

Fitting Things Together: Coherence and the demands of structural rationality. By ALEX WORSNIP. (New York: Oxford University Press, 2021. Pp. xiv + 335. Price \$99.00).

Alex Worsnip's recent book, *Fitting Things Together: Coherence and the Demands of Structural Rationality* provides a sustained, wide-ranging defense of dualism about rationality. Worsnip contends that there are two independent kinds of rationality: *substantive rationality* and *structural rationality*. Structural rationality is a property manifest in a variety of mental states in virtue of the relations between those states. Substantive rationality, by contrast, is a distinct property likewise manifest in a variety of mental states in virtue of their relation to the reasons a subject has given her evidence. Worsnip's book is a first-rate contribution to meta-ethics and epistemology. I'll begin by introducing the distinction, turn to a summary of the book, and then raise two concerns about parts of the overall view.

I. Dualism and the case for it

Consider several cases Worsnip offers to begin his defense of dualism (p. 3).

- *Means–end incoherence.* Val intends to get certified as a Green Zone Ally this year. She also believes that, in order to get certified as a Green Zone Ally this year, she has to attend the training on Monday afternoon. But she does not intend to attend the training on Monday afternoon.
- *Inconsistency:* Farhan believes that he is a great cook. He also believes that if someone is a great cook, they never overcook fried eggs. But he believes that he has overcooked the fried eggs.
- *Cyclical preferences.* Between working on his book and volunteering at the homeless shelter, Alex prefers to work on his book. Between volunteering at the homeless shelter and binge-watching

The Sopranos, Alex prefers to volunteer at the homeless shelter. But between binge-watching *The Sopranos* and working on his book, Alex prefers to binge-watch *The Sopranos*.

These individuals, Worsnip affirms, are irrational in a distinctive way. Their irrationality lies in the way “their attitudinal mental states (where this includes both positive attitudes and absences thereof) fail to fit together” (pp. 3-4). The subjects in these cases display *structural irrationality*.

Why think that the irrationality exhibited by these individuals is distinct from the irrationality of failing to respond appropriately to the evidence? Worsnip offers the following argument based on a *counting intuition*. Consider Tom who believes he is Superman and believes that Superman can fly. Tom is irrational in believing that he is Superman because it conflicts with his evidence. Now stipulate that Tom believes that he can't fly. This further belief is overwhelmingly supported by the evidence that Tom possesses. Yet, when we learn that Tom believes that he can't fly, we see that he is subject to a second charge of irrationality, namely that Tom is inconsistent. Worsnip claims that this rational failing is a *distinct* rational failing (p. 5). Thus, we have two distinct failures: a failure to respond correctly to one's evidence and a failure to have consistent beliefs.

Worsnip proceeds to offer three hallmarks of structural rationality (pp. 7-8). First, judgements about structural (ir)rationality can be made with minimal information about the context of the agents. Call this hallmark *contextless rationality*. Second, judgements about structural rationality can be made in isolation about what is worth doing or the nature of reasons. Call this hallmark *neutral rationality*. Third, structural (ir)rationality can be described by formal patterns exhibited by combinations of mental states. Call this hallmark *formal rationality*. Worsnip claims that all three hallmarks are distinct from substantive rationality which require (i) more information about the

context of the subject, (ii) brings in considerations of value and reasons, and (iii) cannot be described in formal patterns.

The initial case for dualism progresses by distinguishing structural rationality from a host of related phenomena (pp. 10-22). Of note, structural rationality is exhibited across beliefs, preferences, intentions, means-end reasoning, inter-level coherence, etc. It is not a narrow phenomenon restricted to inconsistent beliefs or credences that cannot be represented by a probability function. Moreover, structural rationality isn't a broad sense of coherence exhibited by, for example, explanatory relations between one's beliefs and experiences. A subject who visually experiences a red token on a table but believes there is not a red token on the table is not structurally irrational. While the presence of a red token may explain the experience of a red token, there are competing explanations in the offing. Because structural rationality is both contextless and formal, cases like these which require more context and cannot be exhibited by formal patterns, concern substantive rationality.

