

RETHINKING RATIONALITY ATTRIBUTIONS¹

Lisa BASTIAN

ABSTRACT: Although much has been written about the property of rationality, its requirements, and whether it is normative, rationality attributions themselves have not received much attention. The main aim of this paper is to address this oversight by focussing directly on rationality attributions and their complexities. After offering a diagnosis for why attributions have been largely overlooked, the paper introduces three problems that have plagued the rationality debate as a result: implausible symmetry, conflicts within rationality, and with reasons. Brunero's (2012) answer to the symmetry problem provides the beginnings of a solution but makes it harder to compare the rationality of agents. The remainder of the paper is dedicated to spelling out the new approach: to understand rationality attributions as relative to specific sets of attitudes. This approach circumvents the previous problems and meets a number of desiderata, resulting in a more nuanced and complete understanding of rationality that can enhance our practices of praising or criticising agents based on their rational performance.

KEYWORDS: rationality, rationality attributions, symmetry, requirements

Introduction

If a friend approaches a big life decision in a level-headed way rather than relying on gut feeling, we are likely to congratulate them for being so rational. And if we are trying to poke holes in the arguments of a political opponent, we readily point to any signs of irrationality. When we judge our friend as rational, and our political opponent as irrational, we use these assessments of their rationality as a basis for praising or criticising them. This illustrates that whether we can *attribute (ir-)rationality* to someone plays an important role in our evaluations of agents.

Attributions of rationality or irrationality, and in particular attributions of coherence (or structural) rationality are the focus of this paper. Much has been written about the property of rationality, its requirements, and whether it is normative. But rationality *attributions* themselves have not received a lot of attention, which leads to a number of problems, as shown in section 2. In this paper,

¹ I would like to thank Justin Snedegar, Miguel Egler, Giorgio Sbardolini and my colleagues at VU Amsterdam for providing (sometimes multiple rounds of) comments and for discussing this paper with me.

I aim to address this oversight. By focussing directly on rationality attributions, we can avoid these problems and gain a more nuanced and adequate picture of rationality. Before turning to the attributions themselves, let me first offer a quick diagnosis for why they have been overlooked so far.

1.1 The Requirements-Based Account

On a structural, or coherence view of rationality, being rational amounts to having one's mind in proper order.² In particular, it requires agents to display certain patterns in their propositional attitudes, such as intending the means to their ends, or not believing contradictions.³ As Fogal (2020) helpfully observes, defenders of this view often also subscribe to what he calls the "requirements-based account" of rationality, according to which rationality is "associated with a distinctive set of *rules* or *requirements*" (Fogal 2020, 1035, emphasis in original).⁴ These rules or requirements are taken to proscribe or prohibit coherent or incoherent patterns of attitudes. This account is most explicitly endorsed by John Broome, who claims that "a large part of rationality consists in conforming to normative requirements" (1999, 410), and that "[p]erhaps the most important question a system of rational requirements needs to settle is whether you are rational – have the property of rationality" (2007a, 462), but it also informs the works of Way (2018) and Schroeder (2013), among others. Acceptance of the requirements-based view explains why vast amounts of literature are dedicated to finding the correct formulation of rational requirements – since rational requirements are crucial for rationality, it matters that we get them right. To put the point in Fogal's words, acceptance of the requirements-based account explains why "the contemporary debate over the nature of structural rationality has proceeded on the largely unquestioned assumption that it bottoms out in requirements, rather than something else" (Fogal 2020, 4).⁵

² For some expressions of the coherence view, see Broome (2013), (2005), Kolodny (2005), Bratman (2009), Raz (2005), and Southwood (2008). The coherence view of rationality is rivalled by substantive views of rationality, also known as the reasons view of rationality. Arguments in favour of the reasons view can be found in Nozick (1993), Gibbard (1990), Lord (2018), and Kiesewetter (2017). This view holds that rationality is a matter of correctly responding to reasons. In this paper, I aim to shed light on rationality ascriptions on a coherence understanding. I do not argue here that there could not also be reasons requirements of rationality.

³ Going forward, I mostly refer to 'propositional attitudes' as simply 'attitudes'.

⁴ Fogal rejects this account, but for reasons that do not bear on the discussion at hand, and are not incompatible with the conclusions of this paper.

⁵ As an example for this strong focus on rational requirements, take the vast literature on the question whether rational requirements take wide or narrow scope, starting with Kolodny's (2007) and Broome's (2007c) exchange. Kiesewetter (2017) provides a helpful survey of the broader

If we accept the requirements-based account, the question of whether someone is rational boils down to the question of whether they satisfy or violate the rational requirements that apply to them. We only need to check whether an agent complies with the relevant rational requirement(s), assuming that we have decided on their correct formulation. If agents violate rational requirements, then this “guarantee[s] a negative evaluative property—namely, being irrational”; and if they satisfy such requirements, they are rational in virtue of their compliance (Fogal 2020, 6–7). This has led people to mostly ignore some of the complexities in rationality attributions, leading to the issues discussed in section 2. Looking at a few examples quickly shows how the requirements-based account is implicitly at work in this way.

Take Kolodny, who quickly links an attribution of irrationality to the (potential) violation of a coherence requirement (a requirement to believe the consequences of one’s beliefs):

Look: I think you’re nuts to believe in God, let alone that He created anyone. But given that you believe that God created all people equal, and given that you agree that people whose skin is a different colour from yours are people, you ought to believe that He created them equal too. *It would be irrational of you not to.* (Kolodny 2005, 554, emphasis added)

We can make similar observations when it comes to means-end coherence. Authors directly move to ascribing irrationality based on looming or occurring violations of means-end coherence requirements:

And if he wills to get a philosophy degree and believes that logic is required to get the degree (even though it is not), *he is irrational* if he fails to sign up for logic. (Bedke 2009, 674, emphasis added)

Candice decides to go to the post office this afternoon to send out some mailings, but on the way there, she gives up on this end and decides to go buy groceries instead. But on the way to the market, she yet again trades in this end for another: going to hang out with her friend David. But on the way to David’s house, she once more changes her mind and intends to spend a relaxing afternoon at home, but by the time she gets home the afternoon is gone and she’s accomplished nothing. We’re inclined to accuse Candice of some kind of irrationality. And, since Candice failed to follow through on any of her ends, we’re inclined to *accuse her of being instrumentally irrational*. (Brunero 2012, 236, emphasis added)

