Reasons, Rationality, and Opaque Sweetening: Hare's 'No Reason' Argument for Taking the Sugar

Ryan Doody

Brown University

Forthcoming in *Noûs*. Please cite published version.

Abstract

Hare (2010) presents a compelling argument for 'taking the sugar' in cases of opaque sweetening: you have no reason to take the unsweetened option, and you have some reason to take the sweetened one. I argue that this argument fails—there's a perfectly good sense in which you *do* have a reason to take the unsweetened option. I suggest a way to amend Hare's argument to overcome this objection. I then argue that, although the improved version fares better, there's still room to resist Hare's argument—in a way that raises interesting questions about rational agency. In short, rationality *isn't* about doing what one *has* the most reason to do; rather, it's about aiming to do what *there is* most reason to do.

1 Introduction

Hare (2010) introduced the puzzle of *opaque sweetening*. Here's a version of it.

> Vacation Boxes. There are two opaque boxes: a Larger box (*L*) and a Regular box (*R*). A fair coin has been tossed. If the coin landed heads, a voucher for an all-expenses-paid Alpine ski vacation (*A*) was placed in the Larger box and a voucher for an all-expenses-paid beach vacation (B) was placed in the Regular box; if the coin landed tails, B was placed in the Larger box and A was placed in the Regular box. In either case, you don't know which prize is in which box.
>
> Now imagine that $1 is added to the Larger box. If you choose the Larger box, you will win whichever prize it contains plus $1. Nothing is added to the Regular box. You are asked to choose a box.

	Heads	Tails
L^+	$A^{+\$1}$	$B^{+\$1}$
R	B	A

Let's suppose that you don't prefer either vacation to the other, nor are you indifferent. Because you consider each vacation to be good in a variety of ways, your attitude is one of *ambivalence*.[1]

The puzzle exploits a characteristic feature of ambivalence—namely, that it, unlike indifference, is *insensitive to mild sweetening*: sometimes, if you are ambivalent between X and Y, you are also ambivalent between X and Y^+ (a slightly improved version of Y). In this case, you're ambivalent between the ski vacation and the beach vacation, you don't prefer the beach vacation *plus a dollar* ($B^{+\$1}$) to the ski vacation, and you don't prefer the ski vacation *plus a dollar* ($A^{+\$1}$) to the beach vacation. Small improvements aren't enough to resolve your ambivalence.

Is there a box you should take? Or is it okay to take either? The puzzle brings out a tension between two attractive ideas. One of these ideas supports the verdict that you should take L^+ over R. And the other supports the verdict that it is okay to take either.[2]

This essay concerns a powerful argument for the former view: Caspar Hare's 'No Reason' Argument. Hare argues that you have *no reason* to take R rather than L^+, but that you *do* have a reason—namely, that you'll get a dollar—to take L^+ rather than R. And that, therefore, it's irrational to take R when you can take L^+ instead.

In what follows I will reconstruct Hare's argument and argue that, as it stands, it fails. I will then suggest a way to amend the argument to surmount the objection. I then argue that, although the improved version fares better, there's still room to resist it—in a way that raises some interesting questions about rational agency.

2 The 'No Reason' Argument

Here's Hare's argument, quoted in full:

[1] I use 'ambivalence' to denote a psychological state, related to the phenomenon that others have characterized in terms of 'rough equality' (Parfit, 1984, p. 431), or 'incomplete preferences' (Hare, 2010), or 'incommensurability' (Broome, 2007; Raz, 1986), or 'parity' Chang (1997, 2002, 2009, 2017). Ambivalence is conative, not cognitive: it might correspond to, but is not identical with, a belief about the objective value-relations that hold between the items. Chang's 'parity', on the other hand, is a (non-transitive) fourth value-relation; in addition to "better than," "worse than," and "as good as," comparable alternatives might be "on a par."

[2] The former view is defended by Bader (2018); Hare (2010, 2013) (and, somewhat indirectly, in Doody, 2019a). The latter is defended by Bales et al. (2014); Schoenfield (2014) (and, somewhat indirectly, in Doody, 2019b). (For more discussion of the dialectic, see Doody, 2021).

> There is a consideration of which I am aware that counts in favour of my taking the [Larger box], rather than the [Regular] box: I will get a dollar if I take the [Larger] box, no dollar if I take the [Regular] box. But there is no consideration of which I am aware that counts in favour of my taking the [Regular box], rather than the [Larger] box. So, it is rationally impermissible for me to take the [Regular] box. It is rationally impermissible to do something when I have no reason to do it, and a reason to do something else. (Hare, 2010, p. 240)

We can understand Hare as making, roughly, the following argument:

No Reason Argument

P1 *Reason Principle:* If you have a reason to φ rather than ψ, and no reason to ψ rather than φ, it's irrational to ψ.

P2 In Vacation Boxes, you have a reason to take the Larger box rather than the Regular box (i.e., you'll receive a dollar).

P3 In Vacation Boxes, you have no reason to take the Regular box rather than the Larger box.

C In Vacation Boxes, it's irrational to take the Regular box.

This is a valid argument with three plausible premises. Although Hare speaks of 'considerations of which I am aware that count in favour', we can characterize the same idea in terms of what *reasons you have* for acting.[3] Providing a comprehensive account of what reasons are—and of what it takes to *have* one—is complex and notoriously vexed.[4] To fix ideas, let's say:[5]

[3] In fact, this is the language Hare (2013, p. 47) uses when presenting the argument.

[4] Some philosophers (e.g., Raz, 1999; Scanlon, 1998) think the notion is ultimately unanalyzable. Others, like Broome (2007, 2013), take reasons to be facts that, in some sense, *explain* why it is you ought to do what you ought to do. Still others (e.g., Setiya, 2007, 2014; Smith, 1987, 1995) analyze reasons in terms of *ideal deliberation*: roughly, a reason to φ is a fact such that, were you to be aware of it and ideally rational (and, perhaps, ideal in some other respects as well), it would motivate you to φ. I wish to remain as neutral as possible about these issues.

[5] This analysis is heavily inspired by the one sketched in Bales et al. (2014, p. 466), with four noteworthy differences. First, they don't distinguish between *being a reason* and *being a reason you have*. They claim to be giving an account of the former, but are better understood to be giving an account of the latter. Second, they capture the epistemic component (which is voiced by (3) on our account) by only requiring (1) to hold for *some* way the world might be. It's more perspicuous to tease these apart. Third, they take p—not r—to be the reason, which awkwardly proliferates reasons (e.g., "I'll win a million dollars" is a reason to do anything that you think *might*—however unlikely—result in you winning a million dollars). Lastly, they understand (2) to mean: "p's being true is better than p's being false," which they identify with preferring p to $\neg p$. They point out that,

> r is a *reason that you have* to φ rather than ψ if, and only if, …
>
> (1) *Contrastivity:* r can be recast in the form ⌜if I φ, then $p \wedge$ if I ψ, then $\neg p$⌝.
>
> (2) *Positivity:* For you, p is a good thing.
>
> (3) *Awareness:* You believe that r is true and is a reason for you.

These conditions are meant to characterize the *reasons you have* (i.e., your "subjective reasons") and not the *reasons there are* (i.e., your "objective reasons").[6] I'll have more to say about the distinction in §4.2. There are ways to resist this analysis, but I think it does a decent enough job capturing what Hare must have in mind.

Let me unpack each condition in turn. First, consider *Contrastivity*. In order for r to be a reason for you to φ *rather* than ψ, it must be a consideration that counts more strongly in favor of φing than ψing. The fact that you'll give me a high-five if I φ might be a consideration that speaks in favor of φing, but, if you'll give me a high-five no matter what I do, it isn't a consideration that speaks in favor of φing *rather* than something else instead (e.g., ψing). And so, *that you'll give me a high-five* isn't a reason to φ rather than ψ. Only considerations that *distinguish* between the options can give rise to reasons to do one rather than another.

Second, consider *Positivity.* Reasons are considerations that count *in favor* of something. Clause (2)—that, for you, p is a good thing—is meant to capture the idea that r speaks in favor of, and not against, φing over ψing. If, for you, p is bad or neutral, then the fact that p will obtain if you φ won't count as a reason you have to do it.

There are various ways to understand clause (2), and I intend to remain as neutral as possible between them. In particular, I don't wish to take a stand on whether p must be *objectively* good or merely be something you *consider* to be good. Externalists (who think reasons are provided by objective evaluative facts)

---

because p picks out a sub-maximal state of affairs, there are various ways to understand what it is to "prefer" p to $\neg p$. They endorse a view according to which you prefer p to $\neg p$ just in case, *all else equal*, you prefer p being the case to $\neg p$ being the case. For this reason, they deny **P2**—receiving the dollar, on their view, is *not* a reason take the Larger box (or, at least, not a reason that "can reliably be plugged into the [*Reason Principle*]" without begging the question). I'll say more about this in §3.2.

[6] A famous motivation for the distinction is Bernard Williams' petrol example (Williams, 1979). Since then, it's been widely discussed (e.g., Fogal and Worsnip, 2021; Markovits, 2014; Schroeder, 2007, 2009; Sepielli, 2018; Setiya, 2004; Sylvan, 2015; Way, 2009; Wodak, 2017, 2019). According to a popular family of views, subjective reasons can be analyzed in terms of objective reasons (e.g., Parfit, 2011; Schroeder, 2007, 2009; Sylvan, 2015; Vogelstein, 2012; Way, 2009). For example, according to the Factoring Account, you *have* a reason to φ when *there is* a reason to φ which you (in some sense) *have*—e.g., by being aware of it. (Schroeder (2008) argues the view is false; Lord (2010) defends it.) I don't wish to take sides on this.

and Internalists (who think that only desires provide reasons) are welcome to understand (2) in whichever way they'd like. Feel free, then, to replace "good" with "valuable" or "desirable" or to say instead that "you'd welcome the news that *p*" or that "you desire that *p*" or—as Bales et al. (2014, p. 17) say—that "you prefer *p* to ¬*p*" or anything else in the neighborhood.[7] Clause (2) is only meant to capture the idea that r is a consideration that counts in *favor* of doing something.