II. An overview

Worsnip divides his book into two parts: dualism about rationality defended (chapters 1-4) and a theory of structural rationality (chapters 5-9). After the initial case for dualism in chapter 1, Worsnip turns in chapter 2 to offer a rough account of substantive rationality in terms of actions and attitudes that fit the reasons one has in virtue of the evidence one possesses. He uses Bernard Williams's famous petrol case to distinguish between three different ways we can talk about reasons: fact-relative reasons, evidence-relative reasons, and belief-relative reasons. Fact-relative reasons are

reasons that come from the facts whether one is aware of those facts. Belief-relative reasons come from the subject's beliefs. Evidence-relative reasons come from the subject's evidence.

Consider the evidence-relative reading of Williams's petrol case. A subject is looking at a martini glass that is filled with petrol. It looks and smells like petrol. Nonetheless the subject believes that it contains gin. Worsnip avers that the evidence the subject possesses—the look and the smell of petrol—makes it the case that it is substantively irrational to believe that the glass contains gin. To secure this judgement, we need to assume that the subject knows what petrol looks and smells like, and similarly knows that gin has a different look and smell. The upshot is that substantive rationality concerns evidence-relative reasons. Structural rationality, by contrast, does not depend on responding to one's evidence-relative reasons. A consistent conspiracy theorist may not correctly respond to her evidence but may be perfectly structurally rational.

In chapter 3 Worsnip argues against attempts to eliminate structural rationality or reduce it to substantive rationality. His discussion centers around the following thesis.

The Guarantee Hypothesis (GH). For any set of attitudinal mental states $\{A_1 \dots A_n\}$ of the kind associated with structural irrationality, it is guaranteed that at least one of $\{A_1 \dots A_n\}$ is substantively irrational (i.e., that it is insufficiently supported by the agent's evidence-relative reasons). (pp. 54-55)

If GH is true, then the requirements of structural rationality are superfluous. Worsnip engages with two distinct views, both motivated by GH. Eliminativism maintains that structural rationality is not a genuine property; rationality is just substantive rationality. Reductivism states that while structural

rationality is a property, it can be reduced to substantive rationality. I'll focus on Worsnip's arguments against eliminativism.

Against eliminativism, Worsnip argues that (i) it can't account for the counting intuition and (ii) it can't capture the intuition that structurally coherent agents are rational. The remainder of the chapter is a rich discussion of reasons to reject GH. He considers a bevy of cases from metaethics and epistemology—the preface paradox, Bratman's video game case, misleading higher-order evidence, etc. A complete defense of GH would require discussing many of these cases. A strength of Worsnip's discussion is careful assessment of these cases. He acknowledges that the relative strength of the argument against GH substantively differs for each case.

In chapter 4, Worsnip argues against attempts to eliminate substantive rationality or reduce it to structural rationality. An advocate of reducing substantive rationality to structural rationality is John Broome who holds that rationality is coherence, and not responsiveness to reasons (p. 95). But Broome is using the fact-relative notion of reasons, not the more plausible evidence-relative reasons. Worsnip notes that a motivation for Broome's view is the idea that rationality supervenes on the mind. One way to unpack this is with the idea that mental duplicates have the same rationality status. An account that attempts to eliminate or reduce substantive rationality in terms of relations among mental states would be a *monistic structuralist account*.

What are Worsnip's arguments against monist structuralism? As I see it there are two basic moves. First, monist structuralism has some disagreeable implications. A conspiracy theorist with coherent beliefs would be rational. Second, monist structuralism requires a more substantive theory of evidence; one needs to specify what counts as evidence and what are the relevant support relations.

Worsnip's discussion in this chapter focuses largely on positions in practical rationality, but he does argue against epistemic examples of monist structuralism in epistemic coherentism and Bayesianism.