Correspondingly, if agents satisfy a means-end coherence requirement, they are attributed rationality:

You could intend to do E and fail to do M, and still *be entirely rational*, even if M is the means to E. For one thing, you might not believe M is the means to E, and

debate.

have no evidence that it is. (Broome 2005, 4, emphasis added)⁶

1.2 The Simplistic Approach

In all of these examples, violation or satisfaction of a coherence requirement quickly led to the attribution of rationality or irrationality to the agent. This is what I call the *Simplistic Approach*: agents are attributed (ir-)rationality depending on whether they violate or satisfy a rational requirement. The Simplistic Approach makes sense if we endorse the requirements-based approach, which sees compliance with rational requirements as crucial for being rational. But it also provides fertile ground for a number of problems that have plagued the rationality debate, such as symmetry (2.1), conflicts within rationality (2.2) and with reasons (2.3). Traditionally, with the requirements-based approach in the background, we have looked to the rational requirements themselves to solve these problems, the thought being that any issues surrounding rationality can be solved if we find the correct requirements in the correct formulation.

In this paper, I want to shift the focus away from rational requirements and to the attributions themselves. My proposal moves away from the Simplistic Approach and introduces some additional complexity to rationality attributions. It underscores the insufficiency of taking the requirements-based account and the Simplistic Approach as providing a complete picture of rationality attributions. My approach sheds light on the complexities of how rational requirements connect to rationality attributions and as a result avoids the problems stated above. This does not mean that discussions about rationality's requirements are fruitless, or that the requirements-based account is wrong.⁷ But by focussing on attributions, we gain a more complete picture of rationality as a whole. My favoured approach does more justice to the nature of coherence rationality and also allows for progress regarding some of the existing problems. And since our practises of praising or criticising agents are closely tied up with our judgments of a person's rationality, it is especially important that we are aware of the pitfalls surrounding rationality attributions.

Here is the plan of the paper: in the next section, I introduce three issues that arise from the Simplistic Approach: inability to account for symmetry, conflicting

⁶ In this case, the agent lacks the belief that M is a means to E, which features in the antecedent of the material conditional of a wide scope means-end coherence requirement, making the antecedent false and the material conditional true.

⁷ It is not my aim to reject the requirements-based account. As we will see, rational requirements play an important role in determining whether someone is rational on my approach too. What I want to draw attention to is that the requirements-based account alone does not have the resources to address the complexities surrounding rationality attributions.

rationality attributions, and conflicts with reasons (section 2). I then turn to sketching the beginnings of a solution, using Brunero's (2012) answer to the symmetry problem as a starting point for deriving a *Refined Simplistic Approach*. But as we will see, this approach faces problems of its own, related to our ability to compare the rationality of agents (section 3). Section 4 translates the findings of previous sections into desiderata for a better approach to rationality attributions. The remainder of the section is spent spelling out my favoured approach. The paper concludes that my proposal of understanding rationality attributions as relative to specific sets of attitudes has the ability of meeting the established desiderata and avoids the issues raised in sections 2 and 3.

2. Three Issues

I want to focus on three issues in particular: symmetry, conflicting attributions, and conflicts with reasons. They can all be traced back to the Simplistic Approach of directly deriving attributions of rationality or irrationality from the satisfaction or violation of a rational requirement. This is not an exhaustive list, but the selection allows us to get a clearer idea of how a better approach for rationality attributions should look like.

2.1 Symmetry

Let's turn to *symmetry* first. The symmetry problem takes issue with the apparent but implausible symmetry between different ways of satisfying requirements of structural rationality. For example, means-end coherence requirements typically require agents to avoid being in a situation where they do not intend the means they believe necessary for their ends. This can be done in a number of ways: by intending the means, but also by not believing it necessary, and by not having the end.⁸ All of these are symmetrical in the sense that they all lead to the satisfaction of a rational requirement. But many have noted that they are importantly different: some seem like better ways of satisfying the requirement than others; and some seem outright irrational. As Brunero (2012, 126) puts it, "there appears to be [...] an asymmetry in the rationality of the different ways one can escape from a state of means-ends incoherence".

To illustrate, consider an example:

⁸ This is the case on wide scope formulations of rational requirements, on which the relevant operator governs the entire material conditional. Narrow scope formulations are not subject to the symmetry worry, but face other serious problems. For discussion, see Broome (2007b), Brunero (2010), Brunero (2012), Fogal (2018), and Fink (2018), among others.

Sara without means

Sara intends to send off her paper by the end of the day. She believes that skipping her lunch break is a necessary means for achieving this. But Sara's favourite Italian restaurant runs a pizza lunch special that day and so she goes for a long lunch break instead.

Sara's case seems like a paradigmatic example of irrationality. The Simplistic Approach provides us with an easy explanation for why: we can derive an attribution of irrationality, since Sara violates a means-end coherence requirement of (structural) rationality. Now consider a different version:

Sara without ends

Sara intends to send off her paper by the end of the day, believes that skipping lunch is necessary for this and lacks the intention to skip lunch. But this time, Sara responds by dropping her intention to send off her paper today.

Still using the Simplistic Approach, we would now have to attribute rationality to Sara. After all, she now satisfies the means-end coherence requirement.⁹ This is where we encounter the symmetry problem: giving up one's end seems like an objectionable way of satisfying a rational requirement, especially compared to forming the intention to the means. To see this, let's assume further that Sara believes she has good reason to submit her paper on time since publishing is important for her career. Giving up such a well-supported intention willy-nilly seems like a prime example of irrational behaviour, or at the very least not like the sort of thing that should lead to a rationality attribution.¹⁰ While Brunero (2012, 126, emphasis in original) merely points out that "coming to intend to [skip lunch] seems *more rational* than ceasing to intend to [submit]", Schroeder (2004, 439) likens it to "the distinctive vice of [...] *rationalization*."¹¹ Attributing rationality to someone

⁹ At this point, one might worry that the Simplistic Approach is too simplistic. While it may be plausible that someone is *irrational* simply because they violate a rational requirement, whether they are *rational* may depend on more than whether they satisfy a single requirement. As we will see, I agree that the approach is too simplistic, but it is indeed what the requirements-based account suggests. Take Fogal again: "just as it is in virtue of violating such requirements that agents are irrational, it is in virtue of complying with them that they are rational" (Fogal 2020, 7). In cases like Sara's, where only one requirement is at play, satisfaction of this requirement then leads to attributing rationality. I say more about how I think we should deal with cases where multiple requirements apply in section 4.3.