Finally, consider *Awareness*. Suppose that, unbeknownst to you, the lottery ticket on your left is the Mega Million Jackpot winner, while the ticket on your right is not. Suppose (not implausibly) that, for you, winning the Mega Million Jackpot is a good thing. Offhand, then, there is a reason for you to buy the ticket on your left rather than ticket on your right—if you buy the ticket on your left, you'll win the Mega Million Jackpot; but, if you buy the ticket on your right, you won't. However, this is not a reason *you have*. Because you have no idea that it's the winner, you aren't in a position to recognize it is a reason for you.

Because we are interested in what it's *rational* for you to do, and because rationality is about doing what makes the most sense given your perspective, we are most concerned with the reasons *you have* for doing one thing or another, not the reasons *there are* (of which you might be completely unaware).

3 Assessing the Argument

With this rough explication of what it is to *have a reason* in hand, let's look at each premise.

3.1 Premise 1: The Reason Principle

The first premise, recall, says:[8]

[7] This is not to suggest that the relationships between these notions are straightforward. To the contrary, there are some subtle and interesting issues about, e.g., the relationship between *wanting* p and having a *desire* that's satisfied in exactly the worlds in which *p* is true (see, for example, Grant and Phillips-Brown, 2020, who argue that *desire satisfaction* is "ways-specific"), and the relationship between *wanting* and *preference* (Lewis, 1986; Phillips-Brown, 2021; Stalnaker, 1984). Moreover, some philosophers (notably, Scanlon, 1998) endorse a *Buck-Passing Account of Value*, according to which evaluative notions (like, *good* or *valuable* or *desirable*) are understood as the property of having other, first-order properties that provide one with reasons to hold certain pro-attitudes toward their objects. (For example, on these views, if you consider *p* to be a good thing, there are features of *p* that provide you with *reasons* to value it.) If the *Buck-Passing Account of Value* is correct, then the characterization of reasons offered here appears problematically circular: r is reason only if *p* is good, but *p* is good only if there are reasons to value it. Let this not detain us, however. I don't intend to be giving a *reductive analysis* of reasons (which, if Raz, 1999; Scanlon, 1998, are correct, isn't possible anyway), just a characterization to help us better assess Hare's argument. And, for my purposes, not much turns on the specific formulation of (2).

[8] Hare (2013, p. 47) calls this principle '*Rational Permissibility Tracks Reason*': "If you are rational and two options are open to you, and you have a reason to take the one-rather-than-the-other,

> *Reason Principle*: If you have a reason to φ rather than ψ, and no reason to ψ rather than φ, it's irrational to ψ.

This principle is fairly plausible. It's generally accepted that rationality involves correctly responding to—or being guided by—your reasons.[9] If you have a reason to φ rather than ψ, and no reason to ψ rather than φ, then your reasons decisively tell in favor of φing over ψing. The balance tilts in only one direction. Correctly responding to your reasons, then, is incompatible with ψing when you can φ instead. And so, it's irrational to ψ.

To drive the point home: imagine you know, whichever you choose (φ or ψ), you'll be asked to explain yourself—to offer some justification for why you chose as you did. If you have *some* reason to choose φ rather than ψ and *no* reason to choose ψ rather than φ, there appears to be nothing you can say to explain choosing ψ over φ. Given that there *is* something you can say to explain choosing φ rather than ψ, if you want to be able to explain why you did what you did, you shouldn't choose ψ. But what is rationality all about if not choosing in ways that make sense—that are *explicable*—in light of your perspective?

Although I will question whether this general picture of rationality is correct in §4.2, there's definitely something attractive about it. And so, for now, let's move on.

3.2 Premise 2: You have a reason to take the Larger box

If you take the Larger box, you'll receive a dollar. Is that a reason you have to take the Larger box over the Regular box?

($\mathsf{r}_{\$_1}$) *If I take the Larger box, I will receive a dollar* $\wedge$ *If I take the Regular box, I will not receive a dollar.*

Clearly, $\mathsf{r}_{\$_1}$ satisfies *Contrastivity*—you'll receive the dollar if you take the Larger box, but not if you take the Regular box. Furthermore,receiving a dollar is, for you, a good thing. A dollar ain't much, but it ain't nothing! Finally, you know all that, and so you're aware that $\mathsf{r}_{\$_1}$ is a reason for you. And so, you have a reason—namely, $\mathsf{r}_{\$_1}$—to take the Larger box rather than the Regular box.

That seems right to me, but Bales et al. (2014) disagree. They argue that **P2** is "unsupportable" because, while there is a perfectly good sense in which you

---

and no reason to take the other-rather-than-the-one, then it is rationally impermissible to take the other."

[9] This is a popular (see, e.g., Foot, 1972; Gibbons, 2010; Kiesewetter, 2017; Lord, 2018; Schroeder, 2009; Williams, 1979; Worsnip, 2021) but by no means universally held position (for dissent see, Broome, 2007, 2013; Scanlon, 1998; Setiya, 2004). Importantly, Hare accepts it (e.g., Hare, 2013, p. 47).

might prefer *I will receive a dollar* to *I will not receive a dollar*, in their estimation, it's implausible that such a preference "can generate 'reasons' that can reliably be plugged into the [*Reasons Principle*]—or at best question-begging," (Bales et al., 2014, p. 18). Their argument proceeds by reflecting on two different accounts of what it is to prefer p to $\neg p$ (where p is a sub-maximal state of affairs, like receiving a dollar), and arguing that, in either case, **P2** is dubious. Here's what they say (quoted in full):

> One obvious hypothesis is that I prefer p being true to p being false if and only if I prefer any world in which p is the case to any world in which p is not the case. But on that reading, [**P2**] is false. I do not prefer that I have an additional dollar and A rather than B alone, nor do I prefer that I have an additional dollar and B rather than A alone. Thus, either way, it would not be better if I had an additional dollar, and the additional dollar does not thus provide me with a reason to take the [Larger] box rather than the [Regular] box.
>
> A second hypothesis is that I prefer p being true to p being false if and only if, *other things being equal*, I prefer p being the case to p not being the case. That does seem to correctly describe the phenomenon in [Vacation Boxes]. Other things being equal, I do prefer it to be the case that I have an additional dollar than otherwise. But this account of preference is so weak that it seems inappropriate to use it in this dialectical context, where we are trying to use it in conjunction with Hare's [*Reason Principle*] to derive a conclusion about what we ought to do. On this understanding of preference, it follows that I have a reason to take the [Larger] box because, other things being equal, I prefer to have an additional dollar rather than not. But it is also the case that *I know that other things are not equal*. Indeed, I know that I will only obtain the additional dollar at the cost of forgoing a good of great value to me. So to think that preferences, in this sense, can generate 'reasons' that can reliably be plugged into the [*Reason Principle*] is implausible—or at best question-begging. (Bales et al., 2014, p. 17-8)

I avoided characterizing *Positivity* in terms of preference—but, if we had to, I think we should favor the second hypothesis.[10] Bales et al. (2014) are correct that

[10] And why only consider these two hypotheses? Aren't there other ways to analyze preferring p to $\neg p$? Here's a different (initially plausible but ultimately unpromising) hypothesis: you prefer p to $\neg p$ if and only if $V(p) > V(\neg p)$, where $V(X) = \sum_w C(w \mid X) \cdot V(w)$. According to this hypothesis, you prefer p to $\neg p$ just in case, given your beliefs, you'd welcome the news that p over the news that $\neg p$. But this will not do. In order for V to be well-defined, we must be able to assign

the first hypothesis—according to which you prefer p to $\neg p$ if and only if you prefer *every* p-world to *every* $\neg p$-world—is implausible. Of course I don't prefer receiving a dollar and being tortured forever to foregoing the dollar but avoiding the torture, but that doesn't mean that I don't prefer having a dollar to not having one. You needn't prefer p at its *worst* to $\neg p$ at its *best* in order to count as preferring p to $\neg p$.

It's not entirely clear what it is to prefer p to $\neg p$ "*other things equal*"—but (as Bales et al. (2014) concede) it does seem that, other things equal, you do prefer having an additional dollar to not.[11] The worry, instead, is that you know that other things are not equal. This, though, doesn't threaten the claim that receiving a dollar is *a reason*. Consider, for example, a choice between $A^{+\$1}$ and B. If you opt for the former, you'll get a dollar, but not if you opt for the latter. You know that other things are not equal—you don't prefer $A^{+\$1}$ to B. But that doesn't mean that the dollar isn't *a reason* to take $A^{+\$1}$ over B. It isn't a *decisive* reason, of course, but it's a reason nonetheless.

Instead, Bales et al. (2014) are better understood as objecting to **P1**. Receiving a dollar *is* a reason to take the one box over the other, but—because you know other things are not equal—this reason doesn't generate the obligation to take it. Notice that, when choosing between $A^{+\$1}$ and B, receiving the dollar isn't a decisive reason because you also have reasons to take B instead. But, according to **P3**, in Vacation boxes, you have *no* reason to take the Regular box over the Larger box. And, if that's right, all of the reasons you have tell in favor of taking the Larger box—so, isn't that what rationality requires you do?

---

numbers to represent your preferences over worlds. But, if your preferences are *incomplete*—as we are assuming they are in this case—we cannot represent your preferences in this way.

