Alex Worsnip, *Fitting Things Together*.
New York: Oxford University Press, 2021, 335 pp.

Fitting Things Together defends the distinctive normativity of *structural rationality*, which requires one's mental states to fit together correctly. More specifically, Worsnip argues that structural rationality is '[a] genuine, [b] autonomous, [c] unified, and [d] normatively significant' (x). (a) and (b) are part of his case for *rationality dualism*, 'the view that structural and substantive rationality are two distinct but equally genuine kinds of rationality, neither…reducible to the other' (4). His account of (c) establishes a link between structural rationality and the metaphysics of attitudes: structurally irrational combinations of attitudes are ones that agents are disposed, owing to the nature of the attitudes at issue, to revise under 'full transparency' (133). His take on (d) meshes with this proposal: structural rationality matters because it is fitting to structure deliberation in ways that treat incoherent combinations as 'off-limits' (256).

The book has three parts. The first supports rationality dualism. The second develops Worsnip's conception of the unity and normativity of structural rationality. The third explores the implications for topics in epistemology and metaethics. I'll focus on the first two parts, but mention aspects of the third in passing. Research on many of the topics Worsnip discusses will benefit from attention to his views. But the traffic goes both ways: some predictions of his outlook for these topics may provide reasons to rethink it.

Let's examine Worsnip's case for bifurcating rationality. He isolates structural rationality by noting, as is standard, that *requirements of coherence* seem to explain the irrationality in some examples. The examples Worsnip highlights are mainstream: means-end incoherence, inconsistency, akrasia, and cyclical preferences. These examples highlight the negative, peremptory side of structural rationality, predisposing readers to understand it as the absence of incoherence, as his 'wide-scope in spirit' view (ch.6) holds.

From such examples, Worsnip infers some hallmarks of structural rationality (7):

- It is circumstance-independent and evidence-independent.

- It is 'formal': structural irrationality violates a pattern that abstracts from content.

- Structural rationality judgments 'can be made in abstraction from disagreements about what is worth doing'.

He notes (17-19) that these hallmarks make structural rationality narrow. He narrows it further when separating requirements of structural rationality from *norms of correct reasoning* (180) and *apparent reasons* (35).

Although Worsnip is mostly making explicit a conception Scanlon (1998) and Broome (2013) fostered, it is worth considering whether one can broaden the diet of examples without changing the subject. As he partially acknowledges (18-19), other rationality evaluations have his hallmarks but suggest a notion of 'fitting-togetherness' that is not just the absence of synchronic incoherence. Evaluations of reasoning *qua* reasoning are good examples. Michael Smith has also long recommended *non-arbitrariness* and *systematicity* as demands of structural rationality, which he invokes in a Kantian analysis of reasons. While Smith's program may fail, it draws attention to examples that could be treated as paradigmatic, and revives rationalist themes that merit attention. While I don't think we are rationally required to believe the Principle of Sufficient Reason, there may be

related structural demands. Consider the puzzlingness of believing p, believing there is objective reason to believe p, believing one doesn't understand why there is objective reason to believe p, but lacking any inclination to understand why. Such cases suggest an *aspirational* side to structural rationality. While it is only a rational imperfection to believe without understanding, this suggests an imperfect duty to seek understanding if one finds it lacking.

A different demand that deserves more attention is a principle of sufficient *apparent* reason: rationality demands φ-ing only if it *appears* to you that there are objective reasons to φ. Worsnip (ch.2) thinks this demand is either evidence-dependent or reduces to anti-akrasia. But while structural rationality is independent of *a posteriori* evidential relations, it may not be independent of the *a priori* demand to respect appearances. Following Kant, one might think this demand reflects a structural requirement to harmonize one's receptive and spontaneous mental capacities. This explains why it is hard to imagine subjects whose attitudes float free from the way the world appears to them: it is better to suppose that the world looks different, say, if one has schizophrenia.

Appearances aren't well treated by the categories Worsnip uses to frame substantive rationality in ch.2. He grounds substantive rationality in 'evidence-relative reasons', which he contrasts with 'fact-relative' and 'belief-relative' reasons. Appearances are not fact-relative or belief-relative reasons. They aren't evidence-relative reasons either. These are more demanding: as Worsnip's discussion of higher-order evidence in ch.3 suggests, one can have conclusive evidence for p even though it doesn't appear from one's overall perspective that p (and wouldn't on reflection). Holding that structural rationality is coherence with one's overall perspective, including appearances, seems more fitting.

There are other cases where Worsnip's divisions feel ill-fitting. In ch.6, Worsnip maintains that structural rationality is synchronic and not concerned *per se* with the form of one's *processing*. These claims oppose Kolodny's idea that rational requirements concern the structure of one's reasoning. Worsnip instead suggests (180, n.31) a trinitarian view with a distinctive category of norms of reasoning. This feels artificial: it is more natural to start by treating fallacious reasoning as a paradigm of deficient structural rationality, since it displays the hallmarks.

