his total set of evidence suggests that he is not attractive, or as not really believing that he is attractive. One of the cognitive states may be weaker: it may be a fantasy, or a wish, or a hope, or an assumption, or faith, but not a belief.[21] Part of what it is for something to be a belief, in contrast to these weaker states, is for it *not* to be reflectively sustainable in the face of an acknowledged judgment that it is not supported by the evidence.

I think that something like the story that I am sketching here is implicit in many explanations explain why it is so hard for us to "believe at will". In at least the most paradigmatic cases of (trying to) believe at will, one tries to believe something for pragmatic reasons, despite taking oneself to lack evidential grounds for this belief. Numerous prominent philosophers have claimed that this is because our beliefs are in some sense "controlled" by our evidence.[22] One finds writers saying things like "believing in opposition to one's evidence is motivationally unintelligible,"[23] or "one particular belief-forming process, reasoning, is regulated solely by evidential considerations,"[24] or "belief aims to 'track truth' in the sense that belief is subject to immediate revision in the face of changes in our all-things-considered evidence."[25]

On their most obvious interpretations, however, these claims are false, for a simple reason: we fail to believe in line with what our evidence supports absolutely constantly. Moreover, such failures are *not* always results of non-transparency: they can simply be the result of mistaken assessments of what the evidence supports. However, if the picture I am suggesting is right, there may be a truth in the neighborhood here. What really may be the case, and what these writers should have said (or, if

---

[20] C.f. Elga's (2005) diagnosis of his own state of mind with respect to overrating his own abilities.

[21] This is one way in which apparent transparent violations of (ILC) can be diagnosed as something else. There are others: for example, one may believe without being sure what the evidence supports, or one may believe while thinking that the evidence does not decide between competing doxastic attitudes. Between them, I think these cases take care of apparent counterexamples to the claim that one cannot transparently violate (ILC) – for example, counterexamples involving religious belief. A full defense of this is not possible in the space available here, however.

[22] Cf., e.g., Williams (1973), Foley (1993: 16), Adler (2002b), Hieronymi (2006), Velleman (2000).

[23] Adler (2002a: 8).

[24] Shah (2003: 462). Also (*ibid.*: 469): "my deliberation won't count as belief-formation…unless the deliberation is solely influenced by evidence."

[25] Gendler (2008: 565). See also Noordhof (2004: 247).

we're feeling charitable, meant to say) is that our beliefs, when they are formed reflectively in ways that are transparent to us, are controlled by our *judgments* about the evidence.[26] In other words, we cannot reflectively sustain transparent inter-level coherence. Arguably, that may explain why it is so hard to believe at will, where this involves transparently defying one's own judgment about the evidence.[27]

III. Taking stock: the emerging picture of coherence

Here is the picture that is emerging. Suppose we have an apparent violation of a coherence requirement. On my account, (at least) one of two things must be the case. First, it could be that the agent's mental states are not fully transparent to her on this occasion. In that case, the violation could be perfectly genuine. Second, though, it could be that we ought not really to attribute to the agent the mental states that violate the coherence requirement. The idea here, which I have tried to make plausible, is that an apparent intention, belief or other attitude that, given the agent's other mental states, will put her in sustained, *transparent* violation of a coherence requirement *does not really count* as an instance of that attitude: it is not intention or belief proper, but something less – for example, instead of an intention, a wish or desire; instead of a belief, a pretense or a supposition.

That may seem like too much of a *just so* story. I have tried to motivate it through my individual examples. But I also want to say something broader. On any account of attitudinal mental states, we need something that will distinguish particular attitudinal mental states from others: that will explain just *when* something is not merely a supposition but a belief; not merely a desire but an intention. My proposal is that the way to do this is by appeal to a particular kind of disposition that agents have to revise these mental states. For example, it is part of what it *is* for one to have an intention (rather than, say a desire) that one be disposed such that if one finds oneself *transparently* holding that intention and a belief that some means is required for the carrying out of that intention, but not intending the means, one revises one's attitudes such that one either comes to intend the means, gives up the means-ends belief, or gives up the original intention. If this is not the case, one falls short of genuinely *intending*.[28]

This puts us in a position to answer several possible objections to the present account. First, the charge of *a priori* psychology. There are, of course, many questions about our attitudinal mental states that cannot be answered without detailed empirical investigation. But there are also prior philosophical questions: what *is* it for something to be a belief, or an intention? How do we know what to *look for* when we do our empirical investigation of our belief and intention-forming practices? Now, on the account I have developed, the claims about how it is hard (or even impossible) to hold two attitudes jointly under conditions of full transparency turn out to *fall out* of our best answers to these prior philosophical questions. They are, thus, *not* empirical predictions made from the armchair.[29] To the extent that a distinction between observation and philosophical theory is possible, they do not

---

[26] See also Winters (1979) and Setiya (2008) on this point in the context of the debate about believing at will.

[27] I tried to sketch such an explanation in a chapter of my PhD dissertation (Worsnip 2015: ch. 4).

[28] Must there be some deeper explanation of such dispositions are constitutive of the mental states in question (for example, in terms of the functional profile of the mental states)? I leave this open. It is consistent with my account if there is.