[11] Here's a formal way to analyze preferring p to $\neg p$ all else equal. For each p-world, w, find the $\neg p$-world that's most similar: $f_{\neg p}(w)$. If, for every $w \in p$, $w \succ f_{\neg p}(w)$, you prefer p to $\neg p$ other things equal. This passes a lot of the work on to the similarity-function, and it's not obvious what worlds are most similar to others. For example, suppose that if you had an additional dollar, you'd use it to buy a candy bar. Is the most similar world in which you don't receive the additional dollar a world in which you, nevertheless, have the candy bar anyway? Off hand, a world in which you have a candy bar is less similar to one in which you don't. But such a comparison doesn't seem to be in the spirit of the view. So, instead, perhaps we should compare p-worlds to their $\neg p$-world counterpart that is most similar *up to the point at which p becomes relevant*, but allow the worlds to diverge in all sorts of ways that are "downstream" of p (e.g., we should compare w to a world that is just like it up to the point at which you receive an additional dollar, but that differs from w from that point onward). Even still, this analysis seems to set too high of bar as well. Logical space is vast, and there are surely regions of it in which receiving a dollar would be a bad thing—e.g., suppose there's an eccentric madman on the loose who's out for anyone who just received a dollar. But why would such situations be relevant if you know, or are sufficiently confident, that you're not in one?

3.3 Premise 3: You have no reason to take the Regular box

Is it true that you have *no* reason to take the Regular box over the Larger box? I will argue that it is not: you *do* have a reason to take the Regular box over the Larger box. But, first, why think that **P3** is true? The idea is that—although *there are* reasons to take the Regular box over the Larger box—because you don't know which prize is in which box, you fail to *have* them.

It'll be instructive to look at some candidate considerations, and demonstrate why they fail to be reasons you have for taking the Regular box over the Larger one. Consider, for example:

($\mathsf{r}_{\Diamond A}$) *If I take the Regular box, I might get A* $\wedge$ *If I take the Larger box, it's not the case that I might get A.*

$\mathsf{r}_{\Diamond A}$ isn't a reason you have to take the Regular box because you know it isn't true. It's true that, if you take the Regular box, you might get A. But you also might get A if you take the Larger box. The consideration fails the *Contrastivity* condition. Likewise for "I might get B," "I know there's a 50% chance of getting A", "I know I would get a uniquely good prize," "I know the prize I'd get won't be worse than the one I'd get otherwise," etc. None of these considerations distinguish between the two boxes, and so none of them satisfy *Contrastivity*, and so none of them are reasons to take the Regular box *rather than* the Larger box.

There are, of course, features that distinguish between the two. But these features, given what you care about, aren't *reasons* for you to take the one box over the other. For example, consider:

(r_{reg}) *If I take the Regular box, I will get a prize that was inside a regular-sized box* $\wedge$ *If I take the Larger box, I will not get a prize that was insider a regular-sized box.*

r_{reg} is true, but—because you care about the prizes and not the sizes of the boxes they come in—it isn't a reason to take the Regular box. It fails the *Positivity* condition. For you, getting a prize that was inside a regular-sized box isn't a good thing; it's neither here nor there.

Furthermore, while *there are* reasons for taking the Regular box over the Larger box, these are not reasons that you *have*. For example, if, unbeknownst to you, the Regular box contains prize B, then all the uniquely valuable things about B—e.g., the way the sand will feel beneath your toes, the soothing song of the ocean's waves lapping at the shore, the bottomless Mai Tais at the Tiki Bar, etc.—are reasons to take the Regular box over the Larger one. But because you don't know that B is the prize that's in the Regular box, while these *are* reasons to take the Regular box over the Larger box, they aren't reasons you *have*: you aren't in a position to be aware that they are reasons for you. And so these reasons fail the

Awareness condition. You, of course, are aware that *there are* reasons for you take the Regular box over the Larger box, but these are not reasons you *have*.[12]

And so: you have *no* reasons for taking the Regular box over the Larger box.

I disagree. I think you *do* have a reason to take the Regular box over the Larger box—although one that is, perhaps, not particularly compelling. Let's introduce a name 'the R-prize' that, by stipulation, picks out whichever of the two prizes—A or B—is, as a matter of actual fact, in the Regular box.[13] So, for example, if the coin landed heads, 'the R-prize' rigidly designates prize B; and, if the coin landed tails, 'the R-prize' rigidly designates prize A. A term *rigidly designates* just in case it picks out the same thing in all possible worlds in which that thing exists. (So, if the coin actually landed heads, 'the R-prize' refers to B—even regarding the counterfactual situation in which the coin landed the other way. So, it would be true to say, "Had the coin landed the other way, the R-prize would, instead, be in the Larger box!")

Here, then, is a reason you have to take the Regular box over the Larger box:

(r_R) *If I take the Regular box, I will get the R-prize* $\wedge$ *If I take the Larger box, I will not get the R-prize*

Given what 'the R-prize' denotes, r_R is true: if you take the Regular box, you'll get prize B, but if you take the Larger box, you won't. And getting the R-prize is a good thing—it's a *prize* after all. Finally, because you've fixed the reference of 'the R-prize' in such a way that you are in a position to know that that's what's in the Regular box, you are aware that r_R is true and is a reason for you. And so, you *do* have a reason—namely, r_R—to take the Regular box over the Larger box.[14] **P3** is false.

---

[12] It doesn't follow from your knowing that *there are* reasons to φ over ψ that you thereby *have* a reason to φ rather than ψ. If you know that there *is* a reason to φ rather than ψ, then it's true that, for every possible reason, r_1, r_2, …, r_n, …, etc., you know that, either, r_1 is a reason to φ over ψ, or r_2 is a reason to φ over ψ, or …. But a disjunction of the form $\ulcorner \mathsf{r}_n$ is a reason to φ over $\psi \urcorner$ is not *itself* a reason to φ rather than ψ. This is because, while r_n might be a reason for you to φ over ψ, the proposition that r_n *is a reason for you to* φ *over* ψ is not a reason for you to φ over ψ. It fails the *Contrastivity* condition. If r_n is a reason for you to φ rather than ψ, this will be so whether or not you φ.

[13] Mahtani (2017) makes a similar move in order to argue that the notion of *ex ante* pareto superiority is ill-defined—it delivers different verdicts depending on how the people in question are designated (see Mahtani, 2021, for further discussion). In a review of Hare's *The Limits of Kindness*, Harman (2015) raises a similar concern: whether an action is in someone's expected interest can depend on how we've referred to them. Hare (2016, p. 470-1) responds with a version of the 'No Reason' argument. (See also Setiya, 2020, 2023, for further discussion.) The move is inspired by Gareth Evan's classic discussion of "descriptive names" (e.g., 'Julius' the inventor of the zip, Evans, 1982, p. 31).

[14] One might worry that whether r_R is genuinely a reason depends on some controversial assumptions about singular thought. In particular, according to some views, because 'the R-prize' is a descriptively introduced term, r_R fails to express a graspable singular thought. Instead, on these views, one is able to think singularly about an object only if one is suitably *acquainted* with

4 The 'More Reason' Argument

I've argued that, as stated, Hare's 'No Reason' argument is unsound—**P3** is false. But there's an argument with the same spirit that lives on. That argument goes like this: although you have *a* reason to take the Regular box over the Larger box, you nevertheless have *more* reason to take the Larger box instead; and rationality is about doing what you have the most reason to do. Call this the More Reason Argument.

More Reason Argument

P1*	*More Reason Principle:* If you have more reason to φ than to ψ, it's irrational to ψ.
P4	In Vacation Boxes, you have more reason to take the Larger box than to take the Regular box.
C	In Vacation Boxes, it's irrational to take the Regular box.

This is a valid argument. Is it sound? Let's look at each premise.

4.1 Premise 4: You have more reason to take the Larger box

I've claimed that you *do* have a reason to take the Regular box over the Larger box—you know that you'll receive the R-prize. However, there is a perfectly

the object—either via direct perception, memory, or reliable chains of communication (see, e.g., Evans, 1982; Recanati, 1993; Russell, 1910; Soames, 2005). Consequently, $\mathtt{r}_R$ either fails to express a singular thought at all or it expresses one that you aren't able to properly grasp. These "acquaintance" theories, however, struggle to capture the range of cases in which, pretheoretically, singular thought seems possible (Jeshion, 2002, 2010). Other views, like *Semantic Instrumentalism*, hold that we can generate singular thoughts fairly easily—e.g., by simply introducing a term of the right syntactic type, and fixing its referent with an attributive use of a definite description (see, e.g., Borg, 2007; Harman, 1977; Kaplan, 1979, 1989). But these views, too, are controversial—widely regarded as far too permissive. (In fact, almost no one holds them anymore.)

However, one needn't be trapped between these two extremes. Views like Robin Jeshion's *Cognitivism*, for example, chart a middle path. On that view, "significance"—an object's integration into our goals, plans, emotions, and broader cognitive framework—plays a crucial role in shaping our capacity for singular thought. One can think singularly about an object insofar as one has opened a "mental file" on that object. Jeshion (2010), argues—drawing on empirical work from cognitive science—that whether one has a mental file on an object depends on that object's significance to our goals, plans, emotions, etc. We can't think singularly about just anything, but standing in some special relation of acquaintance isn't necessary either.