Chapter 5 begins the second part of the book where Worsnip offers his theory of structural rationality. In chapter 5 he offers an account of what unifies instances of incoherence. Earlier he rejected GH which offered such a unifying account. In its place, he offers this:

INCOHERENCE TEST: 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. (p. 133)

The Incoherence Test captures the idea that instances of structural irrationality are puzzling. A subject like Tom, who believes he is Superman and believes Superman can fly but believes that he can't fly, is hard to make sense of. To the extent we can make sense of Tom, it involves either inattention to the relevant beliefs or belief fragmentation. But if made fully aware of his beliefs we'd expect Tom to give up at least one of those beliefs. Subjects that don't revise in conditions of full transparency are more accurately described as having some other set of cognitive states. This is apparent in the case of means-end incoherence. If Val intends to get certified and believes the only way to do so is to attend training but doesn't intend to attend training, then when made fully aware of this, she feels no pressure to revise, then she is more accurately described as having a wish rather than an intention to get certified. Worsnip spends the remainder of the chapter unpacking the incoherence test and engaging with objections to it.

In chapter 6, Worsnip fills out the requirements of structural rationality and the form of such requirements. He argues that the requirements are prohibitions against incoherent sets of attitudes (p. 165). This chapter discusses, among other topics, whether the requirements are wide scope or narrow scope and whether they are synchronic or diachronic. Worsnip defends a wide-scope, synchronic understanding of these requirements.

Chapter 7 discusses a challenge to the wide-scope reading of the requirements of structural rationality. In ordinary language one says things like this: ‘If you believe that all taxes are illegitimate, then you shouldn’t also believe that sales taxes are legitimate’ (p. 196). Here the scope of the normative requirement appears to have narrow scope. Worsnip defends his wide-scope account by defending a contextualist semantics for modals.

Chapter 8 tackles the question of whether structural rationality is normative and if so, how it is normative. Worsnip defends the view that structural rationality is normative in virtue of providing reasons and then he defends a specific account of how coherence provides reasons.

Reasons-to-Structure-Deliberation Model. Considerations of coherence constitute reasons to structure deliberation in certain ways. More specifically: the fact that some possible combination of attitudes is incoherent is a reason to treat it as off-limits in one’s deliberation. (p. 256)

Coherence-based reasons and substantive reasons play different sort roles in the deliberative process. The coherence-based reasons are reasons to structure the deliberative process in a certain way: they bear on which attitudes one should take seriously in deliberation. Conversely, the

substantive reasons enable one to adjudicate the merits of the options that one does take seriously, within the deliberative constraints set by the coherence-based reasons.

In chapter 9, Worsnip draws out some of the implications of his dualist view. He discusses a multitude of topics including moral rationalism, rational choice theory, higher order evidence, and conditionalization. The upshot is that Worsnip's dualism about rationality carries significant implications for a wide variety of topics in meta-ethics and epistemology.

III. Evidentialist apprehensions

My preferred view in epistemology and the theory of rational choice is evidentialism coupled with the injunction to maximize expected utility. What should one believe? One should believe what one's evidence supports. What should one do? One should act to maximize expected utility. How would such a combination of views fit with Worsnip's account? It is a form of monism on which rationality consists in appropriate responses to one's evidence and preferences.

In the following I comment on two key moves in Worsnip's case for dualism and explain how an evidentialist might respond.

A. The counting intuition

As we saw above, Worsnip argues that the case of Tom exhibits a failure to respond appropriately to the evidence and a distinct failure to be coherent. But why think that failures to be coherent are not failures to respond appropriately to one's evidence? Tom has evidence that inconsistent beliefs

cannot all be true. In particular, Tom has evidence that the following set is inconsistent {I am Superman, Superman can fly, I can't fly}. There is a second rational failure in Tom's case, viz., an additional failure to heed his evidence that that set of beliefs cannot all be true.

Worsnip also appeals to cases of positive coherence to argue for dualism. Consider Tim who believes he is Superman, believes Superman can fly, and believes he can fly. Worsnip states that Tim is not subject to the distinct criticism that Tom is, even though like Tom, Tim is not responding appropriately to his evidence. Worsnip's dualist view maintains that Tim is rational in a sense and irrational in a distinct sense. How may the evidentialist respond to this? The evidentialist may hold that Tim's belief that he can fly is an appropriate response to *some* of his evidence. His beliefs that he is Superman and Superman can fly can, if true, explain why he can fly. Moreover, Tim may competently deduce that he can fly from his beliefs that he is Superman and Superman can fly. But Tim also fails to heed *all* his evidence. So, while Tim's set of related Superman beliefs appropriately responds to *some* of his evidence, it does not appropriately respond to *all* his evidence. Thus, the evidentialist can explain the two distinct and conflicting senses that Worsnip highlights in terms of Tim's responses to the evidence.