¹⁰ Another reason why we might be hesitant to attribute rationality to Sara here is that her intentions are not persistent enough to play their role—namely to guide our actions and facilitate the successful implementation of our long-term plans (cf. Bratman 1987, 2009).

¹¹ Schroeder talks about changing one's beliefs about what one ought to do.

who displays this vice might strike us as the wrong result, even if the agent in question satisfies a means-end coherence requirement. Finally, consider a third version of the case:

Sara without beliefs

Sara intends to send off her paper by the end of the day. She also believes that skipping lunch is necessary for this and still lacks the intention to skip lunch. This time, she drops her belief that skipping lunch is necessary for submitting and enjoys her pizza lunch break.

Following the Simplistic Approach would yield the same result as before: Sara satisfies a means-end coherence requirement and hence would be attributed rationality. But if this result already seemed problematic in the previous case, it may seem even more objectionable now. The asymmetry seems to be more striking: while giving up the end is worse than intending the means, it is at least better than giving up the means-end belief (Bedke 2009). Brunero (2012, 126) even calls this a “downright irrational way of proceeding”. And yet, the Simplistic Approach would have us attribute rationality to the agent in this case.

These three cases illustrate that the Simplistic Approach of reading off attributions of rationality or irrationality from rational requirements is not satisfactory. It masks the asymmetry between various ways of complying with rational requirements and has us attribute rationality in situations where this seems controversial at best and outright false at worst.

2.2 Conflicting Attributions

The asymmetry between forming the intention to the means and giving up the means-end belief points to a deeper problem: following the Simplistic Approach seems to yield conflicting rationality attributions. What may satisfy one requirement of rationality may violate another. More specifically, what may satisfy a requirement in one *domain* of rationality may violate a requirement in another domain. And according to the Simplistic Approach, the very same behaviour would then have to be called both rational and irrational. This is worrisome for many reasons, but especially unwelcome on a conception that understands rationality as essentially linked to coherence.

Take *Sara without ends*. When Sara drops her intention to submit, she satisfies the means-end coherence requirement and hence can be attributed rationality. But this may make her violate the very same requirement in another instance, e.g. if she also believes that submitting today is necessary for advancing her career, and intends to advance her career. The way in which Sara can satisfy one instance of a rational

requirement can also lead her to violate another one. The Simplistic Approach would yield an attribution of both rationality and irrationality in this case.

Sara *without beliefs* is an example for conflicting attributions in different domains of rationality: what might make her *instrumentally* rational might at the same time make her *epistemically* irrational. Sara drops her means-end belief in order to satisfy the means-end coherence requirement. We have seen that following the Simplistic Approach would have us ascribe rationality to her. But once we broaden our focus only a little bit, we can see that the same behaviour might violate another rational requirement. What if Sara also believes that submitting comparable paper drafts in similar stages of completion has called for skipping lunch in the past? Arguably, if Sara now drops the belief that submitting *this paper* on time requires skipping lunch, she may be seen as violating the rational requirement to believe the consequences of her beliefs. Following the Simplistic Approach, this should result in an attribution of irrationality. The very same behaviour – no longer believing that submitting on time requires skipping lunch – warrants an attribution of rationality, but also of irrationality, if taken in isolation. That the same behaviour could have such drastically different results is a problem, as Schroeder (2004, 346) points out:

The problem [...] is that if [Sara] does respond to [her] situation in this way, [we have] to allow that though [Sara] is being epistemically irrational, [she] is in fact behaving impeccably, when it comes to subjective instrumental rationality. And that is a bizarre thing to say. Surely a good account of subjective instrumental rationality should not tell us that so far as instrumental rationality goes, this kind of behavior is okay.

Not only is the Simplistic Approach insufficiently sensitive to the asymmetries between options of satisfying rational requirements, it also yields conflicting rationality attributions.

2.3 Conflicts with Reasons

A third issue can be found in a general criticism of coherence rationality, namely that there can be conflicts between what is required by coherence rationality and what we have reasons to do.¹² For example, one might be required to believe the consequences of one's beliefs, even if these are all false and one in fact has good reasons not to hold these beliefs. To illustrate, consider the infamous bootstrapping objection (e.g. Broome 2007a; Kolodny 2005). Since any proposition entails itself, in order to believe the consequences of our beliefs, we are rationally required to believe

¹² For a very influential discussion of this point, see Kolodny (2005).

our beliefs—simply because we believe them, and regardless of whether they are true, or supported by the evidence.¹³

Similarly, agents whose ends are immoral or pointless do not seem to have reasons to intend the means to their ends.¹⁴ This, many contend, is puzzling and a serious strike against conceptions that understand rationality in terms of coherence.

To illustrate, take Broome's (2007a, 167) salmonella case:

For example, suppose the fish in front of you contains salmonella. This is a reason for you not to eat it. But there may be no obvious evidence that it contains salmonella. So you might not believe it contains salmonella, and you might eat it, and nevertheless you might be rational. So you are rational even though you do not respond correctly to the reason.

As also observed by Lord (2018, 23), “[t]his case, then, seems like a case where someone is rational even though they don’t do the thing that the reasons decisively support.” And again, the Simplistic Approach plays a role in bringing about this uncomfortable conclusion. An agent satisfies a rational requirement (e.g. to intend to eat the fish as a means to achieving their end of not being hungry) and hence is attributed rationality. At the same time, this positive evaluation of the agent stands in stark contrast with what they have reason to do (namely, not to eat the fish). This is the third issue with the Simplistic Approach: it unduly rewards behaviour that goes against reasons by calling it rational.