[29] See also Blackburn (1998: 54-59).

rule out any particular observational data, but are rather claims about how to philosophically interpret such observational data. It is thus not outrageous to say that these claims apply to *all* agents. If it's part of what it *is* for something to be a belief that an agent who has a belief must be disposed in a certain way, then it is not overgeneralization to say that such dispositions must be present in all agents. Indeed, any theory of a mental state (such as belief) as meeting a condition or set of conditions C rules out, in advance, the possibility of agents who meet C and lack belief (or have belief and fail to meet C). So any objection to my account on these grounds would generalize to any such theory.

Relatedly, here is an objection that can be cooked up to any claim that some particular coherence requirement is hard (or impossible) to transparently violate: "consider person X. Person X reports herself as violating this coherence requirement. Surely person X is possible. But it's an *ad hoc*, theory-driven move to say that person X must be mistaken when she reports her mental states. So it is possible to transparently violate this coherence requirement." Such an objection can be sharpened if we make person X a sophisticated person with a philosophical theory that rationalizes the combination of mental states from their point of view. They might be someone who (albeit wrongly) denies that the states in question are, in fact, incoherent. For example, for the noncontradiction requirement, we can imagine a dialetheist who thinks there can be true contradictions. Surely, it is said, a sophisticated person might transparently violate this requirement by believing p and believing not-p. The dialetheist's mental states, it will be said, may still be irrational or incoherent,[30] but surely they are still *possible*.[31] How can I say that a sophisticated dialetheist doesn't even know her own beliefs?!

I won't disagree that we can always imagine someone who *says* that she (transparently) violates some coherence requirement. The question is whether we are beholden to interpret her attitudes in the way that she herself reports them. Again, the problem with the assumption that we are is that it can be used to generalize the above objection to *all* accounts of individual attitudinal mental states. Suppose you have a theory of belief of the generic form: for an agent to believe some proposition p just is for that agent to satisfy condition C with respect to p. Now, I can object: well, I can imagine someone who reports herself as believing p, but as not satisfying condition C with respect to p. And, I might add, this person has a sophisticated theory of belief, which involves rejecting the claim that believing p is a matter of satisfying condition C. Surely, I may now say, it's an *ad hoc*, theory-driven move to say that she must be mistaken when she reports her mental states. Therefore, I conclude, believing p is *not* a matter of satisfying condition C. Thus, we have a recipe for objecting to *any* theory of belief (or indeed any other attitudinal mental state) in terms of some condition or set of conditions.

Thus, any non-vacuous theory of belief will issue the verdict that particular subjects, even sophisticated subjects, misclassify themselves as having or not having particular beliefs. If this is a problem for my view, it is a problem for every view. Once the generality of the objection is laid bare, I think that it is apparent that it is too quick. Of course someone can *say* that they have an attitude even though it puts them in transparent violation of a coherence requirement; but if we have good theoretical and intuitive grounds to say that subjects misclassify their own mental states, we should not be held hostage to their self-attributions.

---

[30] In fact, as I'll point out a few paragraphs down, the objector *has* to concede this, otherwise the case is not a counterexample to my view.

[31] Bruno Whittle, among others, pressed this objection particularly forcefully.

Indeed, in the case of the dialetheists specifically, I *do* find it hard to make sense of what it means to attribute a transparent state of believing p and believing not-p to someone, no matter how much a person professes those beliefs.[32] And my view in not alone in this: it's not clear how a dispositional theory of belief, for example, or a possible worlds theory of belief, makes sense of such a person. Of course, someone else may disagree with me here, and think that it is easy to make sense of a dialetheist's beliefs, even under transparent conditions. But, I submit that, to the extent that you think that, you think that the dialetheist *isn't really incoherent*, and thus you reject the noncontradiction requirement, at least in its universal and exceptionless form. In that case, the example is no threat to the theory of incoherence that I have offered here. If intuitions about the incoherence of states on one hand, and the sustainability of such states under conditions of transparency on the other, pattern together, despite disagreement about each individually, then we actually have *confirming evidence* for the view that I am advancing here.

Next, the present account makes explicit why coherence requirements, independently of their *normative* status, have an important philosophical role to play; one that makes good on our earlier promise to provide ontological justification for them (against eliminativists), and that helps us to see their distinctness from substantive norms. Specifically, coherence requirements – or, more specifically, the assumption that agents are disposed to obey such requirements under conditions of full transparency – play a constitutive role in our attributions of mental states. We can attribute mental states to individuals by backgrounding other mental states of theirs and seeing what is needed to make coherent sense of their intentions as manifested in their behavior. Likewise, the present notion of coherence will play an important role in explanation and prediction in social science.

If this sort of story sounds familiar, it should: it is highly reminiscent of the "interpretationist" theory of belief and other mental states associated especially with Donald Davidson.[33] Davidson memorably claimed that in order to be able to interpret other agents, we have to assume that these agents are rational. However, the view developed here represents a particular version and precisification of this view. When Davidson says that we must assume that agents are rational, many have interpreted him as claiming that to attribute mental states to agents, we must see them as conforming to *substantive* norms: as doing and believing what they really have reason to do and believe; as believing the truth, and pursuing the good.[34] This assumption seems to many to be crazy. Human beings are very bad at doing what they have most reason to do. Indeed, on many plausible enough normative theories – views that combine a demanding view of morality with the view that morality gives us weighty and categorical reasons – we almost *never* do what we have most practical reason to do. It would be very odd if our philosophical psychology ruled these substantive normative theories out in advance. Likewise, the idea that humans conform to substantive norms seems a very poor basis on which to explain or predict their behavior, which is why many social scientific claims about

---

[32] Importantly, this isn't to say that the dialetheist can't (transparently) hold the theoretical belief that there are true contradictions. The view I am suggesting doesn't say that this theoretical belief is incoherent; rather, it says that it is incoherent to believe a contradictory pair of propositions. So it's not like the view says that *dialetheism*, the theory, is impossible to (transparently) believe.

