Coherence
Alex Worsnip

Forthcoming in Jonathan Dancy, Ernest Sosa, Matthias Steup & Kurt Sylvan,
A Companion to Epistemology (3rd. ed.), Wiley-Blackwell

Penultimate version – please cite published version if possible

The term 'coherence' (and its antonym 'incoherence') is used in a bewildering variety of ways in epistemology (and in philosophy more broadly). This entry attempts to bring some discipline to uses of the term by offering a taxonomy of notions of coherence (and incoherence), and then surveying which of the resulting notions is (or should be) at work in the various different contexts in which it is deployed.

1.   Notions of coherence and incoherence: a taxonomy

We can start with the following general and vague characterization of coherence and incoherence: a set of mental states is coherent when the mental states in question "fit together" or "hang together," and incoherent when they fail to do so – when they "clash." Note that coherence and incoherence, thus defined, are properties of *sets* of mental states, regarding the relations that these mental states bear to each other. With this general characterization in hand, we can distinguish (at least) two dimensions along which more precise notions of coherence (and incoherence) can differ. First, they can differ according to which mental states they allow to enter into relations of coherence and incoherence. Second, they can differ according to how they construe the relevant relations of "fitting (hanging) together" and "clashing". Let's take these in turn.

      Beginning with the first dimension of difference: some notions of coherence will allow *any* kind of mental state to enter into relations of coherence and incoherence, while others will only allow what we might call *attitudinal* mental states to enter into such relations. Attitudinal mental states include certain doxastic states such as belief, credence, suspension of judgment, as well as certain conative states such as desire, intention, and preference, as well as some attitudes that might be classified as either doxastic or conative (or a hybrid), such as hope. They do not include other mental states such as perceptual experiences, memories, and pains. What is distinctive of attitudinal mental states, as opposed to these other mental states, is that they are each a kind of *stance* that one can adopt toward an object (often a proposition; but sometimes something else, as when the object is desire is, say, a chocolate bar) on the basis of reasons. Call a notion of coherence that allows only attitudinal mental states to enter into relations of coherence (e.g. that of Worsnip 2021: 16-17) an *attitudinal* notion of coherence, and a notion of coherence that also allows other, non-attitudinal mental states to enter into mental states (e.g. that of Wedgwood 2017: 4) a *broadly-encompassing* notion of coherence. The latter notion might, for example, allow certain combinations of beliefs and perceptual experiences to count as coherent or incoherent.

Even on a purely attitudinal notion of coherence, we will also likely want to allow *absences* of attitudes to enter into relations of coherence and incoherence. This is because many cases that are widely treated as instances of incoherence involve a combination of attitudes and absences thereof. (Consider, for example, {believing that it's raining, believing that if it's raining then the streets are wet, not believing that the streets are wet}.)

Turning now to the second dimension of difference, we can distinguish two broad families of notions of "fitting (or hanging) together" and "clashing," which I will call *formal* and *substantive*. Examples of doxastic attitudes that *formally* clash plausibly include {believing p, believing not-p}, {having credence 0.7 in p, having credence 0.4 in not-p}, and (somewhat more controversially) {believing p, having credence < 0.5 in p}. There are also plausibly formal clashes between conative attitudes, such as {preferring A to B, preferring B to C, preferring C to A}, and between combinations of doxastic and conative attitudes, such as {intending to Φ, believing that to Φ one must Ψ, intending not to Ψ}. Such clashes are formal in the sense that the attitudes in question clash in virtue of their *form*, or a general pattern that they instantiate. We can see this easily in the foregoing examples because each of them includes variables such as p, A, and Φ, representing propositions, objects of choice, and actions, respectively. In each case, we can see the incoherence in question just by seeing the form or pattern that the attitudes instantiate; we don't need to know what their full content is – that is, what the relevant propositions, objects of choice, and actions are. Relatedly, formal clashes between attitudes (at least on most views) obtain regardless of the circumstances or evidence of the agent who has them: for example, a belief in p and a belief in not-p will always clash, regardless of the circumstances or evidence of the agent.

On a formal notion of incoherence (e.g. Kolodny 2007, 2008; Easwaran & Fitelson 2015; Worsnip 2021: 8), all instances of incoherence involve formal clashes. Moreover, it is typically assumed that on this formal notion, what it is for attitudes to be coherent is merely for them *not* to be incoherent (though see Fogal 2024 for dissent).