My claim that **P3** is false appears to rely on the falsity of *Acquaintance Theory* (or, at least, some of its more stringent versions). I think there are independent reasons to reject these views (see, e.g., Jeshion, 2002). That said, I am not unsympathetic to the idea that *acquaintance* does play an important role in—if not whether $\mathtt{r}_R$ is a reason at all—its significance as a reason. I will discuss this in more detail in section §4.1. I'd like to thank an anonymous reviewer for pressing this worry.

analogous reason for taking the Larger box. Let 'the L-prize' rigidly designate whichever prize—A or B—is, as a matter of actual fact, in the Larger box. (So, for example, if the coin landed heads, 'the L-prize' rigidly designates prize A; and, if it landed tails, 'the L-prize' designates prize B.) Consider the following reason:

(r_L) *If I take the Larger box, I will get the L-prize* $\wedge$ *If I take the Regular box, I will not get the L-prize*

Given what 'the L-prize' denotes, r_L is true. For you, getting the L-prize is a good thing. And, you are in a position to know both of these things. Consequently, r_L is a reason for you to take the Larger box over the Regular one.

This means that you have a reason to take the Regular box over the Larger box, and that you have a reason to take the Larger box over the Regular box. But do you have *more* reason to take the Larger box over the Regular box? It's true that you have, in a sense, *two* reasons to take the Larger box over the Regular box (i.e., r_L and $\mathsf{r}_{\$1}$), and only one reason to take the Regular box over the Larger box (i.e., r_R). But reasons are to be *weighed*, not *counted*. And it's not obvious that the former reasons *outweigh* the latter.

Here's an example to bring this out. Consider a choice between $\mathrm{A}^{+\$1}$ and B. You have reasons to take B over $\mathrm{A}^{+\$1}$ (e.g., the way the sand will feel beneath your toes, the soothing song of the ocean's waves lapping at the shore, bottomless Mai Tais at the Tiki Bar, etc.). But you also have reasons to prefer $\mathrm{A}^{+\$1}$ to B (e.g., the crisp mountain air, the adrenaline rush of ripping down a black diamond, bottomless hot chocolate at the pine lodge, etc.). Furthermore, you have an additional reason to take $\mathrm{A}^{+\$1}$—you'll get a dollar. But it's not true that you have *more* reason to take $\mathrm{A}^{+\$1}$ over B than to take B over $\mathrm{A}^{+\$1}$. On balance, your reasons do not weigh more heavily toward one rather than the other.

Intuitively, the dollar doesn't make the difference because the reasons that favor B over A are not *perfectly balanced* by—or, exactly equally as strong as—the various reasons that favor A over B. If they were, adding that extra dollar would make a difference—you would then have *more* reason to take $\mathrm{A}^{+\$1}$ than B. Instead, the reasons you have to choose B over A are neither stronger than, weaker than, nor perfectly balanced by the reasons you have to choose A over B.[15]

Let's say that a reason is *incommensurable* with another just in case (i) neither is stronger than the other, and (ii) it needn't be a mistake to act on the basis of

[15] In other words, on this view, 'at least as much reason as' is *incomplete*. This is controversial. For example, Dorr et al. (2023) have argued that *all* comparative expressions ('*F*-er than', 'more *F* than', 'equally *F* as') are *comparable*: For any two *F*-assessable *x* and *y*: either *x* is at least as *F* as *y* or *y* is at least as *F* as *x*. (In a companion paper, Dorr et al. (2021) argue that, if all comparatives are comparable, rational preferences are complete.) Although I won't address it here, I'm not convinced by their argument. If you are, though, the argument in §4.1 won't be convincing. The argument in §4.2, however, doesn't rely on this assumption.

the one while in full recognition of the other. The reasons you have that favor B over A are incommensurable with the reasons you have that favor A over B. Alternatively,

Perfectly Balanced Reasons: A reason *perfectly balances* another just in case (i) neither is stronger than the other, and (ii) it would be irrational to act on the basis of the one while in full recognition of the other.

In assessing **P4**, the question then is whether r_R and r_L are perfectly balanced or incommensurable. A strong case can be made for the former (and that's probably correct). But, for the sake of argument, I will try to defend the position that they are, in fact, incommensurable.

Let me say more about what it is to act *on the basis* of a reason, and about how it could be irrational to do so while in full recognition of another reason. You act *on the basis* of a reason, r, when it is r in particular that motivates you to act as you do. You act on the basis of a reason, r, while in full recognition of another, r', when it is r that motivates you to act as you do, but you at the same time know that r' is a reason to act differently. In some cases, there are reasons that it seems *irrational* to act on the basis of while in full recognition of the another. Here's an example.

Suppose you are to choose between, not bales of hay, but two envelopes, each containing exactly one dollar bill. There's nothing particularly noteworthy about either—they are normal envelopes containing normal dollars. The envelope on the left contains a dollar bill with serial number `MD11114377-B`, while the one on the right contains a dollar bill with the serial number `PA19544444-A`. You know all this, and must choose which envelope to take home. Here's a reason to take the envelope on the left:

($\mathsf{r}^1_{\$}$) *If I take the envelope on the left, I will get a dollar with serial number* `MD11114377-B` $\wedge$ *If I take the envelope on the right, I will not get a dollar with serial number* `MD11114377-B`.

$\mathsf{r}^1_{\$}$ satisfies the three conditions—it distinguishes between the two options, getting a dollar (with serial number `MD11114377-B`) is a good thing, and you are aware of these things—and, so, it is a reason you have to take the left envelope over the right.

Here's a reason to take the envelope on the right:

($\mathsf{r}^2_{\$}$) *If I take the envelope on the right, I will get a dollar with serial number* `PA19544444-A` $\wedge$ *If I take the envelope on the left, I will not get a dollar with serial number* `PA19544444-A`.

This too satisfies the three conditions, and so is a reason to take the right envelope over the left.

Suppose you decide to take the envelope on the left. And suppose the *reason* you do so is $r^1_\$$—that's the basis on which you act. You take $r^1_\$$ to *justify* your decision to take the left envelope over the right; it's what motivates you to do as you do. The claim is that this is *irrational*—at least insofar as you are in full recognition of $r^2_\$$. You'd like a dollar, but you don't care about their respective serial numbers. You don't regard either dollar to be better or worse than the other. And there's something irrational about being moved by $r^1_\$$ when you fully recognize that $r^2_\$$ is a perfectly good reason—one that is perfectly symmetric—that points in the other direction.

The claim *isn't* that it's irrational to take the left envelope. You should take one of them, and the one on the left is as good as the one on the right. The claim is that it's irrational to be motivated by $r^1_\$$ (when in full recognition of $r^2_\$$). It's not clear that there are any reasons that could rationally motivate you in this case—perhaps all you can do is *pick* rather than *choose* (e.g., Ullmann-Margalit and Morgenbesser, 1977). But even if there were good reasons to choose the left envelope (e.g., if it is red, and that's your favorite color), $r^1_\$$ is not one of them if you're also aware of $r^2_\$$. The two perfectly balance each other.

On the other hand, consider the choice between prize A and prize B. There are plenty of reasons to choose A over B, and there are plenty of reasons to choose B over A. Here are two examples:

(r_A) *If I take A, I will get to enjoy the crisp mountain air* $\wedge$ *If I take B, I will not get to enjoy the crisp mountain air.*

(r_B) *If I take B, I will get to enjoy the sand beneath my toes* $\wedge$ *If I take A, I will not get to enjoy the sand beneath my toes.*

Suppose you are moved by r_A (while in full recognition of r_B) to choose A. (Furthermore, suppose—as we've been assuming—that the set of considerations that favor one of the vacations isn't stronger than the set of considerations favoring the other. And that, in particular, neither r_A nor r_B is stronger than the other.) Would this be irrational?

Here's a picture, inspired by Ruth Chang's account of practical normativity (viz. Chang, 2009, 2013, 2017), according to which it wouldn't be. On Chang's view (which she calls "hybrid voluntarism"), when your "given reasons" run out (as we assume they have), "you have the normative power to create new will-based reasons for one option over another by putting your agency behind some feature of one of the options" (Chang, 2017, p. 16). The feature highlighted in r_A—the enjoyment of the crisp mountain air—is the kind of thing that one can 'stand behind' and, in so doing, use one's normative powers to create for oneself a decisive

reason to choose A.

She says,

> By putting your will behind a feature of an option—by standing for it—*you* can be that in virtue of which something is a will-based reason for choosing that option.
>
> …You have the normative power to create for yourself a new will-based reason to pursue one option over the other. And you may now have most all-things-considered reasons—that is, taking into account both given and will-based reasons—to choose one option over the other. Distinctive of Hybrid Voluntarism is the claim that agents can *make it true* that they have most reason to do one thing rather than another directly through an activity of their own wills. (Chang, 2017, p. 16-17)

But—the thought goes—one cannot put one's agency behind just any feature (at least, not *rationally*). For example, one cannot rationally "stand behind" a feature that one, at the same time, considers irrelevant or disvaluable. But, for you, the enjoyment of crisp mountain air is not like this—it's a good feature of A. Furthermore, although you fully recognize that the feature highlighted in r_B—the enjoyment of the sand beneath your toes—is *also* a feature you could put your agency behind, that doesn't impede your ability to commit to the crispness of the mountain air instead. And so, it's not irrational to act on the basis of r_A—by putting your agency behind that feature—even in full recognition of r_B. Not so, though, with $\mathsf{r}^1_\$$ and $\mathsf{r}^2_\$$: getting a dollar with a particular serial number is *not* a feature you can *stand behind*—if you fully recognize that the other option also promises a dollar, albeit with a different serial number, but that serial numbers aren't something you care about.

The question, then, is: Are r_R and r_L more like r_A and r_B, or more like $\mathsf{r}^1_\$$ and $\mathsf{r}^2_\$$? Is the feature highlighted in r_R—the enjoyment of the R-prize—something you can *stand behind* while in full recognition of r_L?