[33] See, e.g., Davidson (1980: essay 12, 2004: essay 6). See also, *inter alia*, Dennett (1971).

[34] I am not sure that this common interpretation of Davidson is correct; arguably he intends to restrict his interpretationist doctrine at least primarily to rationality in the sense of coherence. See especially Davidson (2004: essays 11-12, e.g. p. 170).

rationality (especially those in economics), when interpreted as claims about substantive norms rather than coherence requirements, sound so absurd.

My claim is that things are quite different when it is coherence requirements that are at issue. On my view, it is not a precondition of interpretation that we assume that agents do or believe what they have most reason to do. But it is a precondition of interpretation that we assume that agents are disposed to be coherent under conditions of mental transparency.[35, 36] Take a variant of a classic, well-worn example: I know that Tim intends to drink a beer, and I see him heading for the fridge (manifesting his intention to open the fridge). On that basis, I attribute the belief that there is beer in the fridge to Tim. I am assuming that Tim's intentions, desires and beliefs all fit together coherently in the right ways here. If I didn't think that that, I would have no reason to favor attributing the belief that there is beer in the fridge, rather than, say, the belief that the fridge is empty and that the only available beer is in the garage. This latter interpretation literally doesn't *make sense* of Tim's behavior, namely his heading for the fridge rather than the garage. An assumption of coherence is thus needed to attribute mental states to Tim, and to explain and predict his actions.

However, I need not assume that Tim is really responding to his substantive reasons. Perhaps he *ought not* to be drinking beer; perhaps he ought to be attending his child support hearing. Perhaps he ought not even believe that there is beer in the fridge: in fact, he is basing his belief on a vague memory of having put the beer there, and he could well have drunk it last night and forgotten, or it could well have been taken by his brother Billy in the intervening time. Neither of these possibilities interfere with me reading the belief that there is beer in the fridge off of his behavior. They do not make that behavior unintelligible in the way that his actually *believing* that there is no beer in the fridge would.

The present account thus does assume that humans are at least generally *coherent*. But it does not assume that that they even generally successfully do or believe what they have most reason to do or believe. Whether it assumes that they are rational or not, then, depends on whether the property of being rational is understood in terms of coherence or in terms of reasons-responsiveness. Though it is often claimed by psychologists, behavioral economists and others that humans are persistently irrational, it is far from clear that the instances of "irrationality" that they point to generally amount to transparent violations of coherence requirements, as opposed to violations of coherence requirements under conditions of non-transparency, or failures to respond to substantive reasons.[37]

---

[35] In claiming that it is the notion of rationality as coherence, and not that of responding to one's substantive reasons for action, that is of primary importance in this explanatory context, I am agreeing with Ridge (2014: 238-9).

[36] The "under conditions of mental transparency" qualification also represents a revision of interpretationism as standardly understood. Stich (1990: 39-41) helpfully distinguishes three versions of the view: (a) the view that *perfect* rationality has to be assumed for interpretation; (b) the view that satisfaction of only a "fixed bridgehead" of core requirements needs to be assumed for interpretation; and (c) the view that, for each requirement, we need only assume for the purposes of interpretation that it is satisfied *most* of the time. To this we can add (d) the view that we need only assume that the requirement is satisfied *under particular conditions*. My version of interpretationism is a combination of views (b) and (d), with the possible addition of (c) for some cases (see section IV below).

[37] Sometimes where incoherence or "fallacious" reasoning is quickly claimed by psychologists or economists, closer philosophical reflection reveals that it is really a more substantive question of failing to respond to one's reasons, rather than some more generic form of incoherence, that is at issue. See, e.g., Kelly (2004) on the "sunk cost" fallacy, and Sen (1993) and Rulli & Worsnip (2016: 220) on violations of the "independence of irrelevant alternatives".

Strictly speaking, I have made two claims about incoherent combinations of attitudes. First, I've said (as the account says) that when attitudes are jointly incoherent, it's constitutive of those attitudes that agents are disposed not to sustain them jointly under conditions of full transparency. Secondly, I've said that incoherent combinations of states should not be *attributed* to agents under conditions of full transparency. I think of the first claim as being the deeper truth that explains the second. Because it is in the nature of such states that agents are disposed not to hold them jointly under conditions of full transparency, we should not interpret agents as being in such states under conditions of full transparency.

IV.   A hard case: enkrasia

Time for a hard case. We want the present account to deliver an extensionally adequate account of incoherence that covers the most obvious examples of coherence requirements. We cannot consider every putative coherence requirement. So let's focus on one that threatens to make particular trouble for the account: the "enkratic" requirement.