By contrast, on more *substantive* notions of the sort of fitting together involved in coherence (e.g. Harman 1973, BonJour 1985, Berker 2015), avoiding formal clashes is not sufficient for attitudes fitting together. Rather, attitudes fit together when substantive relations of *support* obtain between them. For example, it might be thought that in some good sense, the combination {believing the cat has muddy paws, believing the cat has been outside} fits together better than the combination {believing the cat has muddy paws, believing the cat has not been outside}. Plausibly, this is because a relation of substantive evidential support obtains between the content *the cat has muddy paws* and the content *the cat has been outside*, but not between the content *the cat has muddy paws* and the content *the cat has not been outside*.

On one view (though not by any means the only possible one), this is because the content *the cat has been outside* explains the content *the cat has muddy paws*; the relevant notion of coherence is thus sometimes called "explanatory coherence" (Harman 1973).

Note that the combination that is said to be incoherent (or, perhaps better, *less* coherent) in this substantive sense – that is, {believing the cat has muddy paws, believing the cat has not been outside} – lacks the key features of *formally* incoherent combinations of attitudes. First, it is hard to see any *general* incoherent pattern that these attitudes instantiate: we need to know the specific content

of the attitudes in question to see the relevant "clash." This is no accident: the substantive notion of (in)coherence makes (in)coherence a matter of the substantive support relations between the *contents* of attitudes, so it is unsurprising that we need to know the content of the attitudes to see whether these relations hold. Second, it seems that the clash doesn't obtain for *all* possible circumstances (or evidential situations) that the agent might find themselves in. For example, if there's a large tub of mud sitting inside, and the agent knows that the cat has been playing in it, there is no clash at all between believing the cat has muddy paws and believing the cat has not been outside.

While the cat example in particular probably won't be a site of great controversy, it is worth noting that given a substantive notion of coherence, which states are coherent or incoherent will turn on substantive normative questions about what really evidentially supports what that might in certain cases be very controversial. For some, this might seem to stretch whether the substantive notion is really a notion of *coherence*, properly speaking, at all.

Putting the two contrasts together, we have four possible notions of coherence and incoherence (no doubt further variations are possible within each box):

a) Attitudinal & formal	**b)** Attitudinal & substantive
c) Broadly-encompassing & formal	**d)** Broadly-encompassing & substantive

While all four notions of coherence are intelligibly definable, it's at least *prima facie* unclear what use notion (c) would have. This is because it is doubtful whether non-attitudinal mental states ever enter into *formal* relations of incoherence. One possible candidate might be {having an experience as of *p*, believing not-*p*}, which involves one attitudinal mental state and one non-attitudinal one. But it's clear that attitudes of this pattern do not *always* clash, for all possible contents and circumstances: consider, for example, {having an experience as of the spoon in the glass of water being bent, believing that the spoon in the glass of water is not bent}. So this pattern lacks at least one typical feature of formal incoherence.

With this taxonomy in hand, we can now turn to a few of the different philosophical contexts in which the notion of coherence is invoked, and see which notion of coherence is most plausibly in play.

2. Contexts in which the notion of coherence is invoked

(i) *The foundationalism-coherentism debate*

There is a longstanding debate in epistemology between foundationalists and coherentists about justification [cross-reference here]. We can ask what notion of coherence is at work in the view known as coherentism.

Roughly, coherentism is the view that what it is for beliefs to be justified is for them to be coherent. First: is the notion of coherence here formal or substantive? If it were purely formal, coherentism about justification would be very implausible. To return to our earlier example, suppose that you believe (and suppose the belief is true, if you like, though this shouldn't matter for a coherentist about justification) that the cat has muddy paws (and that there is nothing unusual going

on like a tub of mud sitting inside). A coherentist about justification would need to say that in such a situation, the belief that the cat has not been outside is *no less justified* than the belief that the cat has been outside. After all, neither belief enters into any *formal* relation of incoherence with the belief that the cat has muddy paws (or with any of your other mental states). This seems a very difficult result to accept (pace, perhaps, some subjective Bayesians). By contrast, if the coherentist about justification employs a substantive notion of coherence, they can say, far more plausibly, that the former belief is less justified than the latter. And indeed, coherentists about justification have generally wanted to say this. This strongly suggests that they are employing a substantive notion of coherence.

Now let's ask whether coherentists about justification are employing an attitudinal or an broadly-encompassing notion of coherence. If they were employing a broadly-encompassing notion, it would become hard to distinguish coherentists from their foundationalist opponents (cf. Berker 2015: 333-334). Foundationalists believe that whether your beliefs are justified depends on whether they are (ultimately) supported by some privileged evidential foundation that enjoys a special epistemic status that not all your beliefs enjoy. One popular candidate for this privileged evidential foundation, or a part of it, is your perceptual experiences. But if we allow perceptual experiences to enter into coherence relations, as an broadly-encompassing notion of coherence does (and employ a substantive notion of coherence, as I've just suggested the coherentist about justification should), then for your belief to be supported by your perceptual experiences just is for it to cohere with certain other mental states of yours (viz. those perceptual experiences). Thus, the contrast between foundationalism and coherentism seems to disappear.