3. A First Solution

In his 2012 paper, Brunero develops a response to the symmetry problem which could be construed as a modification of the Simplistic Approach and may be extended to also deal with other issues raised in the previous section.¹⁵ Ultimately, I do not think his response is satisfactory, but it is on the right track and helps to motivate my favoured solution.

¹³ A more recent and sophisticated case where being coherent may conflict with being responsive to the evidence is put forward by Worsnip (2018). My approach has the resources to deal with this, as described in 4.4.

¹⁴ This point goes back to Raz (2005) and is taken up by Schroeder (2005), among others.

¹⁵ This is of course not the only proposed solution. For example, Broome (2013) argues that there are certain “basing prohibitions”, which rule out that certain attitudes (e.g. no longer believing you ought to x) can be based on others (e.g. no longer intending to x). If some ways of satisfying a requirement involve violating basing prohibitions and others do not, then we have accounted for their asymmetry. But Broome’s solution seems somewhat ad hoc; and he himself states that he does “not know how to give a proper analysis of *basing*” (Broome 2013, 141, emphasis in original).

Brunero (2012, 130) suggests that “[i]t’s perfectly consistent to claim both that (1) as far as [means-end coherence] goes, giving up an end is just as rational as giving up the belief, and (2) most often, all-attitudes-considered, it will be more rational to give up the end than give up the belief”. His strategy is to differentiate between two levels of evaluation: *local* and *all-attitudes-considered*. If we only look to the local level, that is, only to those attitudes immediately relevant for means-end coherence, we could maintain that all three ways of satisfying the requirement are rational, since they all lead to being means-end coherent. And so all versions of Sara, whether she gives up her end, or her means-end belief, or intends the means to her end, could be called rational. But importantly, Brunero maintains that we thereby don’t have to ignore the asymmetry of the options, since “judgments about whether one way of proceeding is more rational than another are dependent on the attitudes an agent has besides those directly relevant to the requirement of instrumental rationality” (Brunero 2012, 129). Once we consider *all* of an agent’s attitudes, the asymmetry can be accounted for. For example, if we also consider Sara’s belief that her means-end belief is well-supported, we can maintain that dropping this belief is a less rational way of satisfying the requirement than, say, dropping the intention to the end, or forming the intention to the means. Which of these last two options is the most rational will depend on Sara’s other attitudes.¹⁶ If, for example, she also believes that focussing on her well-being is important, the most rational way of being means-end coherent could be to drop her intention to skip lunch. Conversely, if Sara has no well-being focussed beliefs or intentions, forming the intention to skip lunch could be more rational. These differences in rationality would then be reflected in different rationality attributions. If Sara skips lunch but also believes in the value of prioritising lunch breaks, we can attribute rationality to her. If Sara drops her means-end belief while also believing that it is well-supported by evidence, we cannot attribute rationality to her.

3.1 The Refined Simplistic Approach

We can take Brunero’s suggestions to derive a *Refined Simplistic Approach* which keeps track of the two levels. On the local level, the approach remains unchanged. We can still attribute rationality or irrationality, depending on whether the agent satisfies or violates a rational requirement, based on the attitudes directly relevant to that requirement. In addition, we introduce the all-attitudes-considered level. On this level, we need to consider *all* of the agent’s attitudes, and whether any other

¹⁶ Brunero (2012, 130) acknowledges that “we would need some further explanation of what makes one way of proceeding more rational than another, all attitudes considered”.

requirements are satisfied or violated. This Refined Simplistic Approach makes it possible to yield a rationality attribution on the *local* level, but not on the *all-attitudes-considered* (*aac*) level for the same behaviour. This would allow us to keep track of the asymmetry.

If we further add an indexation of the levels to our attributions of rationality or irrationality, we could also address the second issue. The Refined Simplistic Approach then no longer necessarily yields conflicting attributions. In the case of *Sara without ends*, we could simply say that she is rational_{local} in dropping her end to submit, but not rational_{aac}. The same goes for conflicts between domains of rationality. The Refined Simplistic Approach would yield a local rationality attribution for *Sara without beliefs*, who is instrumentally rational, but not an all-attitudes-considered attribution, since she is epistemically irrational.

Unfortunately, the Refined Simplistic Approach does not seem to alleviate the third issue. The problem was that we positively evaluate behaviour that goes against reasons by ascribing rationality. But eating the fish may be rational both locally and all-attitudes-considered. As the case is presented, the agent does not have any beliefs about the fish containing salmonella, or any intentions that are incoherent with eating it. And so it may well be that eating the fish warrants a rationality attribution on both levels. Yet, it still conflicts with the reasons not to eat the fish, because it contains salmonella. The Refined Simplistic Approach yields the same result: behaviour that conflicts with reasons is ascribed rationality.

3.2 Another Issue: Comparisons

At this point, we can preliminarily conclude that the Refined Simplistic Approach, based on Brunero's suggestions, makes some headway towards addressing the issues. Appealing to different levels of evaluation is on the right track. But having just two levels of evaluation, as Brunero suggests, introduces an additional problem. The two-level approach does not give us the granularity we would want in our rationality attributions. In particular, it does not allow us to make comparisons between agents at points somewhere in between these levels. By having just two levels—one very specific local one that only considers the attitudes directly relevant to a requirement, and another maximally general level, considering the entirety of an agent's attitudes—we limit ourselves to the ends of a spectrum and neglect the entire space in between.

This is problematic because very often the space in between is what we are most interested in. Rarely is it the case that *all* of an agent's attitudes are relevant, since this can include all kinds of intentions or beliefs which might be tangential or entirely unrelated to the question at issue—such as a belief that green is a pretty

colour, or an intention to do a somersault. And similarly, we often are interested in more than just the attitudes directly relevant to requirements such as means-end coherence. The fact that we worry about asymmetry is evidence for this. Symmetry worries only arise once we include further attitudes in our evaluation of an agent’s rationality (such as the belief that another belief is well-supported). Once we do so, we have already left the local level of evaluation.