According to numerous theorists of practical rationality, there is a coherence requirement forbidding akrasia: that is, forbidding one from simultaneously believing one ought to Φ but not intending to Φ.[38] But there is widespread consensus that clear-eyed akrasia is (all too) psychologically possible. One can think that one ought to do something, and have this thought quite clearly at the front of one's mind, but realize that one's intentions fail to match up to what one ought to do. In such a case one may feel a pressure to revise one's intentions in the simple sense that one believes one ought to do so, but this may exert little motivational force; one's intentions may remain as they are without there being much of a puzzle about what is going on. This appears to be a violation of a coherence requirement without any kind of failure of transparency. Yet we still want to describe akrasia as irrational, in the sense of rationality that deals with coherence. We thus seem to have a counterexample to my claim that two or more states are incoherent only all agents must be disposed not to sustain them under conditions of full transparency.

One could try to straightforwardly resist this counterexample to my account in one of two ways. One way would be to hold a very hardline version of motivational internalism about normative judgment, on which if a putative normative judgment does not produce an intention to comply with that judgment (under conditions of transparency), it isn't a normative judgment after all. This may have been R.M. Hare's view.[39] This mirrors the treatment we gave of putative transparent violations of the instrumental requirement and of inter-level coherence, where we said that part of what certifies something as a genuine intention or belief is that it conforms to the relevant coherence requirements governing those states. On this view, a similar thing can be said about normative judgment (i.e. normative belief) in particular and the enkratic requirement. Putative normative judgments that do not produce intentions should actually be classified as other cognitive states (perhaps as purely descriptive beliefs or as what Hare called "inverted commas" judgments).[40]

---

[38] Cf. esp. Broome (2013); see also (among many others) Kolodny (2005), Scanlon (2007), and Setiya (2007).
[39] Cf. Hare (1952: 19-20, 169-70).
[40] Hare (1952: 164-5).

While I have more sympathy for this line of thought than many philosophers do, I still believe that as it stands it is unreasonably strong. In the cases of instrumental irrationality and inter-level coherence, it was (I think) not too much of a stretch of the ordinary notion of belief and intention to say that the putative transparent violations of the requirements in fact involved something less than full-blown belief and intention. But it is a real stretch to say that the ordinary notion of normative judgment does not allow for an intelligible notion of (clear-eyed, or transparent) akrasia. We want to reach a reflective equilibrium in our theorizing about mental states between of our ordinary folk notions of these states, as they play a role in our practices of mental state attribution, and the notions of such states that will allow us to give a systematic, illuminating and unified theoretical account of them. The present proposal gives the second aspect too much weight over the first to keep us in genuine equilibrium.

The second line of resistance would be to simply deny that akrasia really is incoherent, at least in the sense that we have identified and in which violations of the other requirements we have discussed are incoherent. Again, I have some sympathy with this. There is, I think, a good sense in which someone who says "there's conclusive evidence that giving to charity saves lives, but giving to charity doesn't save lives" is incoherent in a deeper way than someone who says "I ought to give most of my earnings to charity, but I'm not going to do so".[41] It is precisely the fact that the former is harder to make sense of than the latter than seems to make it appropriate to brand the former as a more radical kind of incoherence. This reveals an asymmetry between practical akrasia and inter-level incoherence that questions whether they should really be thought of as pure analogues of one another (as calling the latter "epistemic akrasia" suggests).[42] Nevertheless, again as it stands the proposal feels too strong. Rejecting the enkratic requirement wholesale, and saying that there is nothing irrational about believing one ought to do something but not intending to do it, is a drastic move.

So what I propose is a kind of compromise between the two lines of resistance that moderates each one. We should allow that incoherence is something that comes in degrees, and that violations of some requirements are more incoherent than others.[43] For example, violations of inter-level coherence (or indeed of the instrumental requirement) are more incoherent than violations of the enkratic requirement; but the latter are still somewhat incoherent. This can be accommodated by our account of incoherence by saying that, correspondingly, the strength of the disposition not to sustain attitudes jointly (under conditions of full transparency) can also come in degrees. The most incoherent sets of mental states are ones whereby the disposition is so strong that it cannot be blocked; these sets

---

[41] Again, Wallace (2001) brings this out.

[42] In Worsnip (forthcoming), I argue that there is a rationale for inter-level coherence that finds no analogue in the practical case.

[43] Fogal (ms.) also thinks that incoherence comes in degrees, but he takes this to count against an account of incoherence that is framed in terms of requirements, and in favor of an account that is framed in terms of "pressure". I am not persuaded that the degreed nature of incoherence makes talk of requirements inappropriate. It may be that the strength or 'force' of a requirement can itself come in degrees, or alternatively that something's being a requirement is a matter of its forbidding states that are *sufficiently* incoherent, where there is some minimum threshold for this sufficiency. However, my requirements-talk in this paper could be translated into pressure-talk if it needed to be. For an illuminating study of degrees of incoherence of credence functions in a Bayesian setting, see Staffel (2015).

of states will be *impossible* to sustain jointly under conditions of full transparency.[44] But in less incoherent cases, such as akrasia, the disposition is weak enough to sometimes be blocked.

We can then agree with the motivational internalist that it is partially constitutive of normative judgment that the agent have *some* disposition not to be in a sustained and transparent state of holding that normative judgment while having no intention to comply with it.[45] One of the ways that we get a grip on what it is to make a normative judgment is by focusing on this motivating role. Accordingly, there is *some* pressure on us not to interpret agents as consistently defying their own normative judgments. But this pressure is not always insurmountable: if enough other markers of normative judgment are there, we can attribute clear-eyed akrasia to agents nevertheless.