Some (e.g. Haack 1993) might welcome this result. Nevertheless, it suggests that those coherentists who *do* want to distinguish their view from foundationalism need to employ a purely attitudinal notion of coherence. Again, the work of coherentists about justification bears this out. The view is often put in terms of saying that for a belief to be justified is for it to (substantively) cohere *with one's other beliefs*, where no beliefs enjoy any privileged foundational status (cf. Harman 2003). This strongly suggests an attitudinal notion of coherence. I thus tentatively conclude that (most) coherentists about justification should be understood as employing notion (b) from our taxonomy above – attitudinal, but substantive.

(ii)   Structural rationality

A second, quite different, context in which the notion of coherence is frequently invoked is in theorizing about structural rationality. According to numerous recent theorists (Fogal 2020; Worsnip 2021), rationality has two dimensions, substantive and structural. Whereas substantive rationality (roughly) consists in being *reasonable* or responding correctly to one's reasons, structural rationality consists in being coherent. A common idea is that there are *requirements* of structural rationality (sometimes called "coherence requirements"), which prohibit incoherent combinations of mental states. Such requirements are usually (though not universally) understood as being "wide-scope," where this means that one can come to comply with the requirement in question by revising any one of the jointly incoherent states: although the requirements of *substantive* rationality might have something to say about which one of them should be revised, the requirements of structural rationality

are themselves neutral about this (Greenspan 1975, Darwall 1983, Broome 1999). This makes sense if we think of these requirements as requirements *to avoid incoherence*, since when two states are jointly incoherent, revising either will get one out of the incoherence in question. Once again, we can ask what notion of (in)coherence is at play: both whether it is attitudinal or broadly-encompassing, and whether it is formal or substantive.

If we construe structural rationality as requiring coherence in an broadly-encompassing and substantive sense, the contrast with substantive rationality arguably disappears. Recall that substantive rationality (roughly) consists in responding correctly to one's reasons. Now suppose that we hold that coherence requires having the beliefs that stand in a substantive relation of support with all of one's mental states, including states like perceptual experiences. Saying that one's perceptual experiences *support* these beliefs plausibly *just is* to say that they constitute or supply *reasons* for these beliefs; consequently, the requirement to be "coherent" in this sense seems to just amount to a requirement to respond correctly to (a particular class of) one's reasons.

As with the foundationalism-coherentism debate, some (e.g. Wedgwood 2017: 11-12) would welcome the result that substantive and structural rationality – coherence and reasons-responsiveness – come to the same thing. However, the fact that coherence *can* be defined in a way that makes being coherent effectively the same thing as being reasons-responsive doesn't show that it can't also be defined in a way that draws an important contrast and isolates a phenomenon of distinctive interest. Particularly if the theorist of structural rationality is not claiming that structural rationality *exhausts* rationality as a whole, there is no pressure on them to construe structural rationality as requiring coherence in the broadest and most demanding sense. Given this, it is legitimate for the theorist of structural rationality to stipulate that the notion of coherence they intend when they say that structural rationality requires coherence is more circumscribed, such that the contrast between substantive and structural rationality does not disappear. To do this, they need to say that the notion of coherence in play for structural rationality is attitudinal and/or formal.

Consider first saying that it is attitudinal, while still maintaining that it is substantive rather than formal. Would this also collapse the distinction between substantive and structural rationality? The answer is that it would do so only if we are coherentists (in the sense at play in the previous subsection) about substantive rationality. If we are coherentists about substantive rationality, then we think that substantive rationality itself consists in being coherent in the attitudinal, substantive sense, and so understanding structural rationality this way too would dissolve the contrast. However, if we are *not* coherentists about substantive rationality, then a notion of structural rationality as concerned with coherence in an attitudinal, substantive sense would still make it distinct from substantive rationality. In particular, the idea could be that whereas substantive rationality requires having the attitudes that stand in (substantive) support relations with our *reasons*, structural rationality requires having the attitudes that stand in (substantive) support relations with our other *attitudes*, without regard to whether those attitudes are themselves (substantively) rational (cf. Fogal 2020; Sylvan 2021). Understood this way, structural rationality is something like *hypothetical* substantive rationality: it requires the attitudes that *would* be substantively rational if one's other attitudes were themselves substantively rational (again, see Fogal 2020; also Pryor 2004: 363-5).