Proponents of the More Reason Argument think: No, the enjoyment of the R-prize is *not* a feature you can rationally stand behind (while also appreciating that you could enjoy the L-prize instead). And so, it would be irrational to act on the basis of r_R while in full recognition of r_L. And so, because neither reason is stronger than the other, they perfectly balance. And thus, because $\mathsf{r}_{\$_1}$ is an *additional* reason you have for taking the Larger box, you have more reason to take the Larger box than to take the Regular box.

Why is the enjoyment of the R-prize not a feature you can rationally stand behind? The thought, I take it, is that there's an important sense in which you

don't know what the R-prize is.[16] Because you don't know what it is, you aren't in a position to focus in on any of the distinctively valuable features of the R-prize. Although you know it has such features, you don't know what they are! And, the thought goes, you can only "put your agency behind" features you know about.

That's probably right. But for the sake of argument, let me offer two responses. First, the argument turns on the following constraint:

Known Features: When choosing between two options, it can be rational to put your agency behind a feature of one *only if* you know what that feature is.

It's not clear *what it is* to "know what a feature is". Saying what it takes to *know that p* is hard enough; saying what it takes to *know what p is* is even harder. So, while it might be right that, in order to rationally stand behind some feature, you must stand in the right sort of epistemic relationship to that feature, it isn't obvious what that epistemic relationship is. To pick a term, let's say that you must be *acquainted* with the feature.

This raises an interesting question about how acquainted with a feature one must be in order to put one's agency behind it. Suppose you've never been to the beach and that you've never been skiing. You've never felt the sand beneath your toes or the crisp mountain air in your lungs. Would it, then, be a rational mistake to take B over A on the grounds that there'll be sand at the beach but not on the slopes? Or, perhaps you've experienced both, a long time ago, and can now no longer remember what they're like. Must one be able to vividly imagine the feature in order to stand behind it?

Perhaps. But we risk setting the bar far too high. For, even if I've been on many ski trips before, I don't know *precisely* how *this* one will feel. Even in choosing between an apple and a banana, when I've had both many, many times, there is a range of plausible apple-eating experiences I might have if I choose the apple, and I am not in a position to know exactly which one—if any—I would have. Nevertheless, can't I stand behind the way the apple tastes (without knowing *exactly* what that will be)?

Knowing *exactly* what a feature is like can't be necessary. But surely you must know *something* about the feature—to be acquainted with it in some way—in order to rationally stand behind it.[17] In our case, though, you do know *something* about the R-prize. You know that it has distinctively valuable features, and that

[16] Regarding a similar example, Hare (2016) says: "You do not have a reason [to φ] until you *recognize* the way in which it is better if you [φ]. And just dubbing it ['the R-prize'] will not help you *recognize* it" (p. 471, emphasis added). There's an obvious connection between *recognizing* something and knowing what that something is. Hare thinks that only considerations that can sensibly move you to act are properly reasons. And if you don't know what the R-prize *is*, it cannot sensibly move you to act—and so, it's not a reason you have after all.

[17] If you're not convinced, imagine that you've been abducted by aliens. They take you to their home planet, where you'll live out the rest of your days. They offer you a choice about how to

those features are either the valuable features of A or the valuable features of B. And *those* features are ones you are acquainted with. You don't know whether the R-features are like *this* or like *that*. But when choosing an apple, you also don't know *exactly* how it will taste—instead, there is a *range* of apple-tasting experiences that might result. So, it's not impossible to stand behind a *disjunction* of features that you are acquainted with.

There's a disanalogy between the R-features and the range of apple-tasting experiences, however. The range of apple-tasting experiences that might result from choosing the apple and the range of banana-tasting experiences that might result from choosing the banana are wholly distinct. And so you can place your agency behind the one in full recognition of the other. In contrast, what you know about the R-features are (under that description) the same as what you know about the L-features. And so, even if you don't need to know *exactly* what a feature is like in order to stand behind it, it's not rational to stand behind a feature in full recognition of another if, given your impoverished epistemic perspective, you are unable to distinguish between the two.

But—and this brings us to the second response—we've assumed that *standing behind* a feature is categorical rather than conditional. Perhaps, that needn't be so. Suppose instead that you can put your agency behind a feature in a way that's analogous to making a *conditional commitment*. By putting your agency behind the R-prize, you conditionally commit to standing behind those features that make—as it happens, but currently unbeknownst to you—B distinctively valuable. (Of course, had the coin landed the other way, placing your agency behind what you would call 'the R-prize' would involve conditionally committing to standing behind those features that make *A* distinctively valuable.) As a motivating example, consider the attitude prospective parents take toward their as-of-yet unborn children. No one can know exactly what their child will be like, but one can nevertheless commit to loving their child *whatever they're like*. Analogously, why can't you commit to the R-prize in a similar way? There's a fact of the matter about what the R-prize is like—it's beachy—and, even though you don't know that, you can stand behind it nevertheless by *conditionally* committing to embracing those features, whatever they turn out to be, that make the R-prize distinctively valuable.

---

spend the rest of your life. You can choose to become a Gazingaborp or, alternatively, to become a Mikaelsour. They explain that if you become a Gazingaborp, you'll get plenty of baggilums. And, that if you instead opt to become a Mikaelsour, although you'll have to put up with a fair share of yeakizros, you'll be rewarded with as many farfanudles as you'd like. Would it be a rational mistake to choose to be a Gazingaborp over a Mikaelsour on the grounds that you'll get plenty of baggilums given that you fully recognize that by choosing, instead, to become a Mikaelsour you'll be rewarded with farfanudles? You have no idea what any of these things are—I just made them up!—and, furthermore, it's unlikely that you've encountered any similar experiences on Earth. So, it doesn't seem rational—or even psychologically possible—to *stand behind* becoming a Gazingaborp, given that it is (quite literally) an alien experience to you.

Proponents of the More Reason Argument will surely disagree. If you can put your agency behind the R-prize (perhaps in a *conditional* way), it needn't be irrational to act on the basis of r_R while in full recognition of r_L. And so, the two reasons are *not* perfectly balanced—they are, instead, incommensurable. And, as we already know from reflecting on the choice between $\mathrm{A}^{+\$1}$ and B, mildly sweetening one option with a dollar doesn't necessarily give you more reason to take it. **P4** is cast into doubt.[18]

4.2 The More Reason Principle

Let's, at least for the sake of argument, grant that **P4** is true: In Vacation Boxes, you have more reason to take the Larger box than to take the Regular box. It's true that you have *a* reason to take the Regular box (you'll get the R-prize), but let's suppose—contrary to what I hoped to show in the last section—that reason is perfectly balanced by the fact that, if you take the Larger box, you'll get the L-prize.

I think, nevertheless, that the argument can be resisted by rejecting **P1***. It says,

> *More Reason Principle:* If you have more reason to φ than to ψ, it's irrational to ψ.

Underlying this principle is the idea that rationality is about doing what you have most reason to do. If you have more reason to φ than to ψ, you clearly don't have *most* reason to ψ. And so—if rationality is about doing what you have most reason to do—it would be irrational to ψ.

In turn, *that* idea is a particular realization of the more general thought that rationality consists in correctly responding to reasons—a thought that, while not universally accepted, boasts a fair number of adherents. I am not sure whether that general thought is correct, but will—for the sake of argument—grant that it is. I want to argue that, even so, there is room to reject this *particular* instance of that thought. We can accept that rationality consists in correctly responding to reasons, and yet reject that rationality exclusively consists in correctly responding to the reasons that *you have*.

What you have most reason to do and what rationality requires will often coincide. But I think it's possible for the two to come apart. In particular, when you know that the reasons you have are not all the reasons there are, I think it's

[18] Schoenfield (2014), in discussing Hare's argument, appears to agree: "Reasons interact in complex ways and they don't always add up as one might expect them to," (Schoenfield, 2014, p. 273). I hope to have provided some positive reason to think that, in this case in particular, it's not obvious that your reasons add up in a way that *clearly* favors taking the Larger box over the Regular box.

okay for you to exercise *rational humility*: to defer to what you know about the reasons *there are*.

This is not to endorse the view that rationality consists in doing what *there is* most reason to do. It needn't be irrational to wash down Broome's salmonella-laced fish (Broome, 2007, p. 352) with a hearty swig from Williams' famous glass of petrol (Williams, 1979) if you think that the fish is fine and that the petrol is water. Rationality is about doing what makes sense from your *perspective*, and so your beliefs are surely relevant to what it's rational for you to do. This is what motivates the idea that rationality concerns the reasons you *have*—they're appropriately sensitive to your perspective. And, while I don't deny that the reasons you have are *relevant* to what it's rational for you to do, I want to suggest an alternative picture according to which it needn't be irrational to do something other than what you have most reason to do—if, for example, you know of some specific option that there is sufficient reason to do it instead. Your *beliefs* about what there is sufficient reason to do is, of course, also sensitive to your perspective. And so one can deny the *More Reason Principle* and yet still accept that rationality supervenes on one's perspective.

Let's say that there is *sufficient reason* to φ just in case there isn't more reason to do something else. Consider the following principle, linking rationality to what one knows about the reasons there are.

> *Sufficient Reason Link:* If you know for sure that there is sufficient reason to φ, it's not irrational to φ.

The reasons there are—the so-called 'objective' reasons—bear on what one *objectively* ought to do. They hold a special position of authority. They are what we ideally hope to be guided by. Often, we don't know what there is sufficient objective reason to do. What this principle says is: if you *do* know what there is sufficient reason to do, it wouldn't be irrational to act on that knowledge. Importantly, the principle concerns what you *know for sure*, not merely what you take to be *likely*. It very well might be irrational to turn down a bet that, while likely to lose, promises a stupendous windfall otherwise. But, in that case, although it's likely there's sufficient reason to reject the bet, you do not know this for sure.