Some may wish to say that this is the right way to handle the other examples of coherence requirements we considered – allowing the disposition not to hold such states jointly under conditions of full transparent to be blocked in certain cases. This would allow for the metaphysical possibility of sustained and transparent violations of those requirements also. It would be enough that agents must, to count as having such states, must have some disposition not to engage in such sustained and transparent violations – and so that the assumption that agents fulfil these requirements plays *some* constitutive role in our mental state attribution. Although I am tempted to take a harder line on at least some coherence requirements, as shown by my treatment of cases in section II, I would still count this stance as a version of my view.

V.   Naturalism and normativity

This completes my defense of my account of (in)coherence and of coherence requirements. For any individual requirement, there is of course room for dispute. But I think that, in general, to the extent that someone disputes that agents must be disposed not to transparently sustain violations of some putative requirement, they will also be inclined to dispute that it really is incoherent to violate this putative requirement: that is, to dispute that the requirement is genuine. As I have already said, if this is so, it is actually confirming evidence for my account.

As I said at the start, the account I have offered might be understood as a form of reductive realism about coherence and coherence requirements. It identifies the property of coherence, as it attaches to sets of mental states, with a psychological property of these sets of mental states: that agent's being disposed not to sustain them jointly under conditions of full transparency. Reductive accounts in metaethics tend to be identified as 'naturalistic', and I do think this label can be applied to my view, though one hesitation one might have concerns whether psychological states themselves are

---

[44] I am assuming here that the case where it is for an agent *impossible* to sustain a combination of mental states jointly under conditions of full transparency, this can still be made sense of in terms of talk of dispositions. A disposition that always manifests itself under its triggering conditions, and that cannot be defeated, is a very strong kind of disposition. So an account in terms of agents' dispositions can make room both for the cases where sustaining the mental states jointly is *impossible*, and those where the strength of one's disposition does not amount to an impossibility to do otherwise. Thanks to Geoff Sayre-McCord here.

[45] For congenial views, see Jackson & Pettit (1995: esp. 35-38) and Blackburn (1998: 61).

understood "naturalistically" on the kind of interpretationist view that I have expressed sympathy for.[46] In any case, my view is as naturalistic about coherence requirements as it is about mental states.[47]

Now, one might in principle agree with me that these two properties are co-extensive – that is, that incoherent states are states that are such that we are disposed not to sustain them under conditions of transparency, and vice versa – without agreeing that they are identical.[48] After all, this move is often made by non-naturalists in metaethical debates about moral properties. I don't have a knockdown objection to this view, and I think that the view I've offered would still be interesting if it were understood as a claim about (mere) necessary coextension rather than property identity. However, I do think that there is one way in which the move from necessary coextension to property identity is on stronger ground when it comes to coherence as compared to morality, as I'll now explain.

Reductive naturalists about *morality* often say that property identity is the best explanation of the co-extensiveness and thus of the supervenience of the moral on the natural.[49] But non-naturalists have a reply to this: that the co-extensiveness of moral rightness with certain natural properties is explained by one or more very general, irreducibly normative, principles that specify that an act is right iff it has some particular natural features. If these principles are necessary, then we get an explanation of the co-extensiveness of moral and natural properties, and of the supervenience of the moral on the natural.[50] This is not the place to adjudicate the adequacy of this reply. What I want to point out is that such a non-naturalist reply is a non-starter in the case of coherence requirements. There is not, on any view I know of, some master normative principle of coherence that enjoins one not to be in states that one is disposed not to sustain under conditions of transparency. So there is no fundamental, irreducibly normative principle that can explain the co-extensiveness of the property of coherence and

---

[46] Dennett and Davidson themselves do something to encourage this way of thinking, but I am not convinced that it is really correct. The thought seems to be that if our being disposed to obey particular requirements of rationality is built into mental states, and our being disposed to obey particular requirements of rationality is in an important sense an *a priori* rather than an empirical truth (as I myself stressed earlier), then those mental states are not themselves empirically investigable using naturalistic methods. But I don't think this follows. As I stressed earlier, on *any* view, there are prior conceptual or metaphysical questions about what *counts* as (for example) a belief, that are not settled empirically, but this doesn't mean that beliefs aren't empirically respectable entities on any such view; having answered the conceptual question, we can go on to do empirical inquiry by looking for things that do and don't meet the conceptual conditions for belief.

[47] A related but distinct worry is that the account isn't *reductive* because it explicates coherence in terms of dispositions to sustain or give up mental states, but those mental states themselves have to be understood partly in terms of coherence requirements. Thus, at best it is a virtuously circular explication; there is no reductive analysis in the offing. (The idea that the mental states have to be understood partly in terms of coherence requirements might also be an alternative way of making good on the worry that mental states aren't "natural" on this account.) But there is a subtle sleight of hand in this objection. Strictly speaking, all that needs to enter into the attribution of mental states is the assumption that we are disposed to combine or not combine mental states in various ways. Nothing *needs* to be said about coherence or about coherence requirements; the dispositions can be described without any reference to such things. My account does say that we should *then* identify incoherence with those sets of attitudes that we are disposed not to combine (under specific conditions). So, it will be true that the dispositions not to combine states are dispositions not to be incoherent. But it is the disposition that comes first, as it were, and the incoherence that arises in virtue of this. Thus, there isn't in fact any circularity here; coherence is still being reduced to psychological dispositions that can themselves be described without reference to coherence.