The full development of the counting intuition lays in Worsnip's articulation of *Reasons-to-Structure-Deliberation Model* (p. 256). Here dualism becomes the view that there are two distinct sets of reasons: reasons to structure deliberation in certain ways and reasons arising from one's evidence. But I found myself wondering how there could be reasons to structure deliberation in certain ways that were independent from one's evidence. Consider the set {p, p implies q, not-q}. One has as evidence that this set cannot be true. As such one has an evidential reason to structure deliberation in ways that avoid having those three beliefs.

One move to make would appeal to fact-relative reasons and argue that there are fact-relative reasons to structure deliberation in certain way. The distinct evidence-relative reasons are reasons to having certain attitudes. But this move would cut against the arguments Worsnip brings to bear on the discussion of Williams's petrol case. I found myself in large agreement with the stress Worsnip gives to the importance of evidence-relative reasons. Given the considerations I've adduced, one might defend a broad form of evidentialism on which all reasons are evidential reasons.

B. Contextless, Neutral Rationality

A second motivation for dualism about rationality comes from Worsnip's hallmarks of structural rationality. Recall that these were the claims that structural rationality is contextless, neutral, and formal. Let's focus on the first two. Worsnip argues that substantive rationality requires a theory about what counts as evidence and what the evidential support relations are. But some ascriptions of irrationality do not seem to depend on such controversial matters. Consider the case of Val, who intends to get Green-Zone Ally certified, believes the only way to do that is to attend the meeting, but doesn't intend to attend the meeting. We don't need much, if any, by way of theory to be puzzled with Val's case.

What should the monist about rationality say about these cases? As stated, Val's case doesn't threaten monism. Evidentialism informed by rational choice theory may hold that Val's act violates the injunction to maximize expected utility. It's natural to fill out Val's case in the following way. Val places very high utility on becoming Green-Zone Ally certified. Val rules out becoming Green-

Zone Ally certified conditional on any act incompatible with attending the Monday meeting. Yet Val does not intend to act in a way that will bring about her best outcome. Hence, Val is irrational.

What we did here is a bit of philosophy that aims to provide an informative and generalizable explanation of Val's irrationality. We began with a clear intuition and then provided a more substantive theory to explain that intuition and apply it to wider range of cases. An evidentialist may see that Val isn't appropriately responding to her evidence. Value judgments reflect evidence that something is valuable. In placing a high-value on Green-Zone Ally training Val has evidence that this is a thing that matters. Moreover, Val has evidence about how to achieve that goal. But she doesn't appropriately respond to that evidence.

Doesn't this shift a contextless, neutral perspective to a perspective that requires more information about Val and more theory? Yes. But this illustrates one of the organizing worries I had about the stress on contextless, neutral rationality; there is no such thing. Consider, for instance, Worsnip's discussion of dialethism (pp. 20-21, 140-143, 190-191). A dialethist, like Graham Priest, holds that there are true contradictions. This involves an openness to accepting instances of p & not p . Yet if anything is structurally irrational, a contradiction is structurally irrational. Worsnip's discussion of dialethism is nuanced, but he ends up defending a view close to the one David Lewis articulated where Lewis maintains that he cannot make sense of the idea that there may be true contradictions.

Being in the company with David Lewis isn't a bad thing at all. As I understand Worsnip's view, it is that there are epistemic requirements on the ascription of certain mental states, in particular those picked out by the incoherence test. One cannot have the belief that p and the belief that not- p without having the disposition to revise one of those beliefs under conditions of full transparency.

But an evidentialist may hold that, given enough changes in the background evidence, it is possible to believe p and believe not-p in conditions of full transparency without any disposition to revise.

Worsnip might respond with a defense of classical logic and its normativity, but that appears to muddy the waters surrounding the idea that structural rationality is contextless and neutral.

These concerns aside Worsnip offers substantive and engaging work in metaethics and epistemology. I anticipate reflection on his work will bring much fruit.

TED POSTON
University of Alabama