Being restricted to only two levels, at opposite ends of a spectrum, leads to a further problem: the loss of our ability to make comparative judgements. Say we want to compare three agents, *A*, *B* and *C*, with regard to their rationality. *A* is perfectly rational, they do not violate any rational requirements. Their attitudes are perfectly coherent. *C* only satisfies a means-end coherence requirement and violates all other rational requirements. Their only coherent attitudes are the ones directly relevant to means-end coherence. *B* sits between these two. Their attitudes are mostly coherent, but include some incoherencies. Hence, they satisfy a number of rational requirements, including a means-end coherence requirement, but also violate some. How do our three agents fare when we evaluate their rationality, and decide whether to attribute rationality to them, using the two levels of the Refined Simplistic Approach?

At the local level, where we only include attitudes directly relevant for means-end coherence, all three agents satisfy the requirement and hence merit a rationality attribution (Table 1). When we move to the all-attitudes-considered level and consider all the agents’ attitudes, only the perfectly rational agent *A* satisfies all rational requirements, and merits a rationality attribution. Both agent *B* and *C* would violate rational requirements (though *C* violates many more than *B*). Hence, none of these two merit a rationality attribution. These results are summarised in Table 1.

Table 1: pairwise comparisons

agent	local	aac
<i>A</i>	✓	✓
<i>B</i>	✓	✗
<i>C</i>	✓	✗

If we now turn to pairwise comparisons between the agents, we can see how being limited to only two levels of evaluation is insufficient. The grey cells are of particular interest here. If we compare the perfectly rational agent *A* with agent *C* who is almost entirely irrational, with the exception of the attitudes directly relevant to means-end coherence, the verdicts are identical on the local level. Both agents qualify as rational on this level of evaluation, despite the massive differences in their rational performance. At first sight, this is a strange result.

Now of course, one might respond by pointing out that *A* and *C* are equally rational, from the local point of view of means-end coherence. And so we should expect and endorse the same verdict for both of them. I agree that pointing to the level of evaluation is important, e.g. by adding an indexation to the rationality attributions of the Refined Simplistic Approach. This would make statements like “*A* is rational_{local}” and “*C* is rational_{local}” more plausible. But at the same time, the vast differences between the agents are not captured, which can be misleading. As we will see in section 4, my approach has the means to represent these differences and is therefore preferable.

The picture looks similar if we consider the pairwise comparison between *B* and *C* on the maximally general level of evaluation. If we take into account all attitudes and applicable rational requirements, both agents violate some and hence do not merit a rationality attribution. Using indexation of the levels, we would have to say that “*B* is irrational_{aac}” and “*C* is irrational_{aac}”. But clearly *B* is much more rational than *C*, who violates almost all rational requirements, with means-end coherence as their single small island of rationality. This nuance is completely lost once we turn to the rationality attributions at this level—these two agents now seem on a par with regard to their rationality.

Before turning to my approach, let me note that these problems are exacerbated once we remind ourselves of the role that rationality attributions play in our practices of criticising agents. Since a rationality attribution is usually seen as commendable, and an attribution of irrationality as criticisable, it is especially important that these attributions are maximally informative and not misleading. If we are going to criticise someone for being irrational, we better be sure that our assessment of their rationality accurately reflects the state of their propositional attitudes.

To conclude this section, we have seen that we can derive a Refined Simplistic Approach, based on Brunero’s suggestions. By explicitly mentioning two levels of evaluation, we are able to address worries about symmetry and conflicting rationality attributions. But worries about clashes with reasons and informative

Lisa Bastian

comparisons persist. To fully address these, we will have to further refine our approach.

4. Better Rationality Attributions

The problems discussed in the previous sections can be translated into desiderata for a better approach to rationality attributions:

Specificity Which attitudes are evaluated plays a crucial role for whether a rationality attribution is warranted. Therefore, rationality attributions should be specific to the set of attitudes in question.

Flexibility In choosing the level of evaluation, or set of attitudes to evaluate, flexibility is important. We should be able to evaluate very small sets of attitudes, the entire set of attitudes, and sets in between.

Informativeness Rationality attributions should provide information about the comparative differences between agents.

The need for Specificity was already recognised by Brunero, when he acknowledged that “judgments about whether one way of proceeding is more rational than another are dependent on the attitudes an agent has besides those directly relevant to the requirement of instrumental rationality” (Brunero 2012, 129). He tried to incorporate something like Specificity by distinguishing between a local and all-attitudes-considered level. But we have also seen that allowing for only two levels, located at the ends of a spectrum, is not sufficient, since we are often interested in levels somewhere in between. This is recognised by Flexibility. A lack thereof had the result that we are not always able to compare agents in a way that is informative enough to appropriately represent their differences in rationality, creating the need for Informativeness.

To meet these desiderata, it will be helpful to remind ourselves of the nature of coherence rationality. On this view, rationality is ultimately a matter of propositional attitudes and their patterns. An agent’s attitudes are essential to their rationality. The central role that attitudes play for coherence rationality should be reflected in attributions of rationality or irrationality—they should be explicitly stated. Therefore, I suggest that we understand the commonly used locutions

‘S is rational’ and

‘S is irrational’

as shorthand for the more precise attributions

‘S is rational with regard to subset a_n of her attitudes’ and

‘S is irrational with regard to subset a_n of her attitudes’.^{17, 18}

Hence, the main characteristic of my approach is an explicit reference to a set of attitudes.

4.1 Reference to a Set of Attitudes

What do I mean by the phrase ‘with regard to subset a_n of her attitudes’?¹⁹ A helpful way to visualise this is to think of the set and subsets of an agent’s attitudes as nested circles (see Fig. 1). The outer circles correspond to supersets of the inner circles, with every inner circle being a proper subset of the next outer circle. The smallest circle corresponds to a singleton set containing only one attitude (a_1).²⁰ We can then extend this smallest circle by adding more attitudes. This way, we extend the singleton set by adding elements to it, thereby forming a bigger set containing more attitudes (a_1, a_2, \dots), centred around the singleton set.

¹⁷ In this paper, I restrict myself to the properties of ‘rational’ and ‘irrational’. The latter is sometimes also used interchangeably with ‘not rational’. This can be confusing, since ‘not rational’ could either refer to the property of irrationality, or to a-rationality, e.g. when talking about entities that cannot be rational or irrational, like furniture. Since I am interested in attributions of the positive evaluative property of rationality and the negative evaluative property of irrationality, I refrain from using the locution ‘not rational’ to avoid confusion.