[48] As always with reductive naturalist views, even property identity claim does not commit us to the claim that there are not different *concepts* here that pick out the same property.

[49] See, e.g., McPherson (2012). A stronger and more contentious idea is that the co-extensiveness actually *entails* property identity (cf. Jackson 1998: ch. 5; Brown 2011).

[50] Cf. Enoch (2011: ch. 6); Scanlon (2014: ch. 2).

the property of being sustainable under conditions of transparency. Thus, the claim that those properties are in fact the same property seems the most plausible way to go. This suggests that the naturalist is better off (or *even* better off, depending on your sympathies) in the case of coherence requirements than in the case of substantive moral requirements.

This connects to the more general point I foreshadowed at the start, that I see no reason to think that the account I have offered offers any particularly strong support for reductive realism about other requirements or normative claims, for example substantive moral requirements. When you violate a moral requirement, everything about your mind can be transparent to you, and yet you may feel no pressure to change your attitudes or behavior. This point is slightly delicate. It is, arguably, true that if you *acknowledge that you are violating a moral requirement* (where this ascription is read *de dicto*, such that you acknowledge not only that you are behaving the way in question, but *that your behavior violates a moral requirement*), you will feel some pressure to revise your attitudes or behavior. But if this is so, it is because the *acknowledgment* that you are violating a moral requirement involves you in *incoherence* (specifically, akrasia). Without the acknowledgment that you are violating a moral requirement, the violation need occasion no psychological pressure to change course – even if everything in your mind is transparent to you. So, it is the (transparent) violation of the coherence requirement (of enkrasia) that does the psychological work here, not the transparent violation of the moral requirement itself.

By contrast, when it comes to the violation of a coherence requirement, no *acknowledgment* that you are violating a coherence requirement is required for the relevant psychological phenomenon to take place. For instance, when you realize, transparently, that you intend to tell your friend that her partner is cheating on her, that the only way to tell her this is to call her today, but that you do not intend to call her today, you feel some pressure to revise one of these states. For this pressure to be felt, you need not accept any theoretical claim that failing to revise your attitudes would put you in violation of a coherence requirement.[51] Transparent awareness of the fact *that you have the attitudes in question* (viz., the intended end, the means-ends belief, and the absence of an intention to take the means) suffices. So coherence requirements are not on a par with other sorts of requirements, such as moral requirements, here.

Because the metaethical story I have offered for coherence requirements does not obviously extend to other kinds of requirements, the naturalistic, reductive realism about coherence requirements that I advocate yields at best very weak support to naturalistic, reductive realism about other kinds of requirements. I myself am much less sympathetic to naturalistic, reductive realism about moral requirements than to naturalistic, reductive realism about coherence requirements.[52] The possibility that we might give a very different metaethical story for coherence requirements and for other, substantive requirements – such as moral requirements – reinforces the distinctness of coherence requirements on one hand and these other, substantive norms on the other.

No doubt some will think that what this shows is that, on the account developed here, coherence requirements are not really normative, or perhaps, after all, that they are not really

---

[51] Some philosophers are tempted to say that one is only truly incoherent when one *recognizes* that one is incoherent (see e.g. Davidson 2004: 177, but compare p. 195), but that way a vicious regress lies. See Broome (2013: 91-93).

[52] Similarly, Ridge (2014) advocates a broadly reductive realism about "rationality" (by which he means, coherence), but an "ecumenical expressivism" about *moral* judgments as well as other substantive normative judgments. I think something similar is implicit in Blackburn (1998), though he would not put it in these terms.

requirements, in any good sense. They have, these people will say, stopped functioning as normative requirements and become merely descriptive truths about psychology. As I said in the introduction, I in fact want to be neutral on whether coherence requirements are normative in the most robust sense of that term. Nevertheless, let me make three preliminary points in reply to this worry.

First, to the extent that there is a phenomenon to be saved here, it is not most naturally characterized in terms of the language of 'normative requirements' but rather in terms of the language of 'rationality'. What we want to earn the license to say is that incoherent combinations of mental states are irrational. But the account of coherence that I have given *does* connect with a long and venerable tradition of thinking about rationality that is recognizable in the ways that ordinary people actually use the concept. One important idea is that that rationality consists in intelligibility: irrationality is a way of being harder to make sense of as an agent.[53] Another is that rationality is the norm – in the 'default' or 'statistical' rather than the robustly normative sense of 'norm' – on the basis of which we predict human behavior and ascribe mental states – an assumption engrained both in academic social science and in ordinary talk. So, even if this notion of coherence turns out not to count as normative in the most robust sense, I do not think that it "changes the subject" (though it may productively distinguish itself from the jumble of things that we can be talking about when we talk about rationality).