In Vacation Boxes, you know for sure that there is sufficient reason to take the Regular box. These reasons—which concern the sun, the sand, and the sea—aren't reasons that you *have* because you don't know which vacation is in which box. Furthermore, it's not that, by recognizing that there is sufficient reason to take the Regular box, you (granting the argument for **P4**) thereby cease to have more reason to take the Larger box. This knowledge doesn't suddenly grant you an additional reason to take the one, nor does it somehow defeat one of the reasons you have for taking the other. According to the *More Reason Principle*, it's

irrational to take the Regular box. But, according to *Sufficient Reason Link*, it is not.

I think *Sufficient Reason Link*—or, at least, something in the ballpark—is correct. Underlying the principle is the following thought.

Rational Humility: Rationality isn't about doing what you *have* reason to do; it's about *aiming* to do what *there is* reason to do.

As mentioned, you can fail at doing what there is reason to do and be rational nonetheless; similarly, you can fail to hit the bullseye you were aiming for. Furthermore, because what there is reason to do is *scalar*—the extents to which reasons favor some things over others come in degrees—aiming at it needn't involve doing whatever you think is *most likely* best.[19] In fact, in some cases, it isn't irrational to do what you know there *isn't* most reason to do. Consider, for example:

> The Miners Puzzle.[20] Ten miners are trapped in either Shaft A or Shaft B, but you don't know which—both are equally likely. As floodwaters rush toward the shafts, you have three options: *block Shaft A*, *block Shaft B*, or *block neither*. If you block the wrong shaft, all ten miners will drown; if you block the correct one, you'll save everyone. But if you block neither, the floodwaters will split between the shafts, resulting in the certain death of exactly one miner. As it happens, the miners are in Shaft A.

It's not irrational to *block neither*—in fact, given your uncertainty, that's arguably what rationality requires you to do. But you know that blocking neither shaft isn't what there is most reason to do. If the miners are in Shaft A, there's most reason to *block Shaft A*; if the miners are in Shaft B, there's most reason to *block Shaft B*. They're either in Shaft A or Shaft B. And so, either way, there isn't most reason to *block neither*—but that's what rationality requires.

Opting to *block neither* is also the best way to aim at doing what there is reason to do—even though you know there isn't *most* reason to do it. Here's why. You don't know which shaft the miners are in. If the miners are in Shaft A, there is

[19] As Sepielli puts it, making a related point: "[T]he best try at doing what has a scalar feature needn't simply be the action that is most likely to have the highest degree of the feature. (For example, the best try at making money on the stock market is not simply to throw all of one's initial investment into high-risk, high-return penny stocks.) Rather, it need only be an action that maximizes some function of the degrees of the scalar feature and the action's probability of having each of those degrees, respectively." (Sepielli, 2012, p. 54)

[20] The example comes from Regan (1980), and has since been much discussed (see, among others, Kolodny and McFarlane, 2010; Muñoz and Spencer, 2021; Parfit, ms). Examples with a similar structure (sometimes called "three option cases") have also been much discussed (e.g., Broome, 2013; Jackson, 1991; Lord, 2018; Ross, 2006; Schroeder, 2009).

more reason to *block Shaft A* than to *block neither*. If the miners are in Shaft B, there is more reason to *block neither* than to *block Shaft A*. But the *extent* to which there might be more reason to *block neither* than to *block Shaft A* (i.e., if the miners are in Shaft B) is significantly larger than the extent to which there might be more reason to *block Shaft A* than to *block neither* (i.e., if they're in Shaft B)—it's the difference between saving nine lives and only saving one. So, given that you regard each possibility as equally likely, from your perspective, blocking neither is a better way to aim at doing what there is reason to do than blocking Shaft A is. *Mutatis mutandis* for blocking Shaft B. And so, out of the three options, from your perspective, *block neither* best aims at what there is reason to do.

Nevertheless, isn't the Miners Puzzle at least a counterexample to *Sufficient Reason Link*? You know there is sufficient reason to block a shaft, and yet it *is* irrational to do so—rationality instead requires you to *block neither*. I think there are independent grounds for questioning whether you *do* know there is sufficient reason to block a shaft, but space precludes fully exploring that route.[21] Instead, mirroring a move in Muñoz and Spencer (2021), let's restrict the principle to choices between *fully specific* options:

> *Weak Sufficient Reason Link:* If you know both (i) that there is sufficient reason to φ, and (ii) that the fully specific thing that you would do were you to do φ is among the options that there is sufficient reason to perform, then it's not irrational to φ.

Muñoz and Spencer (2021, p. 85-6) rightfully complain that such a principle, while perhaps true, fails to apply in a range of important cases. Certainly, more can be done to develop the picture I've sketched into a fuller theory.[22] But here I am content to speak the truth, even if it's not the whole truth.

Why prefer this picture to the one underlying the *More Reason Principle*? The latter, by focusing exclusively on the reasons that you *have*, closely tethers rationality to a particular aspect of one's perspective. And while rationality should of

[21] Muñoz and Spencer (2021, p. 89) argue that there is decisive objective reason to *block a shaft* because they suggest that, like 'ought to __', 'there is decisive reason to __' is *upward monotonic*: if *X* entails *Y*, then 'there is decisive reason to *X*' entails 'there is decisive reason to *Y*'. There is decisive reason to *block Shaft A*, and blocking Shaft A entails blocking a shaft. But, *pace* Muñoz and Spencer (2021), I am not convinced that 'there is decisive reason to __' is upward monotonic. And there is a principled reason to think it isn't. You have decisive reason to *block Shaft A* because, if you do, you'll save Person 10's life, and, if you don't, you won't. Is this also a reason to *block a shaft*? To my ear, 'if I *block a shaft*, I'll save Person 10's life' doesn't sound true. It depends on whether, if you block a shaft, you'll block the right one. And so, even if it is true, it isn't something you're in a position to know. But if that's not a reason to *block a shaft*, what is? There's more to say about this interesting issue, but I don't have the space to pursue it further here.

[22] In Doody (ms), I attempt to provide a full *decision theory* that's consistent with the picture sketched here. Also, although characterized in terms of *value* rather than *reasons*, and focused on different issues, a somewhat similar picture is discussed in Bacon (2022); Williams (2023).

course be sensitive to one's perspective, this picture elevates that aspect of one's current perspective to a degree that is *chauvinistic*. To see this, consider the following addendum to Vacation Boxes: you'll be given a peek inside of the boxes, or you can pay a small sum (e.g., 40¢) to decide without peeking. The picture underlying the *More Reason Principle* would seem to countenance the following reasoning:

> I can *peek* or I can *pay*. Peeking will reveal which prize is where. And so, if I peek, I'll come to have sufficient reason to choose either box. Therefore, because I'll respond to the reasons I have, I might end up selecting B over $A^{+\$1}$ (if the coin landed heads) or A over $B^{+\$1}$ (if the coin landed tails)—and, in fact, I regard either choice as equally likely.[23] And so, from my current perspective, *peeking* is like a 50/50 gamble between A and B *plus an expected 50¢*.[24]
>
> On the other hand, I know that if I pay the 40¢ to avoid learning which prize is in which box, I will continue to have more reason to choose the sweetened option ($L^+-40¢$) than to choose the unsweetened one ($R-40¢$). According to the *More Reason Principle*, then, it would be irrational to take the unsweetened box. Because I think I'll continue to behave rationally, if I pay the 40¢, I'm sure I will take the sweetened box—resulting in $A^{+60¢}$ (if the coin landed heads) or in $B^{+60¢}$ (if it landed tails). And so, from my current perspective, *paying 40¢* is like a 50/50 gamble between A and B *plus 60¢*.[25]

---

[23] Regarding either as *equally likely* is not necessary for the argument to work—but it makes the presentation clearer, and it is a natural assumption. Instead, so long as you think the probability, p that you'd select L^+ over R (upon learning what the boxes contain) is *probabilistically independent* of how the coin landed (and that $p < 1$), we can generate a version of the problem by adjusting the fee so that it's less than $(1 - p) \cdot \$1$. Still, the assumption of probabilistic independence can be questioned: surely, learning which box contains which prize can affect which box you might select! (For example, if you preferred A over B, learning that the coin landed heads is fantastic evidence that you'll select L^+, and learning that it landed tails is fantastic evidence that you'll select R.) But, given that you don't prefer either prize to the other, it's less clear what might justify such a failure of independence. In any case, even if one isn't rationally *required* to reason in the manner I'm suggesting, it's *consistent* with rationality that there be someone—if not you, someone like you—who does. And I think that's problem enough.

[24] Here's why. Given a choice between $A^{+\$1}$ and B, you think it equally likely you'd choose either. And likewise given the choice between $B^{+\$1}$ and A. Furthermore, because you'll face the first choice if the coin landed heads and the second if the coin landed tails, it's 50/50 which of the two choices you'd face. And so, *peeking* affords you a 25% chance of getting $A^{+\$1}$, a 25% chance of getting B, a 25% chance of getting $B^{+\$1}$, and a 25% chance of getting A. Given that you care about the prizes and the money (and not how you come to have them), that reduces to a 50% of getting A and a 50% of getting B plus a 50% chance of getting \$1—i.e., a 50/50 gamble between A and B plus an expected 50¢.