While this is a possible way of understanding structural rationality, the much more common way of understanding structural rationality is as both attitudinal and formal. The vast bulk of the literature focuses on precisely the sorts of forms of (in)coherence that I mentioned above in discussing what it is for (in)coherence to be formal: examples such as {believing p, believing not-p}, or {having credence 0.7 in p, having credence 0.4 in not-p}. (See, e.g., Broome (1999, 2013), Kolodny (2007, 2008), and many others that followed them.) This way of understanding structural rationality makes the contrast with substantive rationality much sharper. We can now distinguish structural rationality from substantive rationality with reference to precisely the properties that are unique to *formal* coherence: that we can make judgments of structural rationality while (at least partly) abstracting away from the contents of our attitudes, and that the irrationality of structurally irrational combinations of attitudes holds regardless of the agent's circumstances or evidence. If we understand structural rationality as concerned with coherence in a substantive (albeit purely attitudinal) sense, it doesn't have these properties. Understanding the kind of coherence involved in structural rationality as formal also latches onto what I think is an intuitive and important sense of 'coherent' and 'rational'. Suppose someone fails to believe what their other beliefs support believing, but this is because they hold a mistaken (but intelligible) view about what their beliefs support believing: they are mistaken about the evidential support relations between their beliefs. I think there is a very natural sense of 'coherent' in which such a person is perfectly coherent, and indeed a *sense* in which they are rational. If we understand structural rationality as concerned only with formal (and attitudinal) coherence, it picks this sense out.

(iii)   Probabilistic (in)coherence

A third context in which talk of 'coherence' comes up frequently is in the formal epistemology literature, where it is often used relatively narrowly to refer solely to *probabilistic* incoherence. This kind of incoherence obtains when one has credences that violate the probability axioms, as in the aforementioned example of {having credence 0.7 in p, having credence 0.4 in not-p} (or, more generally, {having credence N in p, having credence M in not-p} such that N+M $\neq$ 1), or – to use a different sort of example – {having credence 0.7 in p, having credence 0.6 in $p \lor q$} (or, more generally, {having credence N in p, having credence M in $p \lor q$}, such that N>M.

We can be fairly quick here: probabilistic incoherence of credence seems to be a *particular kind of* incoherence in the formal, attitudinal sense – and hence, given the last section, a particular kind of structural irrationality – but not the only kind. The only caveat here is that there are some people (e.g. Foley 1992; Caie 2013; Pryor 2018: 133) who argue that probabilistic incoherence is not, in fact, always irrational. If their case is compelling, we might by the same token want to say that it doesn't count as a genuine kind of incoherence after all, at least if we are using 'incoherent' to refer to the attitudes that *genuinely* clash in such a way as to be structurally irrational. This highlights a more general terminological difference: while some may use 'incoherent' stipulatively to refer to certain kinds of combinations of states (e.g., credences that violate the probability axioms) *whether or not* they are irrational, others take it to already be a normatively or evaluatively loaded term such that to call something incoherent is already to call it irrational in a certain sense (viz., the structural one).

(iv)   (In)consistency

Finally, it's worth briefly remarking on the relationship between the notion of (in)coherence and that of (in)consistency. In its strictest sense, 'inconsistent' refers to propositions, or beliefs in such propositions that are *logically* inconsistent: as a matter of deductive logic, they cannot all be true. (*Contradictory* propositions or beliefs are a still narrower class, consisting of pairs of propositions, or beliefs in propositions, whereby one proposition is the negation of another.) As should be evident from the foregoing, not all incoherent combinations of states are logically inconsistent. Nor, plausibly, are all incoherent combinations of *beliefs* logically inconsistent: many think that {believing *p*, believing that one's evidence does not sufficiently support believing *p*} is incoherent, for example, but there is no logical inconsistency here. Whether the converse is true – that is, whether inconsistent beliefs are always incoherent – depends, again, on whether we are using the term 'incoherent' in such a way that states only count as incoherent if they are genuinely structurally irrational, and on whether we think that inconsistent beliefs are always structurally irrational – the latter of which is, again, a matter of controversy (Harman 1986, Ryan 1991, Kaplan 1996: ch. 4, Christensen 2004: chs. 3-4).

Conclusion: the unity of (formal, attitudinal) coherence?