Let's turn to Worsnip's case against rationality monism. Worsnip argues first (ch.3) against substantivist views that reduce structural rationality to substantive rationality or eliminate it, and then (ch.4) against structuralist views that reduce substantive rationality to structural rationality. Focusing on Kiesewetter and Lord, Worsnip assumes substantivists need a *Guarantee Hypothesis*, on which structural irrationality guarantees the presence of some substantively irrational attitude(s). He then makes two moves. Firstly, he argues that even if the Guarantee Hypothesis is true, there are decisive objections to substantivism. Secondly, he gives counterexamples to the Guarantee Hypothesis and argues that substantivists cannot explain them away.

Worsnip's case against substantivist eliminativism appeals to two intuitions. One is the 'counting intuition': incoherent subjects with substantively irrational attitudes make two kinds of mistakes. Another is the intuition that coherent subjects are more rational in one respect than incoherent subjects. Worsnip's case against Lord's and Kiesewetter's diagnoses of these intuitions may succeed. But he neglects the diagnosis of Kolodny (2005)'s 'Transparency Account'. He considers a later statement of Kolodny's view, taking Kolodny to agree that there is one genuine structural requirement—anti-akrasia. But this ignores a key theme in Kolodny (2005: 509): '[t]he normative 'pressure' that we feel, when rational requirements apply to us, derives from…the reasons that, as it seems to us, we have'. Kolodny (2005: 558) stresses that '[t]his account does not appeal to…an additional normative concept beyond that of a reason.'

Friends of this approach can agree that there is a distinctive *non-normative* property of coherence, but hold that it has no essential significance. Preface cases recommend this view: here it seems *more rational* to become inconsistent by humbly allowing that one surely made some mistake, but couldn't find it. This point reveals a problem for Worsnip. His best counterexamples to the Guarantee Hypothesis—preface cases—undermine the intuitions he wields against opponents. Preface cases make it plausible that incoherence is insufficient for irrationality: what could be a clearer example of incoherence, after all, than having beliefs that straightforwardly entail a contradiction? Such cases recommend the following re-diagnosis: incoherence ensures a mistake of rationality *only when,* by being incoherent, one *thereby* invites substantive irrationality.

Let's consider structuralism. Worsnip starts with arguments against structuralist eliminativism, then attacks Humean and Kantian reductive views. He suggests (95-99) that Broome's eliminativism assumes an overly narrow conception of rationality that ignores the rational significance of evidence-relative reasons. But defenders of a similar view could respond that rationality supervenes on apparent reasons, which aren't relative to *evidence* in the natural sense, but rather *appearances*, where the requirement to respect appearances reflects a structural demand to harmonize two mental capacities.

Worsnip's case against reductionism doesn't do it full justice. He assumes the best views are 'counterfactual and idealizing', invoking 'the attitudes that an idealized counterpart of us would have, where the relevant idealization…mak[es] the counterpart structurally rational' (99). Some of Worsnip's objections are compelling: counterfactual views suffer from indeterminacy, and Humean views fail to explain moral reasons. He doesn't, however, give a decisive case against Kantianism.

Worsnip believes that Kantians in the end must appeal to substantive rationality. But Kant's view plausibly rests on a regress argument that makes a point about the structure of practical reasoning. It is only intelligible to pursue the means to one's ends if it makes sense to have them in the first place. How then does practical reasoning get started? Kantians deny that it can intelligibly begin with externally dictated ends: this is *arbitrary* and *heteronomous*. One can terminate the regress of practical reasoning autonomously only if there are ends that one can necessarily will autonomously. There will be, if some ends are *constitutive* of practical reason. Here Kantians use an analogue of the *cogito* argument: some ends cannot intelligibly be questioned if one is to have any end at all. Chief among them is reason's capacity to set laws for itself: willing presupposes the authority of this capacity.

Discussing Markovits's relative of this argument, Worsnip reasonably suggests that the claim that *humanity* is a special end is 'a substantive judgment' (113), but no such judgment features in the version above. He may insist I appealed to a substantive judgment in holding that it is arbitrary and heteronomous to submit to externally dictated ends, but this is less obvious. Anti-arbitrariness and anti-heteronomy can be framed as structural requirements. Kantians could then propose that it is substantively rational to φ iff there is a well-founded pattern of reasoning that begins with the ends that are constitutive of practical reason and concludes with φ-ing. One might doubt this view ensures the substantive rationality of morality. But it shouldn't therefore be excluded. Worsnip's evidence-relative view doesn't do so either, if developed plausibly.