[25] Here's why. If you *pay*, you're 100% certain that you'll take L^+ over R. And so you'll either receive $A^{+60¢}$ (if the coin landed heads) or $B^{+60¢}$ (if it landed tails)—i.e., a 50/50 gamble between A and

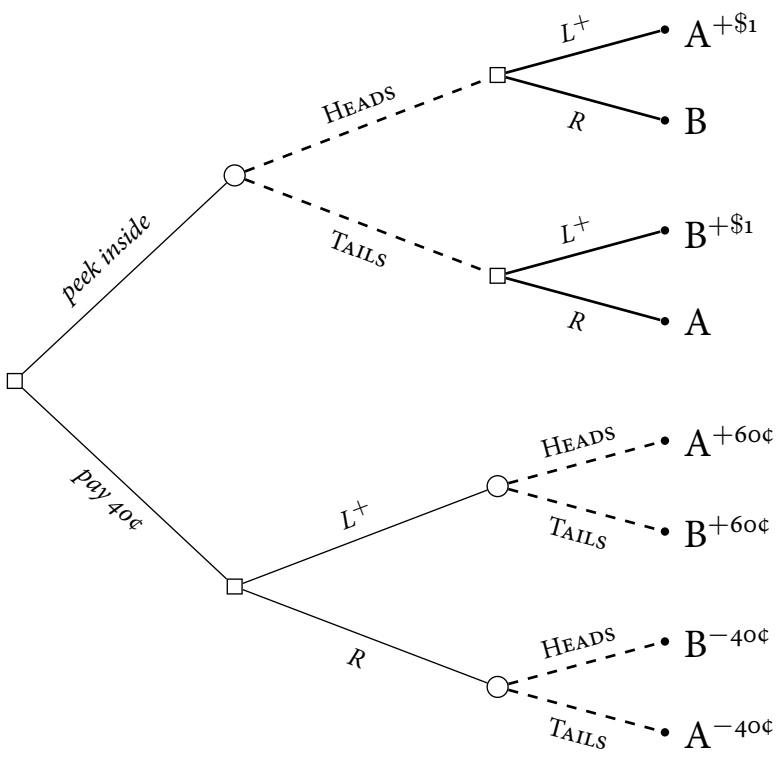

Figure 1: Addendum to Vacation Boxes—to peek or not to peek?

> I have more reason to take a 50/50 gamble between A and B *sweetened with 60¢* than to take one *sweetened with an expected 50¢*.[26] And so, I have more reason to *pay* than to *peek*. Thus, according to the *More Reason Principle*, it's irrational for me to peek inside the boxes.

But that, I claim, is absurd. It's irrational to learn more? You should instead pay 40¢ to remain ignorant—even though doing so is *guaranteed* to make you 40¢ poorer than you could've been?

There's room to resist the argument, of course. (e.g., the reasoning relied on some assumptions about diachronic choice that could be questioned).[27] I won't

B plus 60¢.

[26] This, of course, *needn't* be the case—you might superstitiously hate the number 60, you might love risk (and so prefer risky bets at worse odds to sure-things), etc. But I'll assume that you prefer the sure-thing 60¢ to a 50% chance at $1. If one lacks this preference, a different example (e.g., with a smaller fee) should do the trick anyway.

[27] In particular, the argument assumes that one doesn't have the option of forming a binding resolution to peek-then-take-the-Larger-box come what may. I think it's an interesting question whether a rational agent could form such a resolution—given that they know that, when the time comes, they will have sufficient reason to break it. That said, Hare (2013, Ch. 14) includes an extended discussion of strategies that agents with incomplete preferences may adopt to avoid straying from their resolutions. Whether these strategies work in cases of uncertainty requires a fuller treatment than I can provide here.

offer a full defense of it here, but it doesn't strike me as obviously mistaken.

Alternatively, one could accept the conclusion—it *is* irrational to *peek*—but deny that it's objectionable.[28] After all, it's arguably not objectionable to pay an extra 40¢ for an equally good cup of coffee at a café with less enticing baked goods in order to avoid the potentially-diet-compromising temptations at the bakery. But I think there's an important disanalogy between cases like that and this version of Vacation Boxes. In the former, your 40¢ acts as insurance against a bout of future *irrationality*—you (now, clearheaded) worry that you (then, in the grip of temptation) are at risk of making a *mistake* by splurging on cheesecake croissants.

In Vacation Boxes, however, you don't worry that peeking will lead you to choose *irrationally*—you're certain that you'll always only do what you have most reason to do. Instead, you know that peeking will alter what reasons you have in a predictable way: you'll come to have reasons that support selecting R over L^+. The reasons you currently have tell against coming to have these additional reasons because responding to them rationally, as you know you would, might lead you away from doing what you now have most reason to do. The reasons you currently have are pitted against the ones of your future (better-informed) self—and, according to the picture underlying the *More Reason Principle*, the former should prevail. But I think that's objectionably chauvinistic.

Facts about your current perspective—like, what reasons you have—are obviously relevant to what it's rational to do. But your current perspective isn't exhausted by the reasons you have. Your current perspective also includes what you know about the reasons there are. In this case, you know for sure that there is sufficient reason to *peek*, and you don't know whether there is sufficient reason to *pay*. And so, *Sufficient Reason Link* delivers the more sensible verdict: it's not irrational to *peek*.

I therefore think that, attractive as it appears, the *More Reason Principle* is false, and the 'More Reason' argument fails.[29] One needn't take the sugar after all.

---

[28] Or: one might accept the conclusion and find it objectionable, but trace the fault elsewhere. For example, Gustafsson (2022, p. 24-39) argues that agents with incomplete preferences can be money-pumped—and that, therefore, it's irrational to have incomplete preferences. The problem lies not with the *More Reason Principle* but with the agent's preferences. Still, Gustafsson's case depends on controversial premises I'm inclined to reject. I can't pursue that here.

[29] If the *More Reason Principle* is false, why does it seem true? Here are two thoughts. First, quite often, doing what you have most reason to do coincides with aiming to do what there is reason to do. And so it's easy to run the two together. Second, and more tentatively, *if* you care about being able to easily *justify* your choices to others (under possibly antagonistic conditions), you'll favor actions that can be easily made legible. Citing a clear, contrastive reason (e.g., "I'll get a dollar!") is potentially more compelling than pointing to your knowledge of reasons you can't articulate. There's evidence that decision-makers are sensitive to these kind of social considerations (for example, Barber et al., 2003; Simonson and Nowlis, 2000, found that "reason-based choice" behavior—favoring options with justifications that are easy to articulate—is amplified when decision-makers expect social scrutiny). In fact, according to Sperber and Mercier (2017),

5 Conclusion

I've argued that Hare's 'No Reason' argument for choosing the sweetened option in cases of *opaque sweetening* fails. Its central claim—that there is *no reason* to choose the unsweetened option—is false. The R-prize—the rigidly designated reward in the unsweetened box—provides exactly such a reason.

The paper then examined a modified 'More Reason' version of Hare's argument, which contends that, while there may be *a* reason to take the unsweetened option, it's outweighed by the reasons to take the sweetened option instead. I explore some ways of resisting this claim as well. But, granting it for the sake of argument, I argue that the 'More Reason' version of the argument fails nonetheless. The argument relies on the *More Reason Principle*: roughly, rationality tracks what you have most reason to do. I articulate a competing picture: rationality isn't about doing what you *have* most reason to do, it's about *aiming* to do what *there is* reason to do. You can be moved by your knowledge that *there is* sufficient reason to do something—even though uncertainty prevents you from possessing those reasons. I favor this picture because the *More Reason Principle*, by exclusively focusing on what reasons you *have*, offers objectionably chauvinistic advice—at least when one's choices can predictably alter the balance of reasons that one will come to have.

And so: take the sugar if you wish—but know that, as far as rationality is concerned, its alternative can be just as satisfying.[30]

References

Bacon, Andrew. 2022. Actual Value in Decision Theory. *Analysis*, 82(4): 617–629. URL https://doi.org/10.1093/analys/anac014

Bader, Ralf. 2018. Stochastic Dominance and Opaque Sweetening. *Australasian Journal of Philosophy*, 96(3): 498–507. URL https://doi.org/10.1080/00048402.2017.1362566

Bales, A., D. Cohen, and Toby Handfield. 2014. Decision Theory of Agents with Incomplete Preferences. URL 10.1080/00048402.2013.843576

Barber, Brad M., Chip Heath, and Terrance Odean. 2003. Good Reasons Sell: Reason-Based Choice Among Group and Individual Investors in the Stock

this is precisely the primary function "reason" evolved to have—to convince others. If that's right, we should be careful not to confuse rationality with ease of interpersonal justification.

[30] Thanks is owed to Ralf Bader, Caspar Hare, Simone Gubler, Patrick McKee, Wlodek Rabinowicz, participants of the Value Additivity Workshop at Lund University, and the students in my fall 2023 Puzzles of Rationality course. Thanks as well to two anonymous reviewers.