As I said at the start, coherence requirements can earn their ontological keep by showing themselves to a well-defined, unified category of phenomena that are philosophically significant in some way, where this significance is not exhausted by normative significance, but also includes explanatory significance, significance for the attribution of mental states, and so on. I agree that if the list of coherence requirements were picked on some arbitrary basis, with no well-defined method, nor guiding conception of what coherence in general is, *and* they turned out to have no normative significance, then they would have little significance *tout court*. My primary aim here has been to show that this is not the situation we find ourselves in: there is a more well-regimented and theoretically significant notion of coherence available, irrespective of its normative status. My aim has not been to vindicate the normativity of coherence requirements. If someone concedes that I have been successful in my aims, but simply insists that the word 'requirement' must be expunged from my account, I will feel contented overall.[54]

Second, however, it is not just *obvious* that in giving a reductive, naturalistic, account of coherence requirements, we thereby preclude them from counting as normative even in a robust sense. At least, this should be a matter for debate. Naturalists about *moral* norms are often accused of being unable to account for the normativity of morality[55] – and while such criticisms *may* hit their target, it is not just *obvious* that they do so. Just because we have identified the property of coherence with a psychologically describable property does not immediately entail, without argument, that this property cannot be normative.[56]

---

[53] Cf., again, Davidson (2004) and Ridge (2014: 238-9).

[54] If it is, as I suggested in the previous paragraph, perfectly natural to talk of coherence as a necessary condition for rationality, then there is at least one sense in which there are coherence *requirements*: there are conditions of coherence that one must satisfy if one is to count as being rational. This is what Broome calls the "property" sense of 'requirement' (cf. Broome 2013: 109-110).

[55] Cf., e.g., Nagel (1986: ch. 8); Parfit (2011: volume 2); Enoch (2011: ch. 5, esp. 107-8).

[56] See Schroeder (2005) for a defense of reductionism in the face of such arguments. See also Railton (1989).

Third, as participants in both the debate about the normativity of rationality and the debate about reductive realism in metaethics have noted,[57] the term 'normative' can be used to stand for a range of things, and some requirement might be normative in one sense but not in another. Naturalist realists tend to argue that morality is (necessarily) normative in some but not all of these senses.[58] It may turn out that coherence requirements are as naturalist realists say that moral requirements are, in this respect. Moreover, those who are dissatisfied with anything other than the most robust kind of normativity in the moral case need not take the same view when it comes to coherence requirements. For the intuitive appearance that coherence requirements are normative *is*, in my view, somewhat less robust – both in content and in force – than the intuitive appearance that moral requirements are normative. Ultimately, there may be less normativity to account for in the former case than the latter.

I cannot resolve these questions here. But I hope that the account of coherence that I have offered lends adequate determinacy to that notion for the debate about the normativity of coherence requirements to be conducted in a reasonably orderly manner.[59] And however we eventually resolve this debate, I hope that the account I've given shows that there is an underlying unity behind the talk of coherence in many of its superficially disjunct guises, and that this unified notion is philosophically interesting independently of its normative status.

References

Adler, J.E. (2002a). 'Akratic Believing?,' *Philosophical Studies*, 110: 1-27.

-------- (2002b). *Belief's Own Ethics*. Cambridge, MA: MIT Press.

Blackburn, S. (1998). *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press.

Broome, J. (2005a). 'Does Rationality Give Us Reasons?,' *Philosophical Issues*, 15: 321-337.

-------- (2005b). 'Have We Reason To Do As Rationality Requires? A Comment on Raz,' *Journal of Ethics and Social Philosophy*, 1.

-------- (2013). *Rationality Through Reasoning*. Chichester: Wiley-Blackwell.

Brown, C. (2011). 'A New and Improved Supervenience Argument for Ethical Descriptivism,' in Shafer-Landau (ed.), *Oxford Studies in Metaethics*, *Volume 6*. Oxford: Oxford University Press.

Copp, D. (2007). 'Moral Naturalism and Three Grades of Normativity,' in his *Morality in a Natural World*. Cambridge: Cambridge University Press.

Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.

-------- (2004). *Problems of Rationality*. Oxford: Oxford University Press.

Dennett, D. (1971). 'Intentional Systems,' *Journal of Philosophy*, 68/4: 87-106.

---

[57] For the former, see e.g. Southwood (2008) and Ridge (2014: ch. 8). For the latter, see e.g. Copp (2007).

[58] See, again, Copp (2007). Some naturalist realists, such as the early Railton (1986), argued that morality is normative (in the most robust sense) only derivatively on instrumental normativity.

[59] One problem with the existing framing of this debate in terms of the normativity of rationality is that it is not obvious that the interlocutors have anything like the same notion of rationality in mind. Kolodny (2005), Broome (2013) and Southwood (2008), for example, have a coherentist notion of rationality in mind, but Ross (2012) and Lord (forthcoming) do not. So it is unclear whether these interlocutors are really disputing the normativity of some single, well-defined thing, or whether the lurking disagreement is really about whether coherence exhausts anything worth calling 'rationality'.

Egan, A. (2008). 'Seeing and Believing: Perception, Belief Formation and the Divided Mind,' *Philosophical Studies*, 140/1: 47-63.

Elga, A. (2005). 'On Overrating Oneself…and Knowing It,' *Philosophical Studies*, 133/1-2: 115-124.

Enoch, D. (2011). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.

Finlay, S. (2009). 'Against All Reason? Scepticism about the Instrumental Norm,' in Pigden (ed.), *Hume on Motivation and Virtue*. Basingstoke: Palgrave Macmillan.

Fogal, D. (ms.). 'Rational Requirements and the Primacy of Pressure,' draft manuscript.

Foley, R. (1993). *Working without a Net*. Oxford: Oxford University Press.

Friedman, J. (ms.). 'Inquiry and Belief,' draft manuscript.

Gendler, T.S. (2008). 'Alief in Action (and Reaction),' *Mind & Language*, 23/5: 552-85.

Greco, D. (2014). 'Iteration and Fragmentation,' *Philosophy and Phenomenological Research*, 88/1: 656-73.

Hampshire, S. (1965). *Freedom of the Individual*. New York: Harper & Row.

Hare, R.M. (1952). *The Language of Morals*. Oxford: Oxford University Press.

Hieronymi, P. (2006). 'Controlling Attitudes,' *Pacific Philosophical Quarterly*, 87/1: 45-74.

Horowitz, S. (2014). 'Epistemic Akrasia,' *Noûs*, 48/4: 718-744.

Hurley, S.L. (1989). *Natural Reasons*. Oxford: Oxford University Press.

Hussain, N. (ms.). 'The Requirements of Rationality,' unpublished manuscript.

Jackson, F. & Pettit, P. (1995). 'Moral Functionalism and Moral Motivation,' *Philosophical Quarterly*, 45/178: 20-40.

Jackson, F. (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.

Kelly, T. (2004). 'Sunk Costs, Rationality, and Acting for the Sake of the Past,' *Noûs*, 38/1: 60-85.

Kiesewetter, B. (forthcoming). *The Normativity of Rationality*. Oxford: Oxford University Press.

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

-------- (2007). 'How does coherence matter?,' *Proceedings of the Aristotelian Society*, 107/3: 229-263.

-------- (2008a). 'The Myth of Practical Consistency,' *European Journal of Philosophy*, 16/3: 366-402.

-------- (2008b). 'Why Be Disposed to Be Coherent?,' *Ethics*, 118/3: 437-463.

Lewis, D. (1989). 'Dispositional Theories of Value,' *Proceedings of the Aristotelian Society Supplementary Volume,* 63: 113-137.

Lord, E. (forthcoming). 'What You're Rationally Required To Do and What You Ought To Do (Are the Same Thing!), *Mind*.

McPherson, T. (2012). 'Ethical Non-Naturalism and the Metaphysics of Supervenience,' in Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 7*. Oxford: Oxford University Press.

Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.

Noordhof, P. (2004). 'Believe What You Want,' *Proceedings of the Aristotelian Society*, 101/1: 247-265.

Parfit, D. (2011). *On What Matters*. Oxford: Oxford University Press.

Pettit, P. & Smith, M. (1996). 'Freedom in Belief and Desire,' *Journal of Philosophy*, 93: 433-36.

Railton, P. (1986). 'Moral Realism,' *Philosophical Review*, 95/2: 163-207.

-------- (1989). 'Naturalism and Prescriptivity,' *Social Philosophy and Policy*, 7/1: 151-74.

Raz, J. (2005). 'The Myth of Instrumental Rationality,' *Journal of Ethics and Social Philosophy*, 1.

Ridge, M. (2014). *Impassioned Belief*. Oxford: Oxford University Press.

Ross, J. (2012). 'Rationality, Normativity, and Commitment,' in Shafer-Landau (ed.), *Oxford Studies in Metaethics, Volume 7*. Oxford: Oxford University Press.

Rulli, T. & Worsnip, A. (2016). 'IIA, Rationality, and the Individuation of Options,' *Philosophical Studies*, 173/1: 205-21.

Scanlon, T.M. (2007). 'Structural Irrationality,' in Brennan, Goodin, Jackson & Smith (eds.), *Common Minds: Themes from the Philosophy of Philip Pettit*. Oxford: Oxford University Press.

-------- (2014). *Being Realistic about Reasons*. Oxford: Oxford University Press.

Schroeder, M. (2005). 'Realism and Reduction: The Quest for Robustness,' *Philosophers' Imprint*, 5/1.

Schwitzgebel, E. (2008). 'The Unreliability of Naïve Introspection,' *Philosophical Review*, 117/2: 245-73.

Sen, A. (1993). 'Internal Consistency of Choice,' *Econometrica*, 61/3: 495-521.

Setiya, K. (2007). 'Cognitivism about Instrumental Reason,' *Ethics*, 117/4: 649-73.

-------- (2008). 'Believing at Will,' *Midwest Studies in Philosophy*, 32: 36-52.

Shah, N. (2003). 'How Truth Governs Belief,' *Philosophical Review*, 112/4: 447-482.

Srinivasan, A. (2015). 'Are We Luminous?,' *Philosophy and Phenomenological Research*, 90/2: 294-319.

Staffel, J. (2015). 'Measuring the overall incoherence of credence functions,' *Synthese*, 192/5: 1467-93.

Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.

Velleman, J.D. (2000). 'On The Aim of Belief', in *The Possibility of Practical Reason*. Oxford: Oxford University Press.

Wallace, R.J. (2001). 'Normativity, Commitment, and Instrumental Reason,' *Philosophers' Imprint*, 1/4.

Williams, B.A.O. (1973). 'Deciding to Believe,' in his *Problems of the Self*. Cambridge, UK: Cambridge University Press.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.

Winters, B. (1979). 'Believing at Will,' *Journal of Philosophy*, 76/5: 243-256.

Worsnip, A. (2015). *Rationality's Demands on Belief*. Doctoral dissertation, Yale University.

-------- (forthcoming). 'The Conflict of Evidence and Coherence,' *Philosophy and Phenomenological Research*. DOI: 10.1111/phpr.12246.