There are epistemic analogues of this story. Kantians in epistemology defend *a priori* canons of theoretical reasoning via transcendental arguments that mirror the practical regress argument. Worsnip doesn't consider such views. The coherentist views he considers (116-118) are better cast in this spirit. Rather than holding that 'a belief of yours is supported by evidence iff it coheres with your other beliefs', one could say that a belief is supported by E iff it *makes sense* of E, where there are *a priori* principles of sense-making grounded in the constitution of theoretical reason. Such views needn't appeal to 'substantive judgments about

support relations': it is *incoherent* to think in ways that defy the constitution of theoretical reason. While the approach requires a *disputable* view about the nature of theoretical reason, it is not a *substantive* view in the relevant sense: Worsnip's account of structural rationality rests on a similar view.

Let's consider Worsnip's positive views. His account of the unity of structural rationality seeks to explain why it is *puzzling* (129). Focusing on incoherence, he notes that it is *very* puzzling: incoherent tangles of attitudes seem only attributable given failures of transparency, and are liable to unravel once spotted. This suggests a connection between coherence and the metaphysics of attitudes: '[a] set of attitudinal mental states is jointly incoherent iff it is (partially) constitutive of (at least some of) the states in the set that any agent who holds this set of states has a disposition, when conditions of full transparency are met, to revise at least one of those states' (133). Worsnip notes that this proposal gives some support to constitutivism and the idea that norms of rationality are preconditions of interpretation (150ff). But these ideas allegedly *only* work for structural rationality (156).

The basic shape of this view is compelling, but a less restrictive version explains more. More cases are relevantly puzzling. Consider reasoning that radically flouts inductive and abductive canons, gruesome conceptual schemes, and paranoid beliefs. While such errors are conceivable, it seems inconceivable to engage in theoretical reasoning without *commitment* to the ideals that explain them. It is very puzzling if such shortcomings are noted but don't incline one to return to the drawing board.[1] The underlying explanation is plausibly constitutivist: besides being constrained by constitutive *prohibitions* against incoherent combinations, our reasoning is regulated by constitutive *ideals*. These are not *compliance* constitutive (149)—reasoners needn't be disposed to *meet* them—but they are *commitment* constitutive: we must be *guided* by them to count as having rational capacities.

This story helps with the normativity of rationality. Worsnip thinks structural rationality is normative because there are reasons to structure deliberation in ways that treat incoherent combinations as off-limits. But why structure deliberation this way? Worsnip offers a 'rather uninformative' answer: 'because it's fitting to structure deliberation in ways that respect coherence constraints' (265). He offers a little more by noting that his view of incoherence 'drive[s] home just how little sense these combinations of attitudes make' (266). But similar facts make it unfitting to commit to counter-inductive and counter-abductive canons of ampliative inference, invalid deductive patterns, and gruesome conceptualization. Inference and conceptualization are constitutively regulated by ideals. If one asks why they matter, the answer is that we're committed to them in virtue of being rational beings.

This is not Worsnip's proposal: he thinks it is a substantive fact that it is *fitting* to treat structural irrationality as off-limits. But rational beings cannot but see such combinations as off-limits. If this weren't true, I doubt it would be clear that it is necessarily fitting to avoid them, any more than it is to shun odd socks. Indeed, when inconsistency is something we can invite—e.g., in preface cases—it is accepted as an outcome of our crooked timber. To insist that it still is unfitting to tolerate inconsistency here, as Worsnip does (270-1), is to worship the hobgoblin of consistency. While Worsnip admits that it is *permissible* to be incoherent here (303), this sits unevenly with the claim that it is unfitting: better to say that it is fitting to tolerate some inconsistency, given our inescapably crooked frame.

Further explanation of this fittingness is needed. The Kantian option is to say that *global* consistency and fit with appearances are *constitutive ideals* of reasoning, associated with imperfect duties of rationality. There is no conflict of requirements, just two ideals. A wider constitutivist approach, which includes not just compliance-constitutive requirements but commitment-constitutive ideals, allows for this. This approach takes the question of why

---

[1] A similar puzzlingness is noted in Williams (2020) and earlier work on naturalness (Hirsch (1993)).

rationality matters seriously, without supposing that an answer must bottom out in the promotion of value. To decline the question with an appeal to fittingness seems no less 'tyrannical' than the instrumentalist outlook Worsnip rightly rejects in the Coda.

<div style="text-align: right">

Kurt Sylvan
University of Southampton

</div>

References

Broome, J. 2013. *Rationality through Reasoning*. Blackwell.
Hirsch, E. 1993. *Dividing Reality*. OUP.
Kolodny, N. 2005. 'Why Be Rational?' *Mind* 114: 509-563.
Scanlon, T. M. 1998. *What We Owe to Each Other*. Harvard University Press.
Williams, R. 2020. *The Metaphysics of Representation*. OUP.