¹⁸ Note that I am not committing to the descriptive claim that we always have the more complex and precise attribution in mind when we use the shorter locution. I am defending the claim that the more complex locution is a helpful and more precise way to think about rationality attributions.

¹⁹ I take propositional attitudes to be central for rationality attributions. They are the locus of rationality, and agents are rational or irrational in virtue of the attitudes they do or do not have. In the metaethical, epistemological and Philosophy of Action literature on rationality, beliefs and intentions are seen as paradigmatic propositional attitudes. Work on rationality in these fields almost exclusively focusses on those. There are of course many other propositional attitudes, such as desires, or preferences. But these debates do not take up centre stage in the part of the literature that I am concerned with, and so I do not address them. But it should be noted that preferences in particular could be easily accommodated by the framework developed in this paper, since the requirements for rational preferences can be understood as coherence requirements (see e.g. Zynda 1996).

²⁰ My picture allows for singleton sets only at the lowest level of evaluation (a_1).

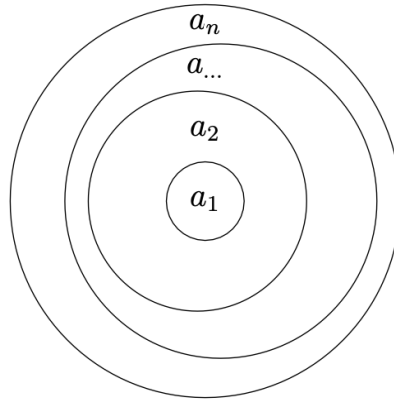


Figure 1: sets and subsets

In principle, any of the agent’s attitudes can form the centre, or the element of the singleton set. Which one we choose depends on pragmatic factors, such as which attitudes we are currently interested in. Some of the bigger circles, or supersets, might not be interesting for the agent’s rationality. They might be random collections of attitudes, like an intention to book plane tickets, a belief that crimson is a dark shade of red, or a belief that the Pythagorean theorem is true.

4.2 Choosing Sets

We have a lot of freedom when it comes to choosing the subset of attitudes to evaluate for rationality. There is not just one correct subset that is the object of evaluation, but many. These subsets can be as varied as our interests when we evaluate agents for their rationality. If in a given situation, we are particularly interested in an agent’s ability to follow their plans, then we will focus on attitudes related to means-end coherence. If we are more interested in the state of an agent’s belief set, the attitude set we are evaluating will contain beliefs. So the set of attitudes we are evaluating for rationality changes with our interests and with the situation. But we need to make sure that we make clear which subset we are focusing on – which is taken care of by explicitly mentioning set a in question in the rationality attribution. This flexibility is a major advantage of my approach.

Additionally, pragmatic factors can influence the set of attitudes we end up evaluating. Often, the way in which a case or an agent’s situation is presented primes us to consider a specific attitude set. This priming can happen directly through *markers*, such as signal words or phrases, or indirectly, by the fact that only some of the agent’s attitudes are *explicitly mentioned*. We have already seen examples of this in the introduction.

Explicit Mentioning

Consider again the case of Candice:

Candice decides to go to the post office this afternoon to send out some mailings, but on the way there, she gives up on this end and decides to go buy groceries instead. But on the way to the market, she yet again trades in this end for another: going to hang out with her friend David. But on the way to David's house, she once more changes her mind and intends to spend a relaxing afternoon at home, but by the time she gets home the afternoon is gone and she's accomplished nothing. We're inclined to accuse Candice of some kind of irrationality. And, since Candice failed to follow through on any of her ends, we're inclined to accuse her of being instrumentally irrational. (Brunero 2012, 236)

It is fair to assume that the attitudes explicitly mentioned in this example do not exhaust Candice's attitudes. Surely, Candice's mental life is a lot richer than could be captured by a mere four attitudes. This suggests that the attitudes that *are* explicitly mentioned are of particular interest, or particularly important in this situation. As a result, the subset we are primed to evaluate for rationality contains an intention to go to the post office, an intention to buy groceries, an intention to visit David, and an intention to spend the afternoon at home.

Markers

Priming can also take the form of markers. As an example, reconsider Kolodny's exchange:

Look: I think you're nuts to believe in God, let alone that He created anyone. But *given that* you believe that God created all people equal, and *given that* you agree that people whose skin is a different colour from yours are people, you ought to believe that He created them equal too. It would be irrational of you not to. (Kolodny 2005, 554, emphasis added)

Similarly to the how Candice's case was presented, the presentation of this exchange primes us to only take into account the explicitly mentioned attitudes of the agent: a belief that God created all people equal, and a belief that people whose skin is a different colour from yours are people. Moreover, the locution 'given that' marks these attitudes as particularly important in this case.

In addition, there are cases where the markers correspond to an entire sub-domain of rationality and hence mark the corresponding attitudes as salient. We can see this in Candice's case as well, which features phrases like 'as far as instrumental rationality goes' and 'relevant to instrumental rationality'. These suggest as salient those attitudes that are governed by requirements of a sub-domain of rationality (in this case instrumental rationality).

Finally, there are markers like ‘overall’. These markers suggest as salient the entire set of an agent’s attitudes, much like Brunero’s all-attitudes-considered level of evaluation. It corresponds to the biggest, outermost circle which represents the entirety of an agent’s attitudes and includes all other subsets. Call this a_n , which is centred around a_1 .

It is important to note that the ‘overall’-marker interacts with the explicitly mentioned attitudes. Absent any other information about the agent, the entire set of an agent’s attitudes corresponds to the set of attitudes we can be certain they have—namely the explicitly mentioned ones. This is not to say that the agent might not have additional attitudes that we simply do not know about. And in fact, they almost certainly have more attitudes, since an agent’s attitude profile is usually not exhausted by only a handful of attitudes.²¹ It is worth noting that in this way, rationality attributions are contingent on the available information.

4.3 Evaluating Sets for Rationality

So much for sets and subsets. The next idea that I have to clarify is what it means to be rational *with regard to a subset of attitudes*. With the requirements-based account of rationality in the background, I take this to simply mean that one satisfies all rational requirements that apply to that subset. A rational requirement applies to a subset of attitudes if it can be assessed for compliance or violation. Put differently, a requirement applies to a subset if the attitudes it governs are elements of that subset. For example, a means-end coherence requirement applies to a subset containing an intention to an end, a means-end belief and an intention to the means, but not to a subset containing the belief that p , the belief that if p , then q , and the belief that q . This latter subset features attitudes that are governed by a requirement to believe the consequences of our beliefs, which would then apply to it. Following the same reasoning, being irrational *with regard to a subset of attitudes* means that one violates at least one of the rational requirements that apply to the subset in question.

The property of rationality reflects the set-subset relation. If an agent is rational with regard to a set of her attitudes (say a_2), she is also rational with regard to the relevant proper subsets (in this case, a_1). Because the subsets that rationality attributions attach to are linearly ordered by inclusion via the subset relation, the attributions are consistent: being rational in a_{n+1} implies being rational in a_n , but not vice versa. The property of irrationality is different: agents can be irrational with

²¹ Brunero even goes so far as to say that such cases are often under-described. While I agree that the description surely is not exhaustive, I maintain that only focussing on some of the agents’ attitudes can be appropriate, given our interests.

regard to the larger set a_2 , but rational with regard to a smaller proper subset a_1 , if they satisfy the requirements that apply to a_1 .

Note that not only present but also absent attitudes impact whether we can assess a requirement for compliance or violation, thereby making the requirement apply to the subset.²² This is a feature of understanding rationality in terms of coherence, that is, of taking rationality to be a matter of displaying certain patterns of attitudes.²³ Whether this is achieved can depend both on present and absent attitudes. Take *Sara without means*. Sara has an intention to an end, a means-end belief but lacks the intention to the means. Due to the absence of an attitude (the intention), she is a paradigmatic example of a means-end incoherent agent. So we can specify further: a requirement applies to a set of attitudes if that set features the presences or absences of the attitudes that are governed by the requirement. And agents are rational with regard to a subset if they satisfy the rational requirements that apply to that subset, and agents are irrational with regard to a subset if they violate one of the rational requirements that apply to it.

4.4 Meeting the Desiderata

Let's see how my approach meets the desiderata and addresses the issues raised in 2. Since the set of attitudes that is evaluated for rationality is explicitly stated in the attribution (' S is rational with regard to subset a_n of her attitudes'), *Specificity* is obviously met. And since there are no a priori restrictions on which subset to select, *Flexibility* is met as well. The set in question can be as large or small as we want, thereby enabling us to make attributions at the local level, at the all-attitudes-considered level, in keeping with the Refined Simplistic Approach, but also at any level in between.

This allows my approach to accommodate worries about asymmetry in a similar but superior way to the Refined Simplistic Approach. We can maintain that some of the more questionable ways of satisfying means-end coherence would not warrant a rationality attribution on a higher-up level, i.e. regarding an attitude set

²² There is precedence for including absent attitudes. Take Kieseewetter (2017, 16), who uses "attitude-state" to refer to "the generic term for both attitudes and lack of attitude". Fogal (2020, 5) specifies that "talk of 'having attitudes' is meant to include the absence as well as the presence of attitudes". Similarly, Brunero (2012, 28) states that "[a]n attitude-state could consist in either the presence or absence of an attitude."

²³ Even on a reasons-view, it is sometimes necessary to include absent attitudes. Assume that reasons-rationality can roughly be understood as the idea that agents should respond correctly to reasons. If an agent then does not form a belief that is supported by their reasons, the absence of this belief has to be taken into account, since it amounts to a failure of rationality.

a_m that includes many of an agent's attitudes, while also holding that all ways of satisfying a means-end coherence requirement would be rational with regard to a much smaller, more local subset a_l (where $l < m$)—we are simply evaluating the agent at different subsets.

To see this, consider again *Sara without means*: Sara intends to submit her paper today, believes that this necessitates skipping lunch, but does not intend to skip lunch. We have noted that there are at least two different verdicts that seem plausible here: we might think that Sara is irrational in not intending to skip lunch because this makes her means-end incoherent. But once we took into account additional attitudes of hers, like an intention to focus on well-being, not intending to skip lunch could be seen as rational. My approach gives us the resources to better deal with this case. We can now see that both these verdicts are adequate and need not be in tension with each other because the attributions simply refer to different subsets of Sara's attitudes. We can maintain that Sara is irrational with regard to the subset of her attitudes that contains her intention to the end, her means-end belief and the absence of her intention to the means. And we can also maintain that she *is* rational with regard to the subset of her attitudes that includes the absence of her intention to skip lunch and the intention to focus on well-being. Importantly, we can do so by evaluating subsets at whichever level is relevant, instead of being limited to only the local and the all-attitudes-considered level, which is an improvement to the Refined Simplistic Approach.

The worry about conflicting rationality attributions can be addressed in much the same way. In *Sara without ends*, we can maintain that dropping the intention to submit on time is both rational and irrational without creating a conflict: it is rational with regard to the subset that only contains the attitudes directly relevant to means-end coherence, but irrational with regard to a subset that also contains a belief that submitting on time is crucial for her career (since the applicable rational requirement(s) are satisfied in first case, but not in the second). The same goes for conflicts between domains of rationality. We can maintain that *Sara without beliefs* is rational with regard to the set of attitudes relevant to instrumental rationality, but irrational with regard to the set of attitudes relevant to epistemic rationality. And again, we are not limited to only the local and the all-attitudes-considered level, but have the flexibility to focus on attitude sets in between.

Further, we can see how my proposed approach can address conflicts with reasons, which was not possible for the Refined Simplistic Approach. The explicit reference to the set of attitudes in question provides the necessary context to alleviate the worry. Instead of simply calling acting against the reasons rational, we now say that eating the fish is rational with regard to the set of attitudes that contains

an intention to no longer be hungry and a belief that eating the fish is a means to this end. We are directly pointing out that the rationality attribution only applies to this set of attitudes, instead of issuing a general positive evaluation of the agent and their rationality. This makes calling the behaviour rational (in relation to the set of attitudes) much more palatable, despite going against the reasons (since the fish contains salmonella). There is no conflict in maintaining that from the perspective of this attitude set, which importantly does not contain a belief about the fish containing salmonella, it is rational to eat the fish, and also that there are reasons against eating the fish. The same reasoning can be applied to immoral or unreasonable ends: agents that pursue the means to these ends can be seen as rational with regard to the set of attitudes concerned with means-end coherence, which is compatible with there being reasons against intending those means.

Turning to comparisons, we can see how having our rationality attribution explicitly refer to the subset at issue provides a more accurate and *informative* picture. When comparing the perfectly rational agent *A* with the almost entirely irrational agent *C*, we indeed maintain the same rationality attribution when we focus on the attitude set that only contains intentions to end and means and the means-end belief. But this will be explicitly acknowledged, so there is no risk of creating the impression that the two agents were equally rational in general. Once we look at slightly bigger attitude sets, the difference becomes clear: *A* still qualifies as rational, since they still satisfy the applicable requirements, whereas *C* can no longer be attributed rationality. The difference between *C* and the not perfectly but fairly rational agent *B* is equally maintained. Both qualify as rational with regard to the same local subset, but only *B* can still be attributed rationality once we move further out and add additional attitudes to the set.²⁴ At the same time, there is no risk of confusing *B* with *A*, since only *A* can be attributed rationality with regard to set of all attitudes a_n .

Conclusion

This paper was motivated by the observation that in discussions about rationality, rationality attributions have largely but erroneously been seen as a given. This can be interpreted as the result of uncritically accepting the Simplistic Approach, motivated by the requirements-based account. But we have seen that this approach

²⁴ While it may be tempting, I currently do not take my approach to provide an account of degrees of rationality. It can capture the differences between agents *A*, *B*, and *C* in terms of different attributions of rationality or irrationality, but not in terms of whether one is more or less rational. Whether the size of a subset to which a rationality attribution is relative can be understood as a measure of degrees of rationality needs to be explored in future work.

leads to a number of issues that make the resulting rationality attributions inadequate. Turning our attention to rationality attributions themselves allowed us to see the need for more specific, flexible and informative attributions. This can be achieved by realising and making explicit that rationality attributions are relative to the set of attitudes in question. Explicitly mentioning the set in question meets the established desiderata and avoids the issues faced by both the Simplistic and the Refined Simplistic Approach. Moreover, providing an account of rationality attributions also addresses a lacuna in the wider discussion, which usually focusses on the form and content of requirements of rationality. Given a coherence understanding of rationality, it is no surprise that propositional attitudes play such a prominent role when it comes to attributing rationality and irrationality. And finally, since (ir-)rationality attributions underwrite our practices of praising or criticising agents, these evaluative practices can be greatly improved by relying on the more precise and informative approach developed here.

References

- Bedke, Matthew S. 2009. "The Iffiest Oughts: A Guise of Reasons Account of End-Given Conditionals." *Ethics* 119(4): 672–98.
- Bratman, Michael. 1987. *Intention, Plans, and Practical Reason*. Center for the Study of Language and Information.
- Bratman, Michael E. 2009. "Intention, Practical Rationality, and Self-Governance." *Ethics* 119(3): 411–43.
- Broome, John. 1999. "Normative Requirements." *Ratio* 12(4): 398–419.
- . 2005. "Have We Reason to Do as Rationality Requires?—a Comment on Raz." *Journal of Ethics and Social Philosophy* 1 (Symposium): 1–10.
- . 2007a. "Is Rationality Normative?" *Disputatio* 2(23): 161–78.
- . 2007b. "Requirements." In *Hommage à Wlodek; Philosophical Papers Dedicated to Wlodek Rabinowicz*, edited by Toni Rønnow-Rasmussen, Björn Petersson, Jonas Josefsson, and Dan Egonsson.
- . 2007c. "Wide or Narrow Scope?" *Mind* 116(462): 359–70.
- . 2013. *Rationality Through Reasoning*. Wiley-Blackwell.
- Brunero, John. 2010. "The Scope of Rational Requirements." *Philosophical Quarterly* 60(238): 28–49.
- . 2012. "Instrumental Rationality, Symmetry and Scope." *Philosophical Studies* 157(1): 125–40.
- Fink, Julian. 2018. "The Property of Rationality: A Guide to What Rationality Requires?" *Philosophical Studies* 175(1): 117–40.

- Fogal, Daniel. 2018. "On the Scope, Jurisdiction, and Application of Rationality and the Law." *Problema* 12: 21–57.
- . 2020. "Rational Requirements and the Primacy of Pressure." *Mind* 129(516): 1033–70.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.
- Kiesewetter, Benjamin. 2017. *The Normativity of Rationality*. Oxford: Oxford University Press.
- Kolodny, Niko. 2005. "Why Be Rational?" *Mind* 114(455): 509–63.
- . 2007. "State or Process Requirements?" *Mind* 116(462): 371–85.
- Lord, Errol. 2018. *The Importance of Being Rational*. Oxford, UK: Oxford University Press.
- Nozick, Robert. 1993. *The Nature of Rationality*. Princeton University Press.
- Raz, Joseph. 2005. "The Myth of Instrumental Rationality." *Journal of Ethics and Social Philosophy* 1(1): 28.
- Schroeder, Mark. 2004. "The Scope of Instrumental Reason." *Philosophical Perspectives* 18(1): 337–64.
- . 2005. "Instrumental Mythology." *Journal of Ethics and Social Philosophy* 1(2).
- . 2013. "Scope for Rational Autonomy." *Philosophical Issues* 23(1): 297–310.
- Southwood, Nicholas. 2008. "Vindicating the Normativity of Rationality." *Ethics* 119(1): 9–30.
- Way, Jonathan. 2018. "Reasons and Rationality." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, 485–503. Oxford University Press.
- Worsnip, Alex. 2018. "The Conflict of Evidence and Coherence." *Philosophy and Phenomenological Research* 96(1): 3–44.
- Zynda, Lyle. 1996. "Coherence as an Ideal of Rationality." *Synthese* 109(2): 175–216.