This entry has tried to make clear some of the different uses to which the terms 'coherence' and 'incoherence' can be put, and which of these uses is at issue in some of the contexts in which these concepts are deployed. There are many important questions about coherence that I am not able to address here. For example, there is a large debate about whether coherence is "normative" in the sense that we necessarily have reasons to be coherent (Kolodny 2007, 2008; Broome 2013: ch. 11; and many others). Another important question is whether, for any of the different concepts that 'coherence' can pick out, there is some general, informative account of what different coherent (or incoherent) combinations of states have in common. In particular, we might wonder whether there is any general feature that all instances of formal incoherence between attitudes have in common.

This is a relatively nascent area of debate, but philosophers have started to offer some possible answers, which I will briefly mention in closing. One view is that incoherent sets of attitudes are all such that it is impossible (no matter the specific circumstances) for *all* of the states to have some other property, such as being sufficiently supported by one's reasons—i.e., substantively rational (Kiesewetter 2017: esp. 235–39; see also Kolodny 2007, 2008; Lord 2018) being "correct" (Singh ms.), to "succeed" (Gibbard 2003; Fink forthcoming), or to be "satisfied" (Fullhart & Martinez forthcoming). Another is that incoherent attitudes are all *exploitable* via devices such as Dutch books or money pumps. A third is that incoherent combinations of attitudes are unified by the fact that (very roughly speaking) it is in some way difficult to sustain them in full, reflective self-knowledge (Worsnip 2021: ch. 5); cf. also Lee 2022). All of these views face challenges and potential counterexamples, but it is early days for this important debate, and we can expect increasingly sophisticated proposals to emerge over the coming decades.[1]

---

[1] Thanks to Daniel Fogal and Kurt Sylvan for helpful comments on an earlier draft.

References

Berker, S. (2015). "Coherentism via Graphs," *Philosophical Issues*, 25(1): 322-352.

BonJour, L. (1985). *The Structure of Empirical Knowledge*. Harvard University Press.

Broome, J. (1999). "Normative Requirements," *Ratio*, 12: 398-419.

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

Caie, M. (2013). "Rational Probabilistic Incoherence," *Philosophical Review*, 122(4): 527-575.

Christensen, D. (2004). *Putting Logic in its Place*. Oxford University Press.

Darwall, S. (1983). *Impartial Reason*. Cornell University Press.

Fink, J. (forthcoming). "The Essence of Structural Irrationality: The Impossibility of Attitudinal Success," *Journal of Ethics and Social Philosophy*.

Fogal, D. (2020). "Rational Requirements and the Primacy of Pressure," *Mind*, 129/516: 1033-1070.

Fogal, D. (2024). "When Things Fail to Fit Together," *Analysis*, doi: 10.1093/analys/anae016.

Foley, R. (1992). "Being Knowingly Incoherent," *Noûs*, 26(2), 181-203.

Fullhart, S. & Martinez, C. (forthcoming). "Coherence as Joint Satisfiability," *Australasian Journal of Philosophy*.

Gibbard, A. (2003). *Thinking How to Live*. Harvard University Press.

Greenspan, P.S. (1975). "Conditional Oughts and Hypothetical Imperatives," *Journal of Philosophy*, 72(10): 259-276.

Haack, S. (1993). "Double-Aspect Foundherentism: A New Theory of Empirical Justification," *Philosophy and Phenomenological Research*, 53(1): 113-128.

Harman, G. (1973). *Thought*. Princeton University Press.

-------- (1986). *Change in View*. MIT Press.

-------- (2003). "Skepticism and Foundations," in Luper (ed.), *The Skeptics: Contemporary Essays*. Routledge.

Kaplan, M. (1996). *Decision Theory as Philosophy*. Cambridge University Press.

Kiesewetter, B. (2017). *The Normativity of Rationality*. Oxford University Press.

Kolodny, N. (2007). "How Does Coherence Matter?," *Proceedings of the Aristotelian Society*, 107(3): 229-263.

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

Lee, W. (2022). "The Real Myth of Coherence", *Erkenntnis*, 87(3): 1211–1230.

Lord, E. (2018). *The Importance of Being Rational*. Oxford University Press.

Pryor, J. (2004). "What's Wrong with Moore's Argument?," *Philosophical Issues*, 14: 349-378.

-------- (2018). "The Merits of Incoherence," *Analytic Philosophy*, 59/1: 112-141.

Ryan, S. (1991). "The Preface Paradox," *Philosophical Studies*, 64(3): 293-307.

Singh, K. (ms.). "Rationality Reunified."

Sylvan, K. (2021). "Respect and the Reality of Apparent Reasons," *Philosophical Studies*, 178(10): 3129-3156.

Wedgwood, R. (2017). *The Value of Rationality*. Oxford University Press.

Worsnip, A. (2021). *Fitting Things Together: Coherence and the Demands of Structural Rationality*. Oxford University Press.