Market. *Management Science*, 49(12): 1636–1652. URL https://doi.org/10.1287/mnsc.49.12.1636.25109

Borg, Emma. 2007. *Minimal Semantics* (Oxford: Oxford University Press)

Broome, John. 2007. Does Rationality Consist in Responding Correctly to Reasons? *Journal of Moral Philosophy*, 4(3): 349–374. URL 10.1177/1740468107083249

——. 2013. *Rationality Through Reasoning* (Wiley-Blackwell)

Chang, Ruth. 1997. Introduction. In *Incommensurability, Incomparability, and Practical Reason* (Harvard University Press)

——. 2002. The Possibility of Parity. *Ethics*, 112(4): 659–688. URL https://doi.org/10.1086/339673

——. 2009. Voluntarist Reasons and the Sources of Normativity. In *Reasons for Action*, edited by D. Sobel and D. Wall (Oxford University Press), pp. 243–271

——. 2013. Grounding practical normativity: going hybrid. *Philosophical Studies*, 164(1): 163–187. URL https://doi.org/10.1007/s11098-013-0092-z

——. 2017. Hard Choices. *APA Journal of Philosophy*, 3(1): 1–21. URL https://doi.org/10.1017/apa.2017.7

Doody, Ryan. 2019a. Opaque Sweetening and Transitivity. *Australasian Journal of Philosophy*, 97(3): 559–571. URL https://doi.org/10.1080/00048402.2018.1520269

——. 2019b. Parity, Prospects, and Predominance. *Philosophical Studies*, 176(4): 1077–1095. URL https://doi.org/10.1007/s11098-018-1048-0

——. 2021. Hard Choices Made Harder. In *Value Incommensurability: Ethics, Risk, and Decision-Making*, edited by H. Andersson and A. Herlitz (Routledge), pp. 247–266. URL 10.4324/9781003148012-17

——. ms. Actual Value Decision Theory. Unpublished manuscript

Dorr, Cian, Jacob M. Nebel, and Jake Zuehl. 2021. Consequences of Comparability. *Philosophical Perspectives*, 35(1): 70–98. URL https://doi.org/10.1111/phpe.12157

——. 2023. The Case for Comparability. *Noûs*, 57(2): 414–453. URL https://doi.org/10.1111/nous.12407

Evans, Gareth. 1982. *The Varieties of Reference* (Oxford: Clarendon Press)

Fogal, Daniel and Alex Worsnip. 2021. Which reasons? Which rationality? *Ergo, An Open Access Journal of Philosophy*, 8(11): 306–343. URL https://doi.org/10.3998/ergo.1148

Foot, Philippa. 1972. Morality as a System of Hypothetical Imperatives. *Philosophical Review*, 81(3): 305–316. URL https://doi.org/10.2307/2184328

Gibbons, John. 2010. Things That Make Things Reasonable. *Philosophy and Phenomenological Research*, 81(2): 335–361. URL https://doi.org/10.1111/j.1933-1592.2010.00373.x

Grant, Lyndal and Milo Phillips-Brown. 2020. Getting what you want. *Philosophical Studies*, 177: 1791–1810. URL https://doi.org/10.1007/s11098-019-01285-1

Gustafsson, Johan E. 2022. *Money-Pump Arguments* (Cambridge: Cambridge University Press)

Hare, Caspar. 2010. Take the Sugar. *Analysis*, 70(2): 237–247. URL https://doi.org/10.1093/analys/anp174

——. 2013. *The Limits of Kindness* (Oxford University Press)

——. 2016. Should We Wish Well to All? *The Philosophical Review*, 125(4): 451–472. URL https://doi.org/10.1215/00318108-3624764

Harman, Elizabeth. 2015. Review: The Limits of Kindness by Caspar Hare. *Ethics*, 125(3): 868–872

Harman, Gil. 1977. How to Use Propositions. *American Philosophical Quarterly*, 14: 173–776

Jackson, Frank. 1991. Decision-Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics*, 101: 461–82. URL https://doi.org/10.1086/293312

Jeshion, Robin. 2002. Acquaintanceless *De Re* Belief. In *Meaning and Truth: Investigations in Philosophical Semantics*, edited by Joseph Keim-Campbell, Michael O'Rourke, and David Shier (Seven Bridges Press), pp. 53–78

——. 2010. Singular Thought: Acquaintance, Semantic Instrumentalism, and Cognitivism. In *New Essays on Singular Thought*, edited by Robin Jeshion, chap. 4 (Oxford University Press), pp. 105–140

Kaplan, David. 1979. Dthat. In *Contemporary Perspectives in the Philosophy of Language*, edited by Peter A. French, Jr. Theodore E. Uehling, and Howard K. Wettstein (University of Minnesota Press), pp. 383–400

——. 1989. Demonstratives. In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein (New York: Oxford University Press), pp. 481–563

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

Kolodny, Nico and John McFarlane. 2010. Ifs and Oughts. *The Journal of Philosophy*, 108(3): 115–143. URL https://doi.org/10.5840/jphil2010107310

Lewis, David. 1986. *On the Plurality of Worlds* (Blackwell)

Lord, Errol. 2010. Having Reasons and the Factoring Account. *Philosophical Studies*, 149(3): 283–296. URL 10.1007/sl1098-009-9359-9

——. 2018. *The Importance of Being Rational* (Oxford: Oxford University Press)

Mahtani, Anna. 2017. The Ex Ante Pareto Principle. *Journal of Philosophy*, 114(6): 303–323. URL https://doi.org/10.5840/jphil2017114622

——. 2021. Frege's Puzzle and the Ex Ante Pareto Principle. *Philosophical Studies*, 178: 2077–2100. URL https://doi.org/10.1007/s11098-020-01524-w

Markovits, Julia. 2014. *Moral Reasons* (Oxford: Oxford University Press)

Muñoz, Daniel and Jack Spencer. 2021. Knowledge of Objective 'Oughts': Monotonicity and the New Miners Puzzle. *Philosophy and Phenomenological Research*, 103: 77–91. URL 10.1111/phpr.12702

Parfit, Derek. 1984. *Reasons and Persons* (Oxford: Clarendon Press)

——. 2011. *On What Matters*, vol. 2 (Oxford: Oxford University Press)

——. ms. What we together do. Unpublished manuscript

Phillips-Brown, Milo. 2021. What does decision theory have to do with wanting? *Mind*, 130(518): 413–437. URL https://doi.org/10.1093/mind/fzaa057

Raz, Joseph. 1986. *The Morality of Freedom* (Oxford: Clarendon Press)

——. 1999. *Engaging Reason* (Oxford: Oxford University Press)

Recanati, François. 1993. *Direct Reference: From Language to Thought* (Oxford: Basil Blackwell)

Regan, Don. 1980. *Utilitarianism and Cooperation* (Oxford: Oxford University Press)

Ross, Jacob. 2006. *Acceptance and practical reason*. Phd thesis, Rutgers University

Russell, Bertrand. 1910. Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society*, 11: 108–128

Scanlon, T.M. 1998. *What We Owe to Each Other* (Harvard University Press)

Schoenfield, Miriam. 2014. Decision Making in the Face of Parity. *Philosophical Perspectives*, 28(1): 263–277. URL https://doi.org/10.1111/phpe.12044

Schroeder, Mark. 2007. *Slaves of the Passion* (Oxford: Oxford University Press)

——. 2008. Having Reasons. *Philosophical Studies*, 139(1): 57–71. URL https://doi.org/10.1007/s11098-007-9102-3

——. 2009. Means-End Coherence, Stringency, and Subjective Reasons. *Philosophical Studies*, 143(2): 223–48. URL https://doi.org/10.1007/s11098-008-9200-x

Sepielli, Andrew. 2012. Subjective Normativity and Action Guidance. In *Oxford Studies in Normative Ethics*, vol. II, edited by Mark Timmons (Oxford University Press), pp. 45–73

——. 2018. Subjective and Objective Reasons. In *Oxford Handbook of Reasons and Normativity*, edited by Daniel Star (Oxford: Oxford University Press), pp. 784–799

Setiya, Kieran. 2004. Against Internalism. *Noûs*, 38(2): 266–98. URL https://doi.org/10.1111/j.1468-0068.2004.00470.x

——. 2007. *Reasons without Rationalism* (Princeton, NJ: Princeton University Press)

——. 2014. What is a Reason to Act? *Philosophical Studies*, 167: 221–235. URL https://doi.org/10.1007/s11098-012-0086-2

——. 2020. Ignorance, Beneficence, and Rights. *Journal of Moral Philosophy*, 17: 56–74. URL 10.1163/17455243-20182841

——. 2023. Other People. In *Rethinking the Value of Humanity*, edited by Sarah Buss and Nandi Theunissen (Oxford University Press), pp. 314–336

Simonson, Itamar and Stephen Nowlis. 2000. The Role of Explanations and Need for Uniqueness in Consumer Decision Making: Unconventional Choices Based on Reasons. *Journal of Consumer Research*, 27

Smith, Michael. 1987. The Humean Theory of Motivation. *Mind*, 96(381): 36–61. URL https://doi.org/10.1093/mind/XCVI.381.36

——. 1995. Internal Reasons. *Philosophy and Phenomenological Research*, 55(1): 109–131

Soames, Scott. 2005. *Reference and Description* (Princeton, NJ: Princeton University Press)

Sperber, Dan and Hugo Mercier. 2017. *The Enigma of Reason: A New Theory of Human Understanding* (Cambridge, MA: Harvard University Press)

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

Sylvan, Kurt. 2015. What apparent reasons appear to be. *Philosophical Studies*, 172(3): 587–606. URL 10.1007/s11098-014-0320-1

Ullmann-Margalit, Edna and Sidney Morgenbesser. 1977. Picking and Choosing. *Social Research*, 44(4): 757–785

Vogelstein, Eric. 2012. Subjective Reasons. *Ethical Theory and Moral Practice*, 15(2): 239–257. URL https://doi.org/10.1007/s10677-011-9286-7

Way, Jonathan. 2009. Two Accounts of the Normativity of Rationality. *Journal of Ethics and Social Philosophy*, 4(1): 1–8

Williams, Bernard. 1979. Internal and External Reasons. In *Moral Luck* (Cambridge University Press)

Williams, J. Robert G. 2023. Aptness and means-end coherence: a dominance argument for causal decision theory. *Synthese*, 201(48): 1–19. URL https://doi.org/10.1007/s11229-022-04017-x

Wodak, Daniel. 2017. Can Objectivists Account for Subjective Reasons? *Journal of Ethics and Social Philosophy*, 12(3): 259–279. URL https://doi.org/10.26556/jesp.v12i3.246

——. 2019. An Objectivist's Guide to Subjective Reasons. *Res Philosophica*, 96(2): 229–244. URL https://doi.org/10.11612/resphil